# MACHINE LEARNING ENABLED DESIGN AND PREDICTIVE MODELLING OF CARBON NANOSTRUCTURED ELECTRODE DERIVED FROM SUSTAINABLE RESOURCES FOR ENERGY STORAGE SYSTEM

**Thesis submitted in partial fulfillment**

**for the Award of Degree**

*Doctor of Philosophy*

by

**STUTI SHRIVAS**

**RAJIV GANDHI INSTITUTE OF PETROLEUM TECHNOLOGY**

**JAIS – 229304**

***Roll No.: 21EE0103***

***2026***

## CERTIFICATE

It is certified that the work contained in the thesis titled ***"MACHINE LEARNING ENABLED DESIGN AND PREDICTIVE MODELLING OF CARBON NANOSTRUCTURED ELECTRODE DERIVED FROM SUSTAINABLE RESOURCES FOR ENERGY STORAGE SYSTEM"*** by ***STUTI SHRIVAS*** has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive, Candidacy, SOTA, and Open Seminar.

Dr. Amarish Dubey
(Thesis Supervisor)

## DECLARATION BY THE CANDIDATE

I, *STUTI SHRIVAS*, certify that the work embodied in this thesis is my own bona fide work and carried out by me under the supervision of *DR. AMARISH DUBEY* from *January 2022* to *January 2026,* at the Rajiv Gandhi Institute of Petroleum Technology, Jais. The matter embodied in this thesis has not been submitted for the award of any other degree. I declare that I have faithfully acknowledged and given credit to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully copied any other's work, paragraphs, text, data, results, etc., reported in journals, books, magazines, reports, dissertations, theses, etc., or available at websites and have not included them in this thesis and have not cited as my own work.

Date:

Place: Jais, Amethi **STUTI SHRIVAS**

## CERTIFICATE BY THE SUPERVISOR

It is certified that the above statement made by the student is correct to the best of my knowledge.

Dr. Amarish Dubey
(Thesis Supervisor)

Head of the Department
(Electrical & Electronics Engineering)

## CERTIFICATE

CERTIFIED that the work contained in the thesis titled "***Machine learning enabled design and predictive modelling of carbon nanostructured electrode derived from sustainable resources for energy storage system"*** by ***Ms. Stuti Shrivas*** has been carried out under my supervision. It is also certified that she fulfilled the mandatory requirement of TWO quality publications that arose out of her thesis work.

It is further certified that the two publications (copies enclosed) of the aforesaid ***Ms. Stuti Shrivas*** have been published in the Journals indexed by –

(a) SCI

(b) SCI Extended

(c) SCOPUS

**Dr. Amarish Dubey**
(Thesis Supervisor)

(Convener, DPGC)

# COPYRIGHT TRANSFER CERTIFICATE

**Title of the Thesis:** Machine learning enabled design and predictive modelling of carbon nanostructured electrode derived from sustainable resources for energy storage system
**Name of the Student:** Stuti Shrivas

## Copyright Transfer

**The undersigned hereby assigns to the Rajiv Gandhi Institute of Petroleum Technology Jais all rights under copyright that may exist in and for the above thesis submitted for the award of the *"DOCTOR OF PHILOSOPHY"*.**

Date:
Place: Jais, Amethi

STUTI SHRIVAS
21EE0103

# ACKNOWLEDGEMENT

Before commencing this thesis, I would like to profoundly thank individuals who have paved the path for my PhD journey with their unwavering support and encouragement. This expedition, with its highs and lows akin to a roller coaster, has been made manageable by the invaluable contributions of numerous individuals, both directly and indirectly involved. First and foremost, I extend my sincerest appreciation to my thesis supervisor, Dr. Amarish Dubey, whose guidance, encouragement, and persistent support have been instrumental in realizing this dissertation. Throughout my doctoral pursuit, Dr. Dubey has provided invaluable insights, constructive suggestions, and meticulous feedback, consistently upholding the highest academic rigor and technical proficiency standards essential for scholarly publication.

Additionally, I am deeply indebted to Prof. Harish Hirani, Director of RGIPT, whose unwavering commitment to fostering a culture of excellence in research has inspired me. My heartfelt thanks also extend to Dr. Shivanshu Srivastava, and Dr. Kalka Dubey for their unwavering support, well wishes, understanding, and motivation. I want to extend my thanks to all the faculty members and RPEC members for their suggestions and to the administrative staff of the Department of Electrical & Electronics Engineering for their help and support.

My parents are equally deserving of my appreciation, whose boundless love, unwavering support, and steadfast belief in my abilities have been a constant source of strength. Furthermore, I am deeply grateful to beloved husband, and son for their unwavering support, encouragement, and sacrifices that have enabled me to pursue my academic aspirations. Lastly, I extend my heartfelt thanks to my friends, Surabhi Verma, Shivangi Srivastava, and other fellow research scholars, Akash, Tanya, Amit and my juniors within the department, whose unwavering support, camaraderie, and cherished memories have been a source of solace and inspiration.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# PREFACE

The increasing demand of energy drives the development of good performance, cost effective, and green energy solutions to reduce the dependence on fossil fuels. Among all energy technologies, energy storage like supercapacitors have extensively explored. However, the materials used for designing efficient energy storage devices play an important role. However, carbon materials have emerged as successful material employed in energy conversion and storage derived from various renewable and non-renewable precursors. For clean energy production, renewable resources derived carbon from natural fiber and biomass, available in abundance in nature, are receiving attention due to environmental impact and cost efficiencies for energy storage devices. In the thesis, review of carbon nanomaterials derived from natural resources-based precursors such as biomass and natural fiber to make electrodes for supercapacitor. Among all energy storage devices, supercapacitors are selected due to their exceptional properties such as high-power density, rapid charge-discharge, and long-life cycle. The device's performance is dependent on their physicochemical and electrochemical properties of their electrode materials. Traditionally, the development of electrodes has been systematically guided by trial-and-error methods. These approaches are mostly time consuming, resource intensive and lack of prediction accuracy for optimizing the complex nature of material characteristics.

To address the challenge, this thesis adopts machine learning (ML) strategy for predictive analysis of material characteristics. The dataset was formed using previously published experimental literature, including physicochemical and electrochemical properties of activated carbon-based supercapacitor. This raw dataset first preprocessed and subjected to feature engineering to prepare it for

training, validating and testing ML models, using both regression and classification techniques which will evaluate the performance of supercapacitor using ML performance metrics such as root mean square error (RMSE), mean absolute error (MAE), and correlation coefficient. The advanced analysis techniques like shapely additive explanation (sHAP) and feature importance score are applied to interpret the feature role in model predictions. The developed ML models successfully identified the key performance features governing supercapacitor performance. The data-driven approach provides powerful insights that can reduce research and development timelines by guiding target electrode material synthesis.

The established ML approach offers great potential for optimizing supercapacitor design, enabling predictive assessment of key performance indicators. This methodology represents a paradigm shift in energy storage material development, fabrication, and facilitate sustainable and efficient supercapacitor optimization through computational intelligence. The thesis presents machine learning methodology into material science field to guide the fabrication and synthesis process of electrode for supercapacitor using simulation by predicting the specific capacitance of supercapacitor for analyzing performance of energy storage device.

# Chapter 1

# Introduction

*[This chapter establishes research foundation, problem statement, objectives and scope for supercapacitor technology. It outlines methodologies for electrode material synthesis, electrode preparation, and utilization of natural resources precursors. This chapter outlines the thesis flow with the brief chapter description and findings.]*

## 1.1 Research Background

One of this generation's most pressing demands is energy. Energy can be stored in a variety of ways, including chemical, electrical, thermal, and electrochemical storage. With the decline of energy fuel, the demand for energy storage devices with renewable energy systems is increasing. Energy storage systems have emerged as the most pressing research problem for sustainable energy development worldwide. Global energy rules have highlighted rising energy demands, and to fulfil future objectives, renewable energy systems have been put in place to deliver the necessary energy in a timely and environmentally responsible manner. When these renewable energy sources are unable to produce enough electricity, a storage device is required to meet the demand.

For the last hundred years, the main energy sources have been fossil fuels. Fossil fuel is a nonrenewable energy resource that gets depleted with time. So, to overcome this problem the research has been shifted toward the generation and storage of energy without carbon dioxide emission. Most renewable sources are continuous and stable therefore it is necessary to develop an energy storage system that will use these sources optimally [1]. Energy storage devices play an important role in modern day technology and sustainability. It has many applications in a wide number of fields in the real world [2]. Some applications are electric vehicles, smartphones, laptops, wearable devices, medical devices, transport like trams and trains for regenerative braking and aerospace. There are different forms of energy storage devices such as batteries, fuel cells, and supercapacitors. The focus is on supercapacitors because of its unique advantages such as fast charging, high power density with good cyclic stability, rapid charging and discharging, and wide operating temperature which make them compatible and flexible for fast energy

storage and used for various applications [3], [4]. In this study, we propose a strategy to prepare porous carbon nanomaterials from different precursors to make carbon electrodes for energy storage devices. These structures have numerous cavities and active sites, make electrolytes more accessible, and boost ion transfer kinetics.

Globally, the demand for supercapacitors experiences a significant surge, expecting an increment in market value. This substantial growth has shown much interest from various companies, driving advancements in supercapacitor research and its commercialization. Currently, the materials used for making supercapacitor electrodes include conductive polymers, metal compounds, and carbons [5]. Among all these materials, carbon is considered the most prominent material due to its properties such as specific surface area, easy availability, high conductivity, and capacitive performance. Previously, porous carbon was produced from fossil fuel but the alternate way to make carbon material is by using Green Earth's abundant natural resources such as natural fiber, agriculture waste and biomass. All these resources are environmentally friendly, cost-effective, green, and sustainable. Due to its physicochemical properties, various forms of carbon material in the form of tubular structures like CNT, fullerene, 2D form, carbon fibers, graphene and their derivatives developed due to their application around energy storage. Carbon nanostructured material derived from natural resources can also be used to enhance functionalities in energy storage applications such as supercapacitors. Supercapacitor is the device which stores charge at the interface of electrode and electrolyte by two capacitive behaviours such as electric double layer and pseudo capacitance. Electric double-layer capacitance occurs due to electrostatic charge transfer while pseudo capacitance occurs due to faradic reaction [6]. The charge

stored in the electrode is directly proportional to the electrode surface and the energy stored is directly proportional to the charge stored in the electrodes [7]. This results in high power density compared to batteries, as the latter rely solely on slower electrochemical reactions. Consequently, while the energy density of supercapacitors is more than that of conventional capacitors. This shows that supercapacitors effectively bridge the performance gap between batteries and capacitors, making them advantageous for applications which require a balance between high storage and fast charging-discharging cycles. Supercapacitor performance depends on the electrode material like its surface area, porosity, electrolyte concentration, current collector, cell assembly and type, potential window, current density, charge-discharge rate, power density and it is limited by their energy density which depend directly on specific capacitance. The research mainly focuses on increasing the energy density and specific capacitance for better performance of supercapacitors

## 1.2 Problem statement and hypothesis

Material science has seen a significant change with the emergence of data-driven methods that combine machine learning and advanced computational technology to accelerate the design, discovery and optimization of novel materials. The paradigm shift is enabled using experimental and simulation driven large datasets that inform predictive models to estimate material properties and guiding the target synthesis of materials with the desired functionalities. These advancements have been influential for development of electrode materials, which will be used in diverse technological applications such as energy storage and conversions, electronic device engineering, and catalysis.

Among the different emerging applications of machine learning in material science, the data-driven prediction of property of electrode material determined impactful advancement. Such methodologies can substantially reduce the time and resources required for synthesis of electrode material while identifying the physicochemical properties of electrode material for energy storage devices like supercapacitor that may be challenging to determine through conventional experimental procedures.

Machine learning is used as powerful tool for discovery of electrode materials with targeted outcome properties. By screening previous published literature for forming datasets and predicting the property of energy storage devices. These approaches further support the design of devices exhibiting specific characteristics such as specific capacitance, power density and stability. Recent studies have demonstrated the use of machine learning in predicting electrode materials properties and identifying new tailored compositions with desired functionalities. Overall, these data driven methodologies hold substantial ground for transforming material discovery, material properties and enabling the development of application specific, high performance electrode material for supercapacitor.

Our hypothesis is that by employing ML techniques in predicting the physicochemical properties of carbon nanomaterial electrode for supercapacitors, thereby identifying material performance for supercapacitors more efficiently than empirical methods. If this hypothesis proves true, then this could predict electrode material performance before fabricating the real time device and encourage innovation within the field of energy storage devices.

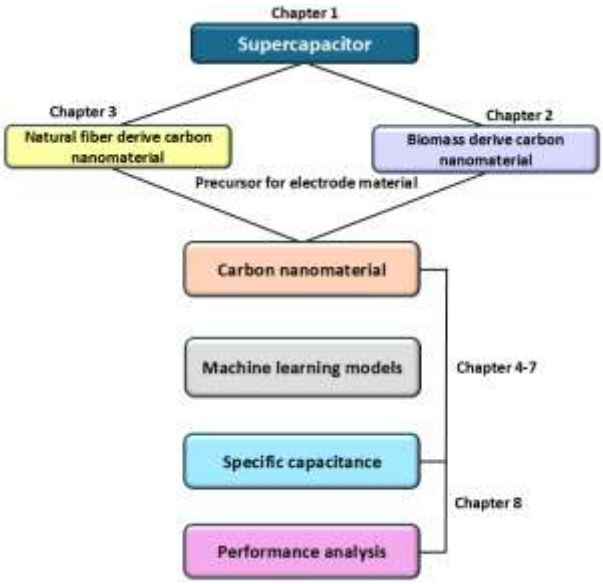


Figure 1.1 Outline of the chapter and thesis

## 1.3 Objective

The central objective of the thesis is to employ ML methodologies for prediction of property such as specific capacitance of carbon electrode derived from various precursors for supercapacitors. The research focuses on identifying and optimizing physicochemical properties of energy storage devices, with the focus on supercapacitors derived from various sustainable precursors.

To address the challenges of ML driven high throughput screening in identifying electrode materials for supercapacitors, this work proposes the formulation of physicochemical-based and electrochemical based descriptors. The workflow is often difficult due to limitation of experimental or simulation data for carbon-based materials that have not been characterized. By developing descriptors obtained from structural features and electrochemical features, the study focusses on developing predictive modelling techniques capable of accurately estimating electrochemical property such as specific capacitance prior to start fabricating the

device. These descriptors are employed to improve the performance of electrode material and optimize the design of high-performance supercapacitors.

To overcome the challenges which are imposed due to limited data availability for carbon material which hinders the dataset formation for accurate ML models, this work employs data preprocessing and feature engineering processes to filter datasets, and strategic choice of ML algorithms. The precise and accurate models derived through this process are used to predict the specific capacitance of supercapacitor

## 1.4 Thesis overview

This thesis focuses on predicting property of carbon material electrode for supercapacitor through high throughput screening methodologies. The thesis chapters are organized in the following manner:

**Chapter 1** introduces the field of supercapacitor, showing the significance of carbon nanomaterial used for making electrodes. It underscores the relevance of ML in predicting the performance of electrode material for supercapacitors. It outlines the problem statement and hypothesis, discusses the objective of the research work, and highlights the expected results.

**Chapter 2** provides the background of using biomass for making carbon nanomaterial-based electrodes for supercapacitors. The chapter discusses material properties, performance of material features, its electrochemical and physicochemical properties, as well as synthesis methods to make electrodes for supercapacitors.

**Chapter 3** provides a detailed review of the natural fiber precursors used for making carbon electrodes for supercapacitors. The properties of carbon electrode are important to select descriptors for applying ML models for performance predictions of supercapacitors.

**Chapter 4** describes the Machine learning methods employed for performance predictions and simulation techniques of activated carbon electrode for energy storage device. This chapter covers the entire flow from dataset formulation to development of ML models for predicting the specific capacitance of supercapacitor. The dataset consists of physicochemical features to train the various ML regression models. Among all the ML models applied, Random forest predicts the specific capacitance with superior accuracy with root mean square error (RMSE) of 61.88 and correlation coefficient ($R^2$) of 0.84. The model analysis identified specific surface area and doping of activated carbon acts as critical features for prediction, establishing ML as potential tool for accelerating energy storage material design. The primary objective of this chapter is evaluating feasibility of predicting specific capacitance using various regression-based algorithms. The main contribution of the chapter is the development of preliminary framework and identification of suitable ML models for evaluating the performance of supercapacitor using activated carbon electrodes.

**Chapter 5** describes the ML based specific capacitance prediction of activated carbon electrode using combinational approach of classification and regression using machine learning algorithms. This chapter describes the various supervised ML algorithms used for training the model using physicochemical descriptors. The dataset of over 500 peer reviewed publications on activated carbon material was assembled to predict material performance through ML performance metrics. The

ML algorithms employed are Random forest, XG boost, Decision tree, Support vector machine, K nearest neighbors, Ada boost, and Gradient boost. The results obtained after applying these algorithms demonstrate that Ada boost works better for class predictions and Gradient boost for value predictions. The Ada boost with root mean square error (RMSE) value obtained for class prediction is 0.687 and Gradient boost model gives RMSE value of 62.09 for value prediction which works best compared to other ML algorithms. Feature importance and sHAP analysis highlight specific surface area, potential window as key important features of analyzing supercapacitor performance. The key contribution of the chapter is integration of class based and value-based prediction methods, which provide comprehensive analysis of prediction and enhances the interpretation of model outputs.

**Chapter 6** focuses on energy storage device performance using data-driven approaches to predict specific capacitance of supercapacitors using boosting regression models such as Gradient boost, XG boost and Ada boost using hyperparameter tunning. The boosting models are predicting good accurate results in the previous chapter, so the comprehensive analysis is done in this chapter. Among these models, XG boost demonstrates superior efficiency, providing RMSE value of 36.60 and means absolute error (MAE) of 28.17 for evaluating the model performance. The subsequent analysis is done to determine the influential features which play an important role in model performance. The sHAP analysis is performed to determine the impact of features in prediction, providing critical insights for guiding the electrode fabrication for supercapacitor. The main contribution is identification of most influential features with hyperparameter

optimization, thereby providing more detailed understanding between material properties and performance of supercapacitor.

**Chapter 7** focuses on datasets of sustainable biomass-based precursors to produce carbon nanomaterial for supercapacitor. The various ML models were applied to predict specific capacitance of biomass derived carbon electrode. Among all evaluated models, Cat boost demonstrates superior accuracy with RMSE value of 22.06 and $R^2$ score of 0.946. The feature prioritization analysis confirmed the model's reliability establishing data-driven methodology and efficient alternative to conventional method for guiding supercapacitor fabrication and optimization.

**Chapter 8** concluded all the methods used to optimize supercapacitor performance. These data- driven models optimize synthesis parameters and reduce experimental reliance. These models can also be used to guide the performance predictions of different properties of energy storage devices such as power density, conductivity, and life cycle. ML application significantly enhances the efficiency and effectiveness of energy storage research and development.

# Chapter 2

# Background of energy storage devices derive from carbon nanomaterial electrode

*[In this chapter, we review how to prepare carbon nanomaterials derived from biomass precursors aimed at fabricating electrodes for Energy storage systems, especially supercapacitors. The comparative analysis of different biomass precursors used for producing carbon, and different synthesis methods will be explained and discussed, how it will change the functionality and electrochemical properties of material so that it will be used for making good electrode for energy storage devices.]*

## 2.1 Introduction

One of this generation's most pressing demands is energy. With the decline of energy fuel, the demand for energy storage devices with renewable energy systems is increasing. When these renewable energy sources are unable to produce enough electricity, a storage device is required to meet the demand. Energy storage technology will enhance efficiency and then fulfil when demand is high. These devices must be more powerful, lighter in weight, smaller in size, and economically viable.

By means of efforts by the research community to develop sustainable electrode materials for storage of energy from soft and hard carbon which can serve as the substitute for graphite such as nanofibers, nanotube and graphene, these materials are used in storage devices due to the properties of surface area, porosity, and capacity [8]. By using traditional synthetic procedures to create these energy storage devices limit their commercial application due to their expensive cost. Therefore, in contrast carbon which is produced from biomass precursors are receiving major attention due to synthesis method, environment friendliness, resource availability and their porosity and controllable morphology [9]. Energy storage devices such as supercapacitors are undergoing extensive study to increase their energy density to keep up with the rapidly rising renewable energy demand.

Carbon nanomaterial was developed by Kroto and co-workers in 1985 after that the work on carbon nano material was followed by Iijima in 1991. Research had been enhanced by the including various properties of Carbon nanomaterials (CNM) like strength, area to volume ratio, energy conversion, power and energy density and storage capacity. CNM is the type of nano structured materials have outstanding mechanical and electrical properties. CNM enhances the conductivity and

electrochemical properties and provide support to the substrate. The CNM based electrode made from biomass precursors have high surface area, increasing the contact of electrode and electrolyte that helps to decrease the weight of the CNM made devices [10], [11]. The application of these energy storage devices is growing day-by-day from batteries to electrical vehicles due to their unique properties of long operational life, rapid charge and discharge cycle, good energy density and capacity retention

## 2.2 Biomass resources

Biomass is a renewable energy source obtained for organic resources produced from food, waste, animal waste, aquatic plants, and agricultural production. It is the one renewable carbon source that generates the most carbon globally. The United Nations estimates that the yearly global production of agricultural biomass is close to 30 billion tons, and that its volume is continuing to rise because of increased agricultural activity [12]. Traditionally biomass has been limited to burning, dumped in the landfills which result in the degradation of environment, but the alternate use of biomass reduces the waste and pollution caused by it. The alternative ways to utilize energy from biomass to meet the increasing need of energy such as cooking, heating, electricity generation. Biomass is used as the indirect resource to develop value added chemicals, including biochemical and carbon materials. Therefore, it is concluded that biomass is used to develop renewable, eco-friendly, green, and sustainable carbon-rich materials. In recent years, biomass derived lignin, cellulose, carbohydrates, and proteins have been extensively utilized to prepare carbon-based materials. The chemical components of biomass are composed of lignocellulose, protein, and other extractives. This chemical composition is essential for selecting end uses of the material produced

and their processing techniques. The lingo cellulosic biomass mainly composed of woody biomass, stalk, saw dust and lignin. The second one is non lingo cellulosic which includes animal and human waste, it consists of lipids, proteins, and carbohydrates. The lignocellulosic biomass is rich in carbon resource.

## 2.3 Materials and methods

### 2.3.1 Biomass precursor for carbon nanomaterials

Biomass acts as a carbon source for producing carbon nanomaterials. There is a different precursor of biomass which is used to obtain carbon nanomaterials by using several synthesizing techniques. A. Dubey et al. [13] in his research took jute and insect derived Tassar cocoon, mulberry cocoon as the biomass product and then bio charring these materials to regain carbon and nitrogen as energy storage material. The insect precursor biomass produced N doped carbon after processing, it got converted into reduced graphene oxide (RGO). The jute after bio charring gave pure carbon form. Those materials showed excellent characteristics of charge storage. Following that, B. Liu et al. [14] used perilla fruitscene leaves as the biomass precursor and a typical pyrolysis process was used with temperatures ranging from 600 to 800 $^{0}$C and the product obtained was Carbon nanorods. The biomass precursor used by W.J. Liu et al. [15] in their research was saw dust char with the CVD method of synthesis using $FeCl_3$ as the catalyst followed the temperature of 600-800 $^{0}$C for producing nanofibers/porous carbon composites.

Table 2.1 Biomass precursor for producing carbon and its properties.

| S. No. | Biomass | Carbon nanomaterials | Properties | References |
|---|---|---|---|---|
| 1 | Corn stalks | Activated carbon | Specific surface area (SSA), porosity | [16] |
| 2 | Coffee grounds | Porous carbon | High SSA | [17] |
| 3 | Banana | Activated carbon | SSA, capacity retention, life cycle | [18] |
| 4 | Algae | Carbon nanotube | Renewable, SSA, porosity | [19] |
| 5 | Peanut shell | Activated carbon | High SSA, low cost | [20] |
| 6 | Bamboo | Porous carbon | High strength, renewable | [21] |
| 7 | Potato starch | Carbon dots | Good electrical conductivity, SSA | [22] |
| 8 | Wood | Activated carbon | SSA, conductivity | [23] |
| 9 | Jute | Carbon tube | Conductivity, porosity | [11] |
| 10 | Tea waste | Activated carbon | Cyclic stability, capacity retention | [24] |

### 2.3.2 Synthesis methods

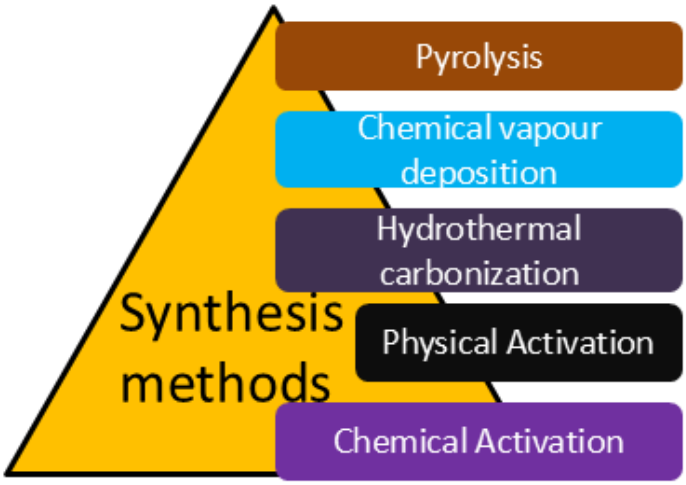


Figure 2.1 The synthesis method of producing carbon nanomaterial from biomass precursors

- **Pyrolysis/Bio-charring**

Conventional pyrolysis is method of heating of biomass without the presence of oxygen to convert it into carbon material also called biochar. It is performed at moderate temperature at or above 500 $^0$C under an inert atmosphere. There is no amount of oxygen is present; therefore, it is not possible for combustion to take place which allow biomass to get decompose into combustible gases. Among the combustible gases, some gases allow to provide heat for the process. Liu and co-worker [25] found that the wrinkle carbon nanosheet was formed by using the perilla fruitscense by performing the pyrolysis process at the temperature of 600-800 $^0$C and the nanosheet obtained with the thickness of 30-40nm.

- **Chemical vapor deposition (CVD)**

The production of CNM using CVD reaction is the two-step process in which the first step is to produce precursors via pyrolysis method. In second step, this precursor is deposited over catalyst usually transition metal such as nickel, iron, and cobalt. The catalyst was also deposited onto the support material including alumina, the nano-meter size particle growth formed over the catalyst by flowing methane and acetylene gas in the CVD furnace.

Acomb et al. [26] observed that the iron as a catalyst displayed greater CNT growth as compared to nickel, cobalt, and copper as the catalyst. The primary explanation for this is that the iron created smaller particles over the support surface. In their study, it was observed that support also play important role in active growth of nanoparticle over the catalyst. The iron over alumina showed better response than iron over silica [26]. In another study by mahanandia and coworker it was found out that the temperature play an important role in fabricating CNT structure, in his

paper normal CNT was synthesized at 700 $^0$C whereas iron-catalyst-loaded CNTs were required a higher temperature of 950 °C [27].

- **Hydrothermal carbonization**

Biomass with high moisture content can be effectively pretreated through a thermochemical method known as hydrothermal carbonization, applicable for various uses. The process occurs over a duration of 5 to 240 minutes within a sealed reactor, maintaining a temperature range of 180 $^0$C to 280 °C while maintaining pressure between 2 to 6 MPa. The primary product obtained after the process is called hydro char. The distribution of the obtained product yield depends on the process condition and biomass precursor used. The hydro char is partially carbonized and has high density oxygen groups. The hydro char obtained from HTC has low specific area and porosity, further activation is needed to modify its chemical and physical properties. F. Chen et al. [28] developed bamboo derived electrode through HTC by synthesizing activated carbon through citric acid assisted HTC followed by KOH activation. It was observed that adding citric acid can improve degree of hydro char, while also enhanced specific surface area and oxygen content in the material.

### 2.3.3 Activation methods

- **Physical activation**

It is performed in gaseous environments mainly in steam, air, or carbon dioxide. The degree, oxidizing agent, biomass source, and activation temperature, and all affect the activated carbon's physical and chemical properties.

When air or oxygen acts as an activating agent, the exothermic interaction of carbon with air and oxygen causes difficulties, which are attributed to it. The robustness in

rate of reaction made it difficult to handle which resulted in uncontrolled burn and decrease its yield. This activation process has the advantage of lower energy activation.

The steam reacts more quickly compared to $CO_2$ because of its smaller molecules and enters the pores of biochar. Throughout this procedure, a sample of biochar is continually flushed with the gaseous oxidizer while being exposed to nitrogen. High porosity carbon produced by this process in which oxidizing agent penetrates the biochar's interior structure during activation, where it reacts with carbon atoms to increase the surface area which depends on activation condition, while $CO_2$ impedes the entry of molecules into the pores. It seems that all pores take part in the activation/gasification reactions in $CO_2$ activation.

- **Chemical activation**

It is conducted within the temperature range of 400 – 900 °C on biomass carbon precursor. It prefers more than physical activation due to low pyrolysis temperature, better carbon yield, more surface area and properly distributed micropores structures which are beneficial for energy storage application. KOH is a chemical activation reagent in which activation carbon matrix engraved by potassium-containing elements, resulting in micro and mesopores in material. Further, the water vapor generated during the activation process contributes to the gasification, thereby improving the porosity. By controlling the precursor proportion, activation time and temperature the porosity can be transformed. Hu et al. [29] in his research used coconut shell as the biomass-based carbon precursor to prepare a high surface area activated carbon followed by KOH activation. The surface area of carbon material enhanced between 16 and 709 $m^2/g$ by using an activation temperature of 600 $^0C$, further increased to 1379 $m^2/g$ when temperature was raised to 800 $^0C$.

$ZnCl_2$ activation is also one of the chemical agents used for activation of biomass derived carbon into highly porous carbon activation. $ZnCl_2$ could facilitate aromatic condensation processes at high temperatures and the dehydration at low temperatures because of its Lewis acid nature. Sun et al. [30] synthesized Boron and Nitrogen doped porous graphitic carbon (BNGC) using chitosan followed by $ZnCl_2$ as activation. The unique microstructure of chitosan combined with $ZnCl_2$ activation results in BNGC with high surface area of 1567 $m^2/g$ and enhanced electrical conductivity.

## 2.4 Biomass derived storage devices and its properties

Systems that store energy for later use are known as energy storage devices. They are essential for improving grid resilience, integrating renewable energy sources, and controlling the supply and demand for energy. Supercapacitors are energy storage devices, especially in applications which require rapid charge and discharge cycle, long operational life. The research is going on in this field focuses on improving its efficiency, reducing its cost and expanding its application in various sectors.

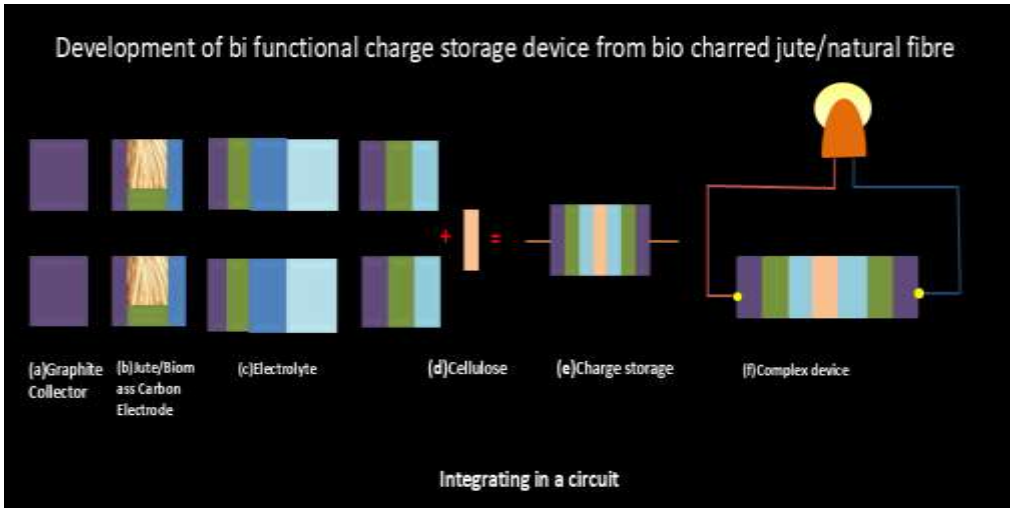


Figure 2.2 Development of storage devices from bio-charred jute.

It is categorized into electrochemical double layer capacitors (EDLCs) and pseudo capacitors (PCs). The EDLCs works on physical absorbing and deabsorbing electrolyte ions in the pores of electrodes when an external voltage is applied. The PCs phenomenon depends on redox reaction that takes place on electrode surface. PCs have more charge storage compared to EDLCs but low cyclic stability. Carbon derived from biomass has been widely used in supercapacitors, either as a substrate to manufacture PCs materials or as EDLCs electrodes. For high performance supercapacitors, efforts should be made to balance pore structure and robust surface properties.

Other than porosity and surface area, surface properties of CNT obtained from biomass also have significant potential in electrochemical performance of the energy storage device. Some research studies show that by adding functional group and heteroatom doping can alter the surface properties like pseudo capacitance, conductivity, and wettability of electrodes. Typically, Nitrogen heteroatoms are widely studied, it was found that mixing N atoms in carbon material enhances the electronic conductivity of CNT. Additionally, other heteroatoms including Boron, Phosphorous, and Sulfur also improved the performance of carbon materials in supercapacitors. B-doping, like N-doping, has the potential to enhance electrical conductivity and introduce pseudo capacitance, leading to high-performance carbon materials.

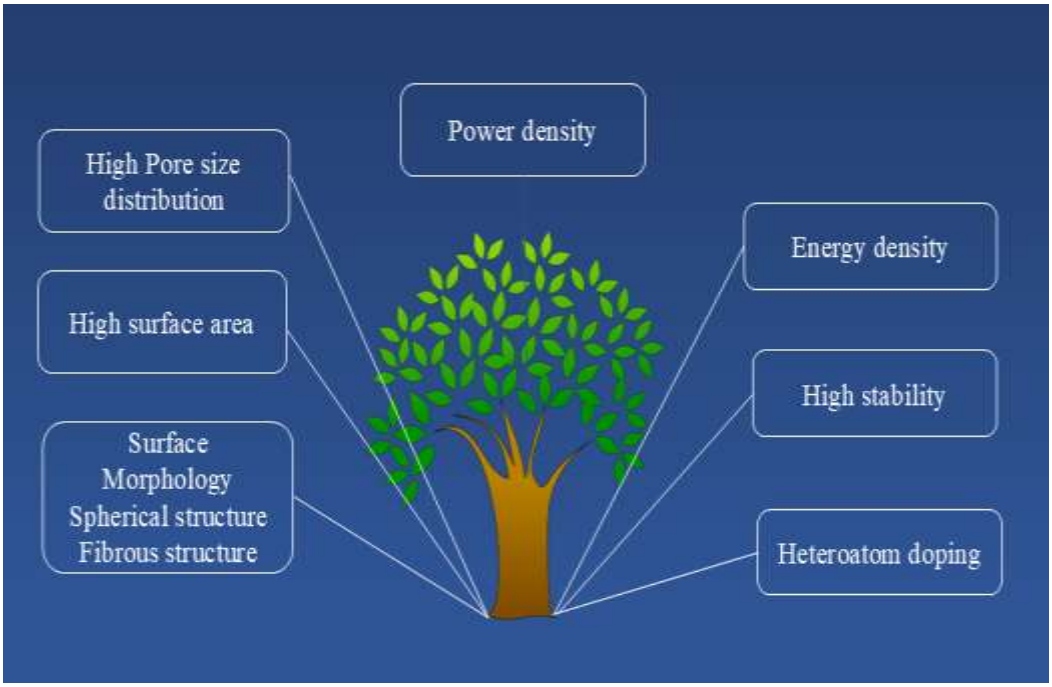


Figure 2.3 Properties of energy storage device derived from biomass precursor

## 2.5 Conclusion

CNMs have been demonstrated in this chapter as potential material for the energy storage application because of their physical and chemical properties. The biomass precursor has been discussed in the chapter to produce carbon as electrode material for supercapacitor. The method of synthesis, properties and morphology is also taken into consideration for efficient electrode for energy storage. The architecture of carbon nanomaterials has been described in this chapter as it plays active role on the electrochemical performance of super capacitors. It is also examined that the small modification like doping, defect creation and the size in its structure give out excellent electrochemical performances. The composite electrodes give high performance like conductivity, high surface area, and enhanced specific capacity. The next chapter discuss the natural fiber-based precursor for making electrodes for supercapacitor.

# Chapter 3

# Carbon nanomaterial-based electrodes from natural fiber and biomass precursors for energy storage devices

*[The growing demand for renewable high-performance energy storage systems made to explore natural fiber precursors to develop carbon nanomaterials for next-generation energy storage devices. This chapter focuses on synthesis of carbon electrodes from natural fibers using various carbonization and activation methods. This chapter reviews development of electrode nanomaterials from natural fiber obtained precursors for supercapacitor applications, highlighting their preparation methods, electrochemical characteristics, performance, renewable and sustainable nature. The derived carbon nanomaterial is engineered to exhibit improved energy and power density, specifically for supercapacitor applications. It discusses recent challenges, properties, characteristics, and future research directions in developing cost effective and environmentally friendly supercapacitors for energy storage. This chapter, therefore, provides a novel contribution to reviewing natural fiber-based carbon electrode in the energy storage field.]*

## 3.1 Introduction

Globally, the demand for supercapacitors experiences a significant surge, expecting an increment in market value. The current chapter focuses on increasing the charge storage capability for better performance of supercapacitors using natural fiber derived carbon. The life cycle and techno-economic assessment of natural fiber derived carbon materials for energy storage is also important which focuses on assessing the environment and cost comparison of the sustainable electrode material. The life cycle involves the entire process from natural fiber precursor to synthesis into carbon form, to device fabrication, to its operation performance and end of life recycling or degradation, focusing on low carbon footprint and renewability. The natural fiber precursors sources such as corn fiber, sugar bagasse fiber, cotton fiber, silk fiber and kenaf fiber will be taken into consideration in this literature review chapter are naturally abundant, renewable and degradable in nature as compared to fossil fuel-based carbons such as coal tar, petroleum, and coke derived materials. The natural fibers precursors obtained from agriculture cultivation practices and byproducts sometimes obtained from waste such as textile waste and industries, can offer low-cost feedstock reducing environment impact. The transformation into carbon material involves carbonization and activation can be performed at lower temperature as compared to graphitization of fossil fuel derived carbon material, that result in low energy consumption. From techno-economic perspective natural fiber derived carbons exhibit low costs due to utilization of waste, whereas fossil fuel derived carbon material involves high processing costs, energy consumption processing and environment taxes associated with greenhouse gas emissions. Moreover, natural fiber derive supercapacitor exhibit good electrochemical performance with high surface area, high porosity and

tunable heteroatom doping without the need for costly treatments. When considering overall precursor to device fabrication costs along with its sustainability, natural fiber provides cost effective and sustainable alternative to conventional fossil fuel derived carbon for high performance supercapacitor applications.

The focus of this chapter is to provide a comprehensive overview on development of activated porous carbon derived from natural fiber precursors using different synthesis and activation techniques. Additionally, it also examines the influence on its properties such as morphology, specific surface area, porosity, and specific capacitance on performance of energy storage devices, mainly supercapacitors. The influence of other factors such as heteroatom doping in the carbon to enhance the performance of energy storage devices, and the methods to enhance their properties to increase their charge storage capabilities and energy density.

## 3.2 Natural fiber resources

Natural fibers displayed in Figure 3.1 are hair-like materials, similar to pieces of thread which are derived from plants, animals or minerals that are used for various purposes such as fabrics, energy storage, medical applications, and textiles [31], [32]. Natural fiber is biodegradable, sustainable and environmentally friendly [33]. Natural fibers are of two types: cellulose-based, and protein-based. Cellulose-based fibers include flax [34], jute [35], cotton, and other protein-based fibers include wool, silk [36], and mohair. There are some plant-based natural fibers such as cotton, Hemp, jute, Coconut Coir, Bamboo, kenaf etc.

Natural fiber resources offer dual opportunities: they serve as rich sources of carbon material for energy storage devices while also constitute as waste management challenge if not used properly. Their sustainable exploitation reduces the carbon

footprint and cost- effective by turning waste materials into value-added products. Smart utilization of natural resources is essential for developing sustainable green energy future. Natural resources are exploited through bioresources including natural fiber, biomass, crops and agriculture-based precursors as renewable carbon sources derived from plants, animals, and waste materials to develop nanomaterials for energy storage devices. Ignoring these resources can become waste, while their utilization promotes sustainability and socio-economic benefits in energy storage systems [37].

Natural resources like sugarcane bagasse, rice husks, and coconut shells are utilized to develop carbon nanomaterials, which have applications in various fields such as energy storage including supercapacitor and fuel cells, environment applications including water treatment and in healthcare fields including biosensors and bio-stimulators, improve efficiency and sustainability while reducing waste and environment degradation [38]. Natural fiber resources such as sugarcane bagasse, waste corn husk residue used to prepare high quality porous carbon nanomaterials with high surface area. These properties make it suitable for application in separation, energy storage and electrocatalysis [39]

Most of the carbon electrode materials were produced from natural fiber through thermochemical treatment with or without going through physical and chemical pretreatment or activation processes [40]. In previous studies by researchers, thousands of natural fiber precursors and numerous processing techniques have been seen and studied to produce carbon nanomaterials with physical and chemical properties and studied their electrochemical performance for energy storage devices [41]. However, very little research has been attempted to develop an exact guiding principle to explicate the relationship between the components in natural fibers,

their chemical processing method, the properties of as-developed carbon material, and their electrochemical performance as electrodes. In some studies, by researchers, it was found that the electrochemical performance of natural fiber-derived carbons is affected by their chemical structure and basic compositions[42]. All plant-based natural fiber mainly consists of hemicellulose, cellulose, and lignin [43] which is explained in further subsection in detail. These three substances are the main constituents of the plant cell wall also known as lignocellulose [44], that is removed by chemical processing. The chemical constituents of these fibers have properties like elongation, tensile strength and young modulus range which is important for various electronic device fabrications. Natural fibers have many beneficial properties; they have low density with relatively high specific surface properties. These fibers with renewable energy resources also have a cost advantage and comparatively easy processing [45]. Even within the same plant species, the chemical composition of fibers varies due to plant growth conditions, geographical considerations, and the method of fiber extraction.

In most natural fibers, cellulose, hemicellulose and lignin form very well-defined microstructures and it is retained after the thermochemical treatment of fiber that results in different morphologies of carbon such as biochar and activated carbon (AC). The thermochemical process used for converting biomass to carbon is known as pyrolysis or biochar. Biochar is the by-product obtained by burning natural fiber at low oxygen levels to produce a porous structure with a high specific surface area. The biochar obtained, and heteroatom presence in it, depends on the natural fiber precursor composition. So, it is very necessary to understand precursor functionalities and compositions to attain the desired properties and characteristics for energy storage [46]. The coating of natural fibers with polymers or activated

porous carbon after their synthesis and activation process enhances their electrical conductivity and makes them suitable for use as efficient carbon electrode material. The carbon obtained from biomass waste of natural fibers is more advantageous than the traditional method obtained because of its cost, sustainability, and availability. The traditional method of producing activated carbon is from coal and petroleum which is costly, and not environmentally friendly despite producing high carbon yield [47]. Natural fiber-based carbon electrode material is used for energy storage applications because of its appealing carbon framework and structure during conversion.

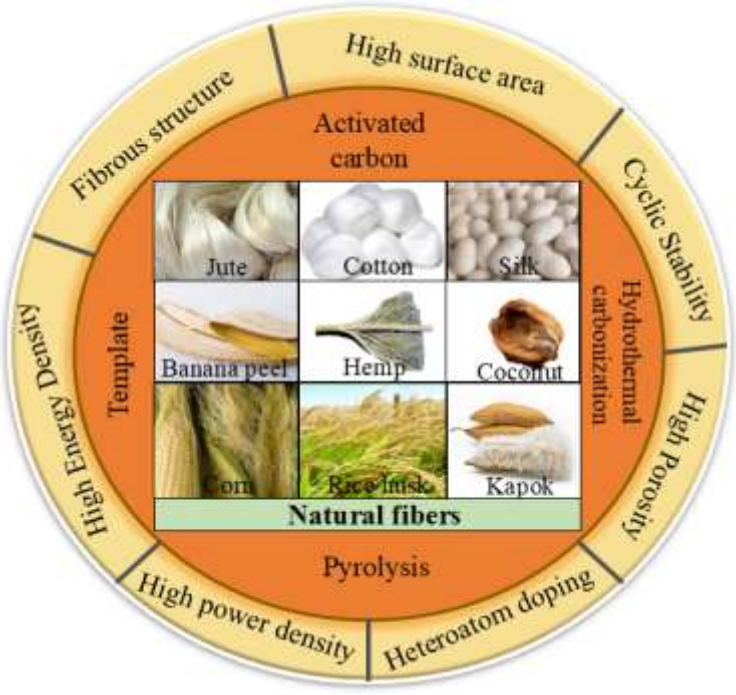


Figure 3.1 Natural fibers precursor, synthesis methods and properties of supercapacitor.

### 3.2.1 Lignocellulose

The natural fiber obtained from plants consists of chemical components of lignocellulose biomass displayed in Figure 3.2 which are composed of hemicellulose, cellulose and lignin, all of which have great potential for energy

generation and storage which will discussed further [48]. These three biopolymers are formed 3-dimensional structures in different degrees. Cellulose is the primary part that makes the plant cell wall. The hemicellulose is an amorphous structure attached to neighbouring cellulose fibrils via cross-links [49]. The lignin is a crosslinked macromolecule that bonds all together and helps strengthen cell walls [50]. The cellulose and hemicellulose in biomass obtained from natural fibers account for 70% of biomass content and are tightly crosslinked to lignin via a bond.

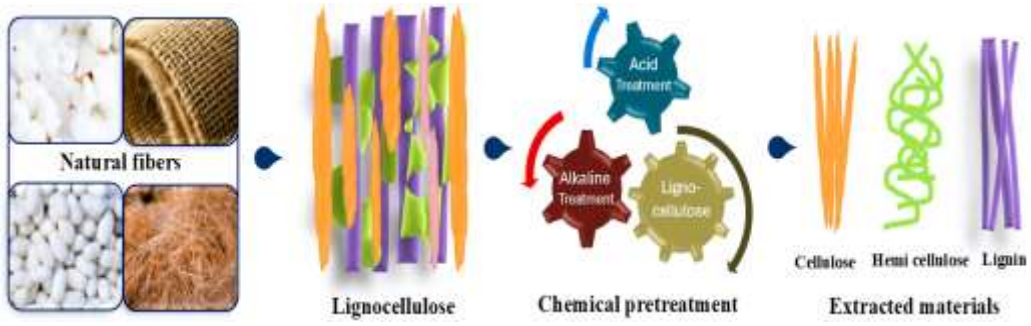


Figure 3.2 Extraction of hemicellulose, cellulose and lignin from natural fiber.

### 3.2.2 Cellulose

The cellulose part of lignocellulose is a carbohydrate polymer consisting of glucose units which act as a source of carbon. To extract cellulose from lignocellulose biomass, chemical pretreatment will be done to remove hemicellulose and lignin. The extracted cellulose is further purified to remove unwanted residual components that will affect the bio-charring process. The purified cellulose is heated using the pyrolysis method of synthesis which will be discussed in the next sections [51]. The synthesis method decomposes the cellulose into carbon material by removing volatile substances. The remaining solid composed of carbon with impurities depending on the natural fibre precursor and process conditions [52], [53] .

### 3.2.3 Hemicellulose

Hemicellulose is a heterogenous polymer composed of various glucose monomers connected to each other's through branches and amorphous structures. The branch frequencies of hemicellulose depend on the precursor nature and compositions. The hemicellulose is linked with cellulose through hydrogen bonds, with lignin through covalent bonds. It is easily hydrolysed during thermochemical pretreatment because of its low degree of polymerization. Hemicellulose has high content of carbon, and it is decomposed at low temperatures between 200 to 300 $^{0}$C which can affect the carbonization process. The presence of hydroxyl and acetyl groups enables the modification of chemical structure and helps to improve the properties of carbon materials obtained from it. The hemicellulose is extracted from lignocellulose biomass using NaOH or ammonia to break the bond between lignin and hemicellulose. The hemicellulose thermal decomposed to remove volatile compounds by obtaining carbon residue with amorphous structure [54].

### 3.2.4 Lignin

It is an aromatic polymer, that provides support and rigidity to the plant cell wall by cross-linking cellulose and hemicellulose fibers. Lignin forms a 3D cross-linked branched structure. It enables chemical modification and doping due to the presence of hydroxyl, methoxy and carboxylic groups. Lignin is extracted using dilute acids before thermal treatment, during the process volatile compounds are removed. The obtained structure converts into carbon material with a high degree of aromaticity that will further be used for making electrodes for supercapacitors [55]. Table 3.1 discusses the lignocellulose constituents of natural fiber derived biomass.

Table 3.1 Lignocellulose constituents of natural fiber derived biomass.

| S. No. | Material | Hemicellulose (%) | Cellulose (%) | Lignin (%) | References |
|---|---|---|---|---|---|
| 1. | Oil palm fiber | 50 | 25 | 70 | [55] |
| 2. | Banana fiber | 14.8 | 13.2 | 14 | [56] |
| 3. | Sugarcane fiber | 32 | 44 | 24 | [56] |
| 4. | Coconut fiber | 23.7 | 0.52 | 3.54 | [56] |
| 5. | Hemp | 17.9 | 74.4 | 3.7 | [57] |
| 6. | Kenaf | 33.9 | 53.4 | 21.2 | [57] |
| 7. | Flax | 12 | 64.1 | 11.8 | [57] |

## 3.3 Natural fiber to Carbon conversion method

Any biomass obtained from natural fiber is broadly synthesized through a two-step process consisting of carbonization followed by activation. The nanostructure and electrochemical performance of natural fiber derived activated carbon are significantly influenced by intrinsic texture of precursors, synthesis, methodology, activation time and temperature, activation concentration that will promote pore development. However, all these excessive activations can also lead to collapse of pore distribution, resulting in reduction of specific surface area and carbon yield [58].

The conversion of natural fiber into functional supercapacitor involves a well-defined process as illustrated in Figure 3.3 and 3.4. Initially, the precursor selection is a first step, where a suitable natural fiber is chosen based on its structural and chemical composition. This precursor undergoes the synthesis and activation process to convert it into carbon-based materials. The resulting CNM exhibit high surface area, tailored porosity and excellent conductivity making it ideal candidate

for fabricating electrode for supercapacitors. The sequential process ensures the effective utilization of natural fiber offering a sustainable approach to develop high-performance supercapacitor materials.

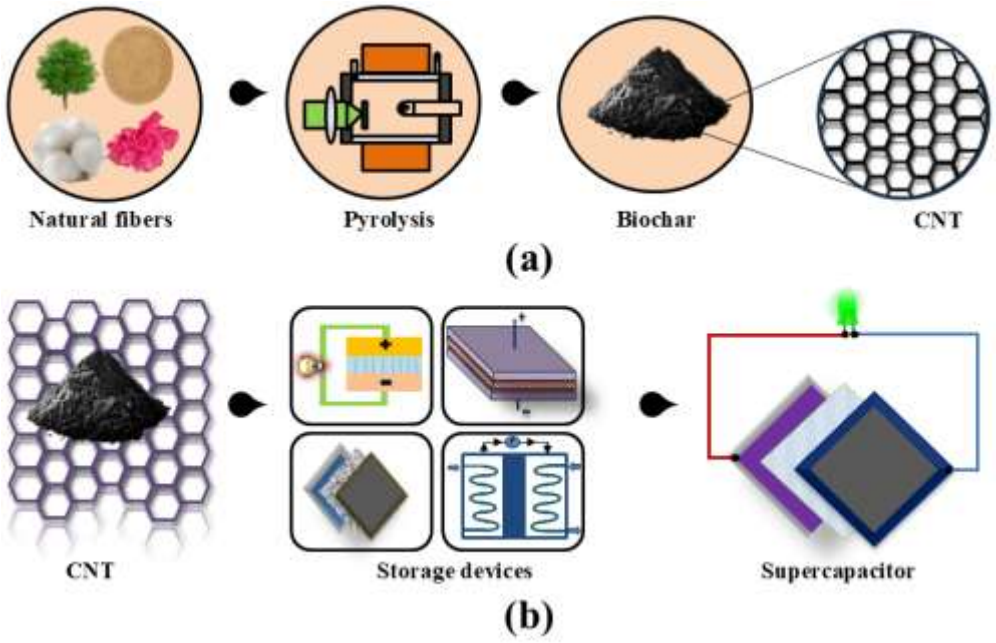


Figure 3.3 Schematic diagram illustrates the conversion of Natural fiber into carbon material (using pyrolysis) for energy storage devices. (a) Process of producing CNT from Natural fiber, (b) Energy storage application made from CNT

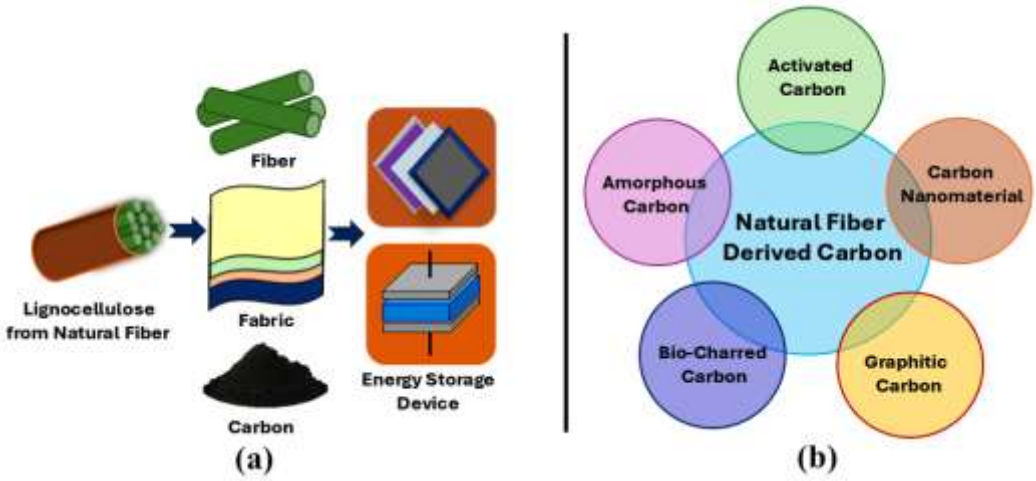


Figure 3.4 Schematic illustration of Natural fiber derived material and structures. (a) Natural fiber to produce different materials to energy storage device (b) Different form of Natural fiber derived carbon

The important features required for good electrode material are pore size range, easy fabrication methods, and good electrical conductivity [59]. Pore size ranges are classified into three categories, i.e., micropores have diameters less than 2 nm; mesopores have diameters between 2nm to 50 nm, and macropores have diameters more than 50 nm [60]. Nowadays, energy storage technologies require accurate control of the type of porosity present in the material.

De et al. in their study [61] developed porous carbon in which the synthesized porous material displayed high specific capacitance with an increment of 50-80 % than the corresponding solid carbon material. In that literature, the SEM graph showed that the particle size of porous materials was smaller than their solid particles which depicted the BET surface area of porous carbon was 220 $m^2g^{-1}$ while solid carbon was 24 $m^2g^{-1}$, this revealed that the high surface area led to increased diffusion of electrolyte into porous electrode results in high specific capacitance. The specific capacitance of porous carbon was 128 $Fg^{-1}$ at 20 $mVs^{-1}$ whereas solid carbon had specific capacitance of 84 $Fg^{-1}$ at 20 $mVs^{-1}$. Porosity requires some management to control the carbon's energy-storing capacities. Activation is also used to enhance the surface area of natural fiber-based carbon material by introducing many micropores. The surface area and porosity are very important factors to consider for the performance of energy storage devices.

## 3.4 Factors influencing Natural fiber derived activated carbon synthesis

### 3.4.1 Pretreatment of Precursor before carbonization

Prior to carbonization process, pretreatment of natural fiber precursors can offer various advantages. For instance, the synthesis of activated carbon from sugarcane

bagasse fiber by ethanol soaking and hydrothermal treatment followed by activation, has been shown to enhance microporosity, specific capacitance and performance rate. These properties improvement of natural fiber derived activated carbon are superior to those activated carbon prepared without any pretreatment process, highlighting the efficiency of tailored precursor modification strategies in optimizing electrochemical performance of activated carbon material for supercapacitor [62].

Swelling assisted synthesis represents an alternative strategy for fabrication of porous activated carbon from fiber derived cellulose. In this method, cotton fiber was pretreated by immersion in NaOH/urea solution, followed by carbonization process. This treatment effectively disrupts inter-intra molecular hydrogen bond of cellulose thereby reducing polymerization degree, enhancing it specific surface area and pore size. Notably the low ionic radius of Na+ ions infiltrates into cellulose matrix. Cellulose chains remain separated from each other and do not aggregate into compact structure. This occurs because urea hydrate from the dissolution process prevents hydrogen bonding between cellulose chains, thereby maintaining them more open form. This structural dispersion is useful for subsequent carbonization processes as it enhances porosity and increases the surface area of resulting hierarchical porous activated carbon [63].

### 3.4.2 Synthesis methods

The main synthesis methods to convert natural fiber to carbon as shown in Figure 3.5 are pyrolysis, hydrothermal carbonization, Chemical vapors deposition (CVD), and activation methods are physical and chemical activation. These synthesis methods are used to produce carbon in numerous forms and structures as depicted in Figure 3.4b. including activated carbon, bio-charred carbon, nanostructured

carbon, amorphous carbon, porous carbon such as carbon nanomaterial, graphene, and fullerene.

- **Pyrolysis**

Carbonization is divided into two main parts: primary pyrolysis and secondary pyrolysis. Primary pyrolysis occurs at temperatures 200 to 400 $^0$C and secondary pyrolysis occurs at 400 to 600 $^0$C. At less than 100 $^0$C temperature, all the moisture vaporizes from the material. Hemicellulose degrades between 250 $^0$C to 350 $^0$C temperature range and then the cellulose degrades at a temperature range of 350 to 400 $^0$C which is a part of primary pyrolysis, and then at higher temperatures, lignin degrades [13]. After this, the other non-desired substances be removed. The residue obtained after the pyrolysis is bio-charred carbon which is porous. Then, this bio-charred carbon will go through the activation process (physical or chemical activation) which increases its specific surface area and pore volume. The mechanical and electrical properties of the produced activated carbon and carbon composites made from the above process, that is biomass waste are very alluring because of the organized biomass structure [64]. To obtain the desired carbon structure for energy storage, there are other various methods also to convert natural fiber into activated carbon. However, the application of the intended product is the major factor in technique selection. The performance of the electrode for an energy storage device is determined by the precursor selection, synthesis method, activation methods and functional groups present on its surface [65].

- **Hydrothermal carbonization**

Pyrolysis is the phenomenon carried out in an inert atmosphere at a specific temperature range at or above 500 $^0$C while hydrothermal carbonization is performed in an aqueous environment at a specific pressure and low temperature

range of 120 to 250 $^0$C [66], [67]. The product obtained after hydrothermal carbonization conversion is called hydro char, hydro char is partially carbonized and contains heteroatom high-density oxygen groups. The hydro-char can be directly used as electrodes for energy storage. The existence of heteroatoms is beneficial in activated carbon material because it can increase charge transfer and specific capacitance in energy storage devices. Jain et al. [67] and coworker obtained BET surface area of 2440 $m^2g^{-1}$ and mesopore surface area of 1121 $m^2/g$ from hydrothermal synthesis of coconut shell at 275 $^0$C temperature.

- **CVD**

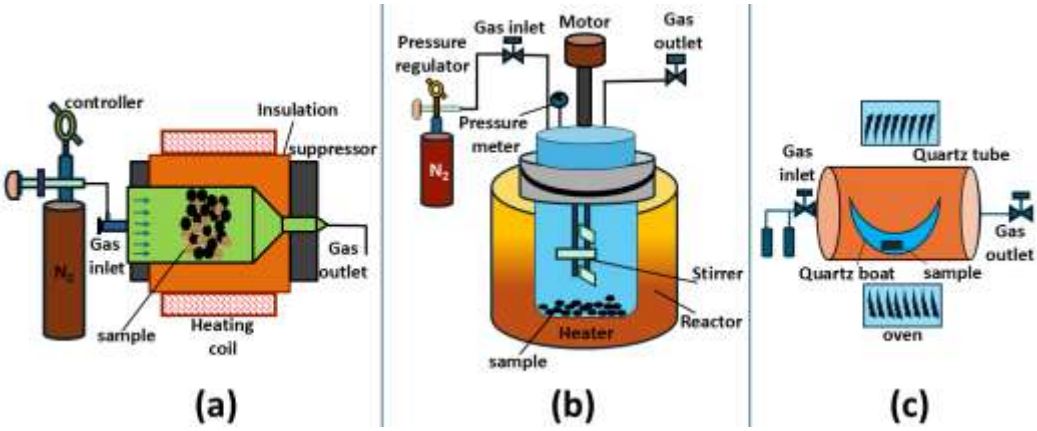


Figure 3.5 Different types of synthesis reactor to produce carbon. (a) Pyrolysis reactor, (b) Hydrothermal carbonization reactor, (c) CVD reactor

The CVD involves two stage process of synthesis. The first stage is pyrolysis method, ensuring a consistent source of carbon material. In the second stage, the precursor is deposited onto the catalyst composed of transition metals such as nickel iron, and cobalt. The catalyst is supported on substrate material mainly alumina, to enhance its stability and facilitate uniform growth. In the CVD furnace, the growth of nanometre size particles is initiated on the catalyst surface by introducing hydrocarbon gases, such as acetylene or methane. The interaction of carbon precursor with the catalyst under controlled temperature and gas flow enable the

formation of high-quality carbon nanomaterials. The two-stage process is beneficial because it ensures precise control over structural and morphological properties of resultant CNMs which are useful for energy storage applications [26], [27].

### 3.4.3 Activation methods

Physical Activation is a two-step activation process in which the first process is pyrolysis of precursors below 700 $^0$C followed by gasification in the controlled environment above 800 $^0$C in the presence of gases including $CO_2$, air, steam, and their compositions shown in figure 6a. The gasification process further encouraged the decomposition of pyrolyzed bio-charred material to generate interconnected porous structures. The degree of carbon bio-charring can be influenced by activation period, gas flow rate, temperature, and furnace selections. Guo et al. [68] investigated the $CO_2$ based activation of coconut shell fibers and observed that elevating the activation temperature promoted pore formation and facilitated an increase in pore diameter. Taer et al. [9] researched the $CO_2$-based activation of rubber wood-derived carbon and studied its electrochemical characteristics and found out that by activation temperature, conductivity and capacitance of activated carbon increased. The pores structure of activated carbon contributed to micropores due to $CO_2$ activation, but steam activation facilitates the widening of micropores and enhances its growth [69]. In the physical activation method, carbon dioxide or steam is mostly preferred as an activation agent due to ease and cleanliness.

On the other way, chemical activation is a single-step process above the temperature of 700 $^0$C combining pyrolysis and activation processes. The activation agents used are KOH, NaOH, $H_3PO_4$ and $ZnCl_2$ as shown in Figure 3.6a. The porous carbon showed high specific surface area and pore volume after activation method Chemical activation preferred over physical activation because of high carbon

yield, less activation time and temperature, high specific surface area, and high porosity as described in Figure 3.6a.

KOH is widely used as an activation agent acting as an oxidant and forming oxygen functional groups. The corn cob and KOH mixed in a ratio of 1:3 as an activation agent. This mixture was thermally prepared under an $N_2$ atmosphere for a certain period. The nitrogen doped activated carbon produced from corncob showed an increase in specific surface area of 2859 $m^2g^{-1}$ [70]. This study reveals that the activated carbon material enhanced the properties of material, also displays a pore structure distribution that had micropores to mesopores distribution, presenting benefits for energy storage applications. The carbon material depicted a specific capacitance of 185 $Fg^{-1}$ at a current density of 0.4 $Ag^{-1}$. The capacitance value falls within the range for electrode materials used in supercapacitor applications, indicating that the carbon material may be suitable for effective energy storage in electrochemical devices.

According to an experimental investigation, Fu et al.[71] conducted a study where rice bran was carbonized in $N_2$ environment at the heating rate of 3°C $min^{-1}$ for 1 hour at a temperature of 700 °C, followed by 850 °C for an additional hour. This process resulted in the production of carbon material with the pore distribution, exhibiting a specific surface area of 2475 $m^2g^{-1}$, and pore volume of 1.21 $cm^3g^{-1}$. The comparative analysis was performed in the study to evaluate the production of activated biochar using one step and two step KOH activation method. This comparison demonstrated a highly microporous structure was developed with highest specific surface area of 2138 $m^2g^{-1}$ compared to 741 $m^2g^{-1}$ during the first step. In one step activation method, the precursor was directly mixed with activation agent and pyrolyzed at 750 °C for 3 hours. In contrast, the two step activation

method involved pyrolysis of precursor at 450 °C for 3 hours, followed by addition of KOH as activation agent to the obtained biochar for further activation. Compared to single step activation, two-stage activation of KOH produces more activated biochar.

Sun et al. [72] used cotton as the precursor for producing porous carbon as electrode material for supercapacitor. For synthesis, MgO used as template and $ZnCl_2$ as an activation agent. The specific surface area and pore volume of absorbent cotton before activation were 127 $m^2g^{-1}$ and 0.1 $cm^3g^{-1}$. The specific surface area and pore volume of absorbent cotton after $ZnCl_2$ activation were 663 $m^2g^{-1}$ and 0.34 $cm^3g^{-1}$. The highest surface area obtained was 1990 $m^2g^{-1}$. For testing, KOH was used as an electrolyte in three electrode systems with specific capacitance of 240 F/g at $1Ag^{-1}$ current density.

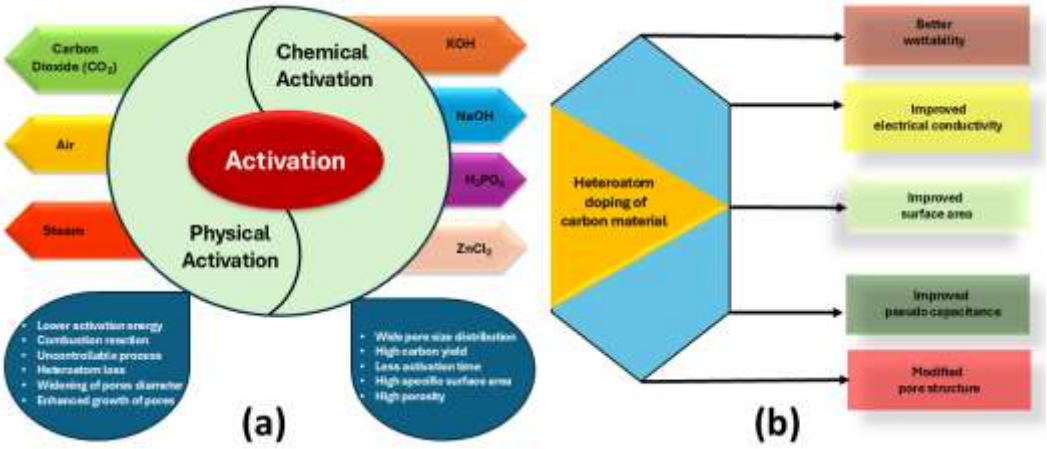


Figure 3.6 The description of activation and heteroatom doping. (a) Natural fiber derived carbon activation techniques involve physical and chemical activation, its types and advantages (b) Natural fiber derived carbon doping and its properties.

Hou et al. [73] focused on zinc chloride $ZnCl_2$-based activation on natural silk fiber for energy storage applications, in their research, they combined natural silk fiber with activation agent in a particular ratio followed by thermal treatment for 1 hour at 900 $^0C$ then the porous activated carbon produced displayed 2D nanosheet

morphology with a specific surface area of 2494 $m^2g^{-1}$, rich in nitrogen doping while the electrochemical analysis showed the specific capacitance of 242 $Fg^{-1}$, energy density of 90 $Whkg^{-1}$, and power density of 875 $Wkg^{-1}$. The activation agent increased the electrochemical properties of activated carbon material used for making electrode.

Lin et al. [74] focused on the role of different activation agents on the properties on activated carbon material obtained from bamboo fiber. The activation agents used were $ZnCl_2$, and $H_3PO_4$. After activation, the BET-specific surface area of $ZnCl_2$ and $H_3PO_4$ activated carbon were 1908.51 and 900.50 $m^2g^{-1}$. The pore volume also enhanced after activation, the pore volume of $ZnCl_2$ and $H_3PO_4$ activated carbon were 1.05 and 1.92 $cm^3g^{-1}$. The specific capacitance obtained for $ZnCl_2$ activated carbon was 169.79 $Fg^{-1}$ at current density of $0.2Ag^{-1}$ whereas the specific capacitance of $H_3PO_4$ activated carbon was 121.27 $Fg^{-1}$ at 0.2 $Ag^{-1}$. The study showed that the $ZnCl_2$ activated carbon perform better than other activation agents.

William et al. [75] demonstrated a comparative analysis of chemical and physical activation on biomass flax fiber. In his study, he showed that there was a significant enhancement in surface area using chemical activation. In physical activation, the surface area was 840 $m^2g^{-1}$ with mesoporous structure while in chemical activation using Zinc chloride as an activation agent results in a surface area of 2400 $m^2g^{-1}$ with microporosity.

## 3.5 Features of Natural fiber derived Activated carbon

### 3.5.1 Surface area and porosity

The high surface area of carbon provides large electrode/electrolyte interfaces for ion adsorption, effectively help in optimizing specific capacitance of carbon

material. In addition, configuration and connectivity of pores influence the ions and electron movement in carbon, thus affecting their electrochemical properties. The narrow pores are pursued to improve the rate performance and power density of device. The SSA and pore size distribution can be tuned and tailored by using different synthesis methods [76].

Micropores (< 2 nm) primarily contribute to charge storage by providing a high surface area for ion adsorption, whereas mesopores (2-50 nm) facilitate ion transport, thereby enhancing electrochemical performance. This suggests that an increase in activation temperature leads to an expansion in mesopore volume and an increase in the average pore diameter, which collectively improves the overall electrochemical properties of carbon materials. This study highlights that the combination of hierarchical pore structure and high porosity results in excellent electrochemical performance with materials showing high capacitance and energy density [77].

The carbonization temperature and KOH activation in which KOH to carbon mass ratio play crucial role in determining morphology and porosity of resulting carbon material. An optimal KOH to carbon mass ratio of 2:1 has been reported to enhance the specific capacitance, due to the formation of well-developed porous network that balance the micropore generation for charge storage and mesopore development for efficient ion transport. This optimization is essential for improving electrochemical performance of natural fiber derived carbon in supercapacitor applications [42].

### 3.5.2 Heteroatom doping

The heteroatom dopants include nitrogen, and sulfur play an important role in increasing the wettability and conductivity of material by transforming their

electronic characteristics which is described in figure 6b [78] . By carbonizing nitrogen-containing materials, such as polyvinyl pyridine $(C_3H_6NO)_n$, nitrogen can be incorporated inside carbon structures to create N-doped carbon. Another method of synthesizing nitrogen-doped carbon is through the thermal treatment of nitrogen-containing natural fiber. N-doped carbon enhances the performance of the supercapacitor by improving pseudo-capacitance. The specific capacitance of nitrogen-doped active carbon material is composed of a pseudo-capacitance and an electric double-layer capacitance. The electric double-layer capacitance occurs due to the physical phenomenon happening at the electrode/ electrolyte interface, and the pseudo capacitance is due to the faradic reaction. Mao et al. [79] in his work showed that the N-doped active carbon displayed capacitance of more than double (683 mF $cm^{-2}$ at 2mA $cm^{-2}$) as compared to undoped carbon when used as the electrode for supercapacitor with the improved long cyclic stability of 96 % after 10,000 cycles. Nitrogen doping also enhances the energy density of supercapacitors. The energy density improved to 6.7 $Whkg^{-1}$ as compared to 5.9 $Whkg^{-1}$ [80].

Li et al. [70] in his study depicts that the nitrogen content of activated carbon increases with an increase in temperature, the nitrogen content is 0 wt% at 0 $^0C$, 2.97 wt% at 400 $^0C$ and further increases to 3.98 % at 600 $^0C$. The enhancement in electrochemical performance is observed when nitrogen content increases from 0 to 2.97 wt% of nitrogen-activated carbon corresponding to temperature from 0 to 400 $^0C$. The specific capacity of 129 $mAhg^{-1}$, specific capacitance of 185 $Fg^{-1}$ at a current density of 0.4 $Ag^{-1}$, capacity retention of 86 % in 500 cycles, energy density of 141 $Whkg^{-1}$ and power density of 1747 $Wkg^{-1}$ was observed. However, when the temperature of activated carbon increases from 400 to 600 $^0C$, there is a decline in

the electrochemical performance of the material showing decrement in specific capacity, surface area, rate capability and capacity retention of nitrogen-doped activated carbon.

Huang et al. [81] synthesized sulfur doped activated carbon (SAC) by coating on the surface of porous carbon using pyrolytic decomposition method. The incorporation of SAC was beneficial as it introduced pseudo-capacitance, modified the pores structures and enhanced conductivity. The total surface area of activated carbon was measured 1094 $m^2g^{-1}$, after doping with sulfur the total surface area decreased to 987 $m^2g^{-1}$, indicating the sulfur partially covered the surface of the pores. The specific capacitance of activated carbon was 72.7 $Fg^{-1}$ whereas specific capacitance of SAC increased to 82.3 $Fg^{-1}$ at current density of 0.07 A/g. The result demonstrated that activated carbon doped with sulfur enhanced the electrochemical performance of material, making it the promising candidate for supercapacitor.

Chen et al. [82] prepared nitrogen and sulfur doped activated carbon from elm flower (EL) using hydrothermal carbonization method of synthesis to develop electrode for supercapacitor. The EL derived activated carbon exhibit nitrogen content of 2.21 atom % and sulfur content of 6.06 atom % respectively. The specific surface area of hydrothermal carbonized EL was 6.13 $m^2g^{-1}$ whereas after nitrogen and sulfur doping, the value of specific surface area significantly increased to 2638.94 $m^2g^{-1}$. The co-doped activated carbon showed a good specific capacitance value of 275 $Fg^{-1}$ at current density of 1 $Ag^{-1}$. The results demonstrated that doping greatly influences the properties for using it as promising material for supercapacitor applications.

### 3.5.3 Nanostructure

Carbon material exhibits diverse nanostructures including one-dimensional (1D) nanofiber, two dimensional (2D) nanosheet, three dimensional (3D) hierarchical framework and naturally form architectures. The strategy design of nanostructure in natural fiber derived carbon enables the optimization of specific surface area, a well-balanced porous network and enhanced electrical conductivity. The electrochemical performance of energy storage devices is increased by these characteristics which guarantee efficient ion adsorption and electron transport.

- **1D carbon nanostructure:**

1dimensional nanostructure derived from natural fiber contains fibrous and tubular morphologies [83], have been recognized as promising structure for high performance supercapacitors [84]. Their elongation geometry ensures maximum electron transport routes and easy ion diffusion, high SSA make them ideal for energy storage devices. The tubular form has high SSA, can serve as reservoir for electrolytes, improving ion accessibility and offering more active sites for electrochemical process. Additionally greater intrinsic 1D morphology of carbon derived from natural fiber have surface chemical and structural characteristics, which enables its integration into variety of materials and devices through tailored fabrication strategies [85]. In addition, the presence of functional groups, such as hydroxyl (OH) groups, attached to the surface of 1D carbon structures can increase interfacial wettability between the electrode and electrolyte, thus significantly influencing degree of infiltration in the electrodes [86]. Wang et al. [87] prepared a hollow carbon material through carbonization of cotton fiber and provide its application in supercapacitors. The electrode produced from above method exhibited a better specific capacitance of 175 $Fg^{-1}$ at current density 1 $Ag^{-1}$ along

with exceptional rate capability, retaining 107 $Fg^{-1}$ even at high current density of 60 $Ag^{-1}$.

- **2D carbon nanostructure:**

2D carbon nanosheets offer vast electrolyte accessible surface area, facilitate low resistance ion transport pathways, exhibit excellent electrical conductivity, and possess outstanding mechanical properties, rendering them advantageous for energy storage applications [88], [89]. Pore carbon nanosheets with hierarchical pore structures provide high internal free volume, which efficiently accommodate volume change during charge/discharge operations, thus providing structural stability and ensuring long cycling life. Such characteristics enable rapid electron and ion transport, which allow high power output while promoting active interfacial sites and charging carrier density, leading to high energy storage capabilities [90]. Hou et al. have reported the synthesis of the highly porous nitrogen doped carbon nanosheet derived from silk by one step process of activation and graphitization of natural silk. This process yields a 2D nanosheet with high SSA of 2494 $m^2g^{-1}$ and high hierarchical pore volume of 2.28 $cm^3g^{-1}$ both of which play an important role in optimizing electrochemical performance. This highly porous nitrogen doped carbon consists of a thin nanosheet morphology with rough thickness of 15 to 30 nm. This structure reduces transport resistance, allowing efficient mass diffusion and charge transfer at electrode/electrolyte interface, and hence improving the performance of supercapacitors. It possesses a specific capacitance of 242 $Fg^{-1}$ and energy density of 102 $Whkg^{-1}$, and shows high cyclic stability, retaining 91% after 10,000 cycles. This excellent performance is due to their high SSA and nitrogen doping, which increases pseudo-capacitance and facilitates ion transfer efficiency [73].

- **3D carbon nanostructure:**

The 3D carbon nanostructure consists of hierarchical interconnected porous network. These structures are useful for energy storage applications because they are designed to maximize surface area, electrical conductivity and ion diffusion pathways. The 3D carbon nanostructure involves 3D graphene network, carbon aerogels, carbon nanotube and activated carbon foams. With an increase in dimensionality a greater proportion of specific crystal become exposed to electrolyte, increasing active surface area. This structural advantage enabling better ion diffusion and charge storage, which leads to high capacitive performance of electrode materials [91].

Song et al. in his work prepared 3D hierarchical porous carbon from corn husk fiber synthesized using carbonization method. This study highlights unique structural properties of 3D hierarchical porous carbon which includes well defined pore distribution that effectively reduces diffusive resistance and promote efficient mass transport. These attributes are highly suitable for supercapacitors' applications. The specific capacitance of 200 $Fg^{-1}$ at current density of $20Ag^{-1}$, this demonstrates its ability to deliver energy rapidly with minimal capacitance degradation. The 3D hierarchical porous carbon exhibit SSA ranging from 814 to 928 $m^2g^{-1}$ along with hierarchical pore distribution of micro, meso and macropores. This well-organized pore structure promotes efficient ion transport and enhances charge storage capacity [92].

## 3.6 Materials and methods used for making Energy storage devices.

Different natural fiber precursors are taken into consideration for study like cotton fiber, banana peel fiber, jute, coconut fiber, and corn fiber for the synthesis of

activated carbon. Other chemicals used for activation purposes are KOH, NaOH, $H_3PO_4$, and $ZnCl_2$ etc.

### 3.6.1 Synthesis

Natural fiber precursors of carbon collected from the environment in desired quantities then it is washed with DI water to remove particles and impurities present, cleaned, and then sample were placed for sun drying or in the oven drying at a controlled temperature to ensure complete removal of moisture content [18]. The dried material was smashed into fine powder form to store it in a dry atmosphere. After removing all the moisture, the dried sample was soaked in ethanol/distilled water depending on the inherent properties of natural fiber. Subsequently, the activation process was carried out in which the obtained material was mixed with the activation agent such as KOH, $ZnCl_2$, $FeCl_3$, in desired ratio (gram) under controlled conditions. For the synthesis of activated carbon, mixture was subjected to thermal treatment in the tubular furnace under the $N_2$ environment for pyrolysis at the ramp rate at a certain temperature above 600 $^0$C for more than an hour. Post synthesis, the material was purified via HCL washing, followed by grinding using a mortar pestle to remove impurities, then the material was rinsed with distilled water until the pH became neutral for filtration to remove acidic residues. The purified material was dried in the oven to obtain the porous activated carbon [93].

### 3.6.2 Characterization

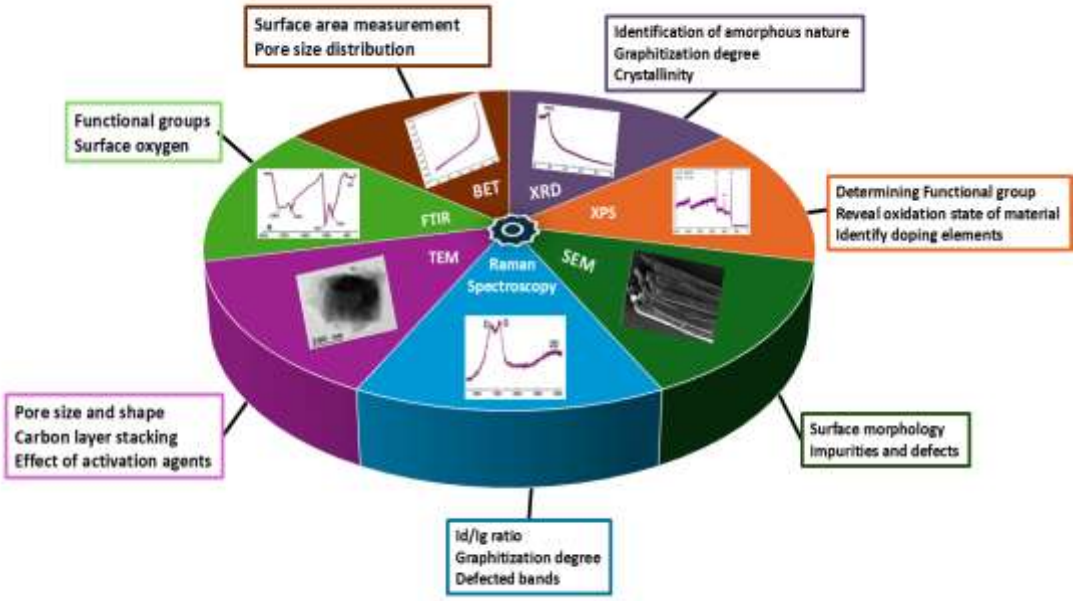


Figure 3.7 Schematic illustration of characterization techniques used for natural fiber derived carbon.

The characterization technique described in Figure 3.7 is used to examine the internal structure of natural fiber derived carbon. An X-ray diffraction (XRD) analysis was employed to determine structural properties, crystalline/amorphous nature, graphitic composition, and phase modifications of the activated carbon. Fourier transform infrared spectroscopy (FTIR) is an analytical technique used to determine molecular structure of activated carbon by measuring their absorption of infrared radiation across spectrum ranging from 4000 to 500 /cm. This method is used to provide information about functional groups, chemical bonds and their molecular structures. Scanning electron microscopy (SEM) analysis used to elucidate surface morphology of activated carbon to determine pore structure characteristics, like pore distribution and morphology. X-ray photoelectron spectroscopy (XPS) used for analysis of surface chemistry of activated carbon, used

to identify functional groups. Raman spectroscopy is used for characterizing electronic and structural properties of activated carbon. Brunauer-Emmet-Teller (BET) used to evaluate the pore volume and specific surface area of activated carbon [94].

To Study the electrochemical characterization techniques of activated carbon, three-electrode configurations were used in which platinum Pt was used as the counter electrode, Ag/AgCl was used as the reference electrode and the working electrode was prepared by mixing the 80 wt.% porous activated carbon with 15 wt.% acetylene carbon black and 5 wt.% of polyvinylidene fluoride (PVDF) acting as a binder. The prepared mixture was mixed with N-methyl-2-pyrrolidone (NMP) to make a slurry. This slurry was coated on Nickel foam (1$cm^2$) and used as a current collector, after that, the prepared electrode was dried in an oven at 100 $^0$C. The mass loading of electrode material was between 2-3 mg/$cm^2$. The electrochemical experiment was carried out in a molar electrolyte solution between a voltage range of 0-1 V [42]. The schematic representation of making carbon electrode is shown in Figure 3.8. The cyclic voltammetry and galvanostatic charge-discharge (GCD) curve was performed to examine the electrochemical properties. The GCD curve is a practical method to determine specific capacitance. The following Equation 3.1 shown below is used to find specific capacitance of electrode based on GCD profile:

$$C_{GCD} = \frac{I\Delta t}{m\Delta V} \qquad \text{Equation 3.1}$$

Where $i = I/m$ is a symbol of current density, m represents active mass, V represents a potential window. The specific capacitance of the prepared electrode material was estimated using Equation 3.2:

$$C_{sp} = \frac{fidV}{2*S*m*\Delta V} \qquad \text{Equation 3.2}$$

In which *i* represents instantaneous current, *dV* represents differential potential and $fidV$ represents the area bounded by the cyclic voltammetry curve. S denotes the scan rate, m is the mass of the material, and $\Delta V$ denotes the potential window.

The power density P and energy density E were determined using the following Equations:

$$E = \frac{CV^2}{2\times 3.6} \qquad \text{Equation 3.3}$$

$$P = \frac{E*3600}{\Delta t} \qquad \text{Equation 3.4}$$

Where *C* denotes specific capacitance in F/g, *E* denotes cell potential in volts and $\Delta t$ denote discharge time in seconds.

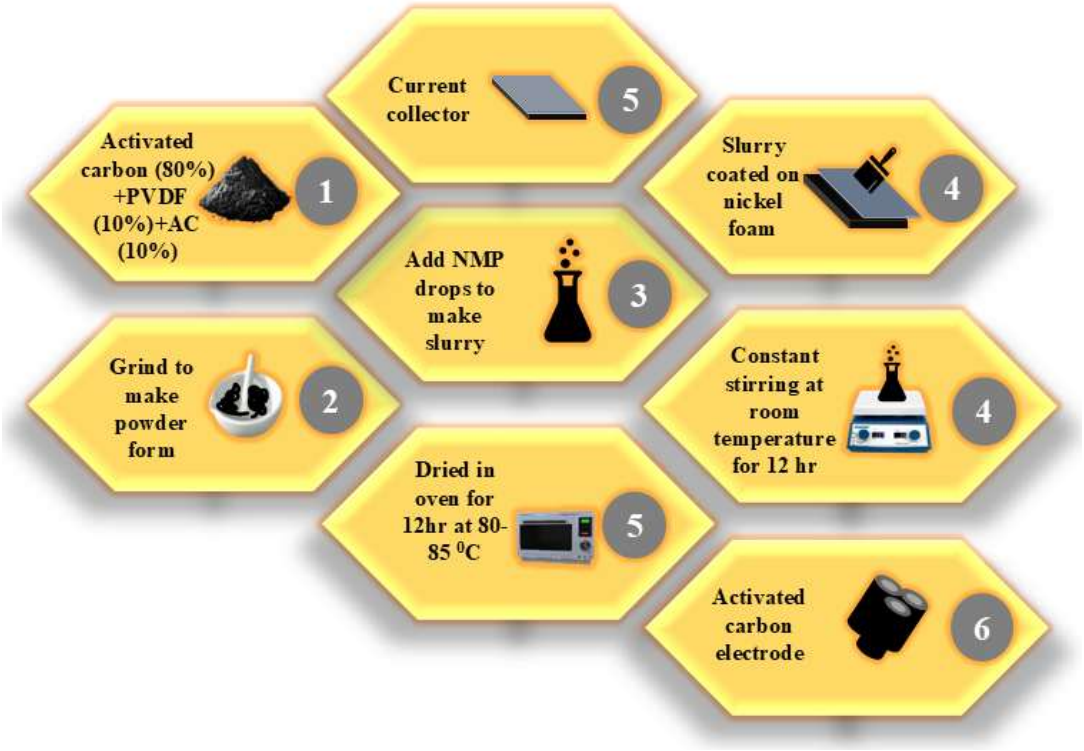


Figure 3.8 Process flow for making carbon electrode.

### 3.6.3 Properties of supercapacitors

Table 3.2 present the characteristics of carbon material derived from natural fiber precursors. The table includes key parameters like carbon precursor, specific surface area (SSA), Specific capacitance (SC), energy density, power density, capacity retention, life cycle, current density, and scan rate. A comprehensive analysis of these properties is critical for determining the suitability of the material for fabricating the electrode in supercapacitors. Such analysis is important to ensure optimal performance and efficiency, making the material suitable for making supercapacitors for commercial applications.

Table 3.2 Properties of natural fiber derived carbon materials for supercapacitors.

| S. No. | Materials | SSA | SC | ED | PD | CR | LC | CD | Ref |
|---|---|---|---|---|---|---|---|---|---|
| 1. | Activated carbon from onion peel | NG | 127 | 13.61 | 200.8 | 109 | 2,000 | 0.75 | [95] |
| 2. | Jute sticks derived AC | 1100 | 150 | 20 | 500 | 99 | 1000 | 1 | [96] |
| 3. | Recycled jute AC | 408 | 51 | 21 | 100 | 100 | 5000 | 0.5 | [97] |
| 4. | Flax derived AC | 1649 | 500 | 4.53 | 10000 | 85 | 150,000 | 0.25 | [98] |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 5. | Flaxseed residue AC | 3230 | 398 | 61.2 | 468.8 | 98.1 | 10000 | 20 | [99] |
| 6. | Amla derived Nanoporo us Carbon | 1946 | 272 | 20 | NG | 60 | 10000 | 1 | [100] |
| 7. | Cotton derived N-P doped AC | 751 | 423 | 28.67 | 200 | 88.9 | 5000 | 5 | [101] |
| 8. | Co doped RGo Derived AC from Silk | NG | 104 | 28.31 | 78.24 | 89 | 10000 | 0.5 | [102] |
| 9. | Kenaf fiber derived AC | 693 | 217 | 6 | 215 | 95.9 | 5000 | 1 | [103] |
| 10 | Kapok fiber derived carbon | 228 | 433 | 8.92 | 10588 | 91.5 | 2000 | 20 | [104] |
| 11 | Bagasse fiber derived carbon | 2169.8 | 315.2 | 13.2 | 2500 | 94.2 | 5000 | 10 | [105] |
| 12 | Corn fiber derived AC | 1804 | 411 | 31.1 | 215 | 85.7 | 10000 | 30 | [106] |
| 13 | Ramie fiber derived carbon | 807.05 | 301.9 | 15.53 | 44600 | 87.67 | 140000 | 5.48 | [107] |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 14 | Coconut fiber derived AC | 93.67 | 145.3 | 7.78 | 3000 | 85.24 | 1000 | 5 | [108] |
| 15 | Hemp derived porous carbon | 1600 | 600 | 25.3 | 4320 | 85 | 10000 | 10 | [109] |

(Note - SSA - Specific surface area ($m^2g^{-1}$), SC- Specific capacitance ($Fg^{-1}$), ED - Energy density ($Whkg^{-1}$), PD - Power density ($Wkg^{-1}$), CR - Capacity retention %, LC- Life cycle, CD - Current density ($Ag^{-1}$), NG- Not given)

Supercapacitors have better power density as compared to other storage devices; however, they are limited by their low energy density as compared to other storage devices. To address the problem, research has been needed to increase its energy density [110]. According to the above Equation 3.3, energy density has a direct correlation to capacitance and voltage square. The voltage of supercapacitor depends on the type of electrolyte used for fabricating the storage device. Using ionic and organic liquids-based electrolytes results in higher operating voltage but this electrolyte is responsible for higher equivalent series resistance which signifies low power density [59]. Thus, the alternative approach is preferred to increase its specific capacitance, which is improved by introducing conducting polymers or doping in the electrode material like nitrogen, boron, sulfur, phosphorous and oxygen on the above or inside the natural fiber-based AC material. Heteroatom doping results in enhancing performance due to the pseudo-capacitive behaviors of supercapacitors without losing power density and life cycle [111].

## 3.7 Natural fiber-derived carbon nanomaterials characteristics.

Zhang et al. [95] in his study synthesized KOH activated carbon derived from cellulose-based corn stalk as electrode material with high specific surface area of

532.88 $m^2g^{-1}$. As depicted in the study, the morphological analysis revealed the presence of amorphous carbon peak and proved that disordered structure of carbon formed during synthesis and activation. The electrochemical analysis indicated that the specific capacitance of 139 $Fg^{-1}$, along with maintaining 1000 charge-discharge cycles. The result proved that developed natural fiber derived carbon possessed excellent electrochemical characteristics, making it suitable candidate for supercapacitor applications.

Zhang et al. [96] presents facile, green and sustainable method to synthesize N and O doped porous carbon from cotton using diammonium hydrogen phosphate as an activation agent and doping at different temperatures of 300 $^0$C, 350 $^0$C and 400 $^0$C. The prepared material at 350 $^0$C exhibit morphology with 8-10 μm obtained by SEM analysis as depicted in Figure 3.9. The TEM image demonstrated an amorphous carbon structure with irregular domains, revealing the presence of functional groups. The material had two broad peaks around $25.0^0$ and $43.5^0$ indicating amorphous structure with low degree of graphitization, which was influenced by N and O doping and oxidation temperature. The morphological characterization showed the surface area of 1022 $m^2g^{-1}$, the pore size distribution ranging from 0.4 to 4 nm, confirming the existence of micro and meso-pores, which is advantageous for ion storage and electrolyte ion transport. The XPS analysis demonstrated the incorporation of N and O into the prepared carbon structure with 5.8 wt% and 11.8 wt% respectively. The prepared material exhibits good electrochemical performance with high specific capacitance of 292 $Fg^{-1}$ at the current density of 0.5 $Ag^{-1}$. It also demonstrates excellent cycling with 87 % retention after 5000 cycles and has high energy density of 18 $WhKg^{-1}$. This strategy developed an efficient and sustainable way for synthesis of porous carbon material

using material engineering, presenting a promising potential as advanced electrode preparation for future supercapacitor applications.

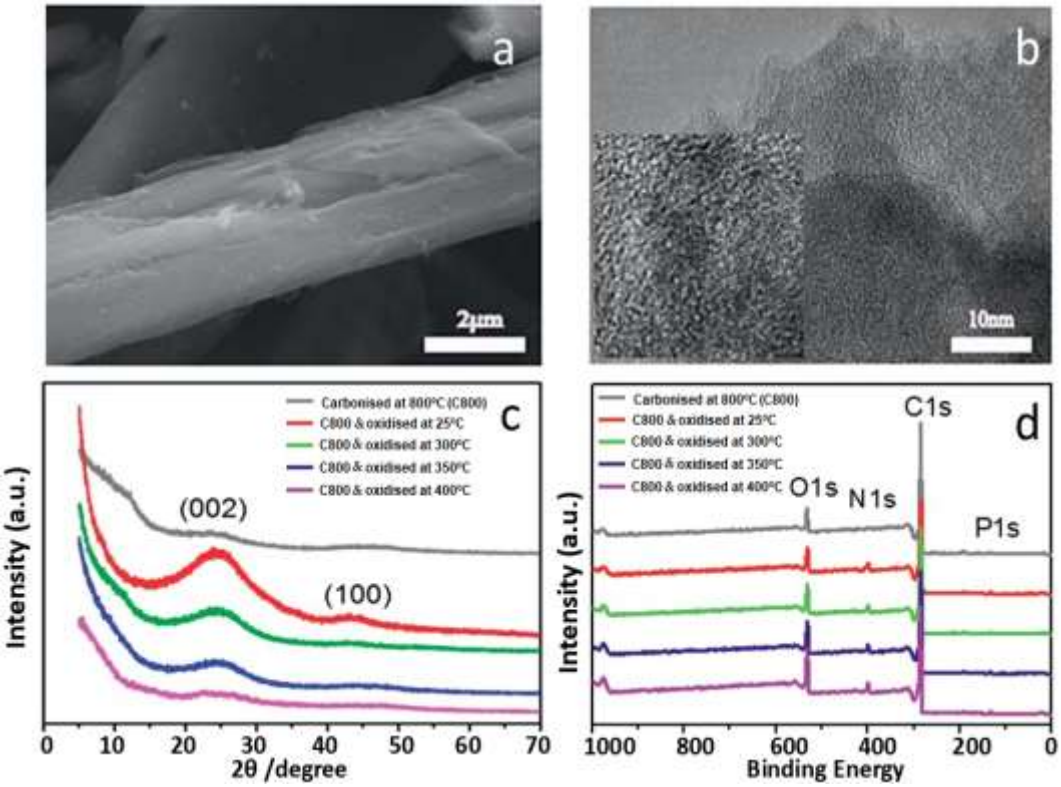


Figure 3.9 SEM , TEM, XRD, and XPS characterization of cotton derived nitrogen and oxygen doped porous carbon derived from cotton fiber (a) SEM image of cotton derived nitrogen and oxygen doped porous carbon derived from cotton fiber, (b) TEM image of cotton derived nitrogen and oxygen doped porous carbon derived from cotton fiber, (c) XRD spectra of cotton derived nitrogen and oxygen doped porous carbon derived from cotton fiber, (d) XPS spectra of cotton derived nitrogen and oxygen doped porous carbon derived from cotton fiber. Adapted from *Zhang et al*,*[112]* under creative common CC BY license.

Park et al. [86] in his research developed kenaf-based activated carbon monoliths (ACMK) with good morphological structures formed through controlled pyrolysis and activation techniques, as shown by its SEM, TEM and BET analysis as depicted in Figure 3.10. These techniques demonstrate slit-shaped mesopores and high surface area that helps to improve ion diffusion and transport with excellent charge storage. Material engineering plays an important role in optimizing temperature at 500 $^{0}$C to tailor surface area and porosity, with value of 693 $m^2g^{-1}$ and 1.31 nm,

thereby improving performance of the material. ACMK -500, pyrolyzed at 500 $^{0}$C temperature exhibited the highest specific capacitance of 217 $Fg^{-1}$ and cyclic stability of 95.9 % retention after 5000 cycles at current density of 1 $Ag^{-1}$. The presence of oxygen functional group in the material enhanced its charge storage beyond electric double layer capacitance. The inclusion of nanostructured features in prepared carbon monoliths provides a high specific surface area, highlighting the importance of nanomaterial in advancing high performance supercapacitors. With the excellent properties of energy and power densities of 6 $Whkg^{-1}$ and 215 $Wkg^{-1}$, the electrode demonstrates practical applications as sustainable electrode materials for supercapacitors.

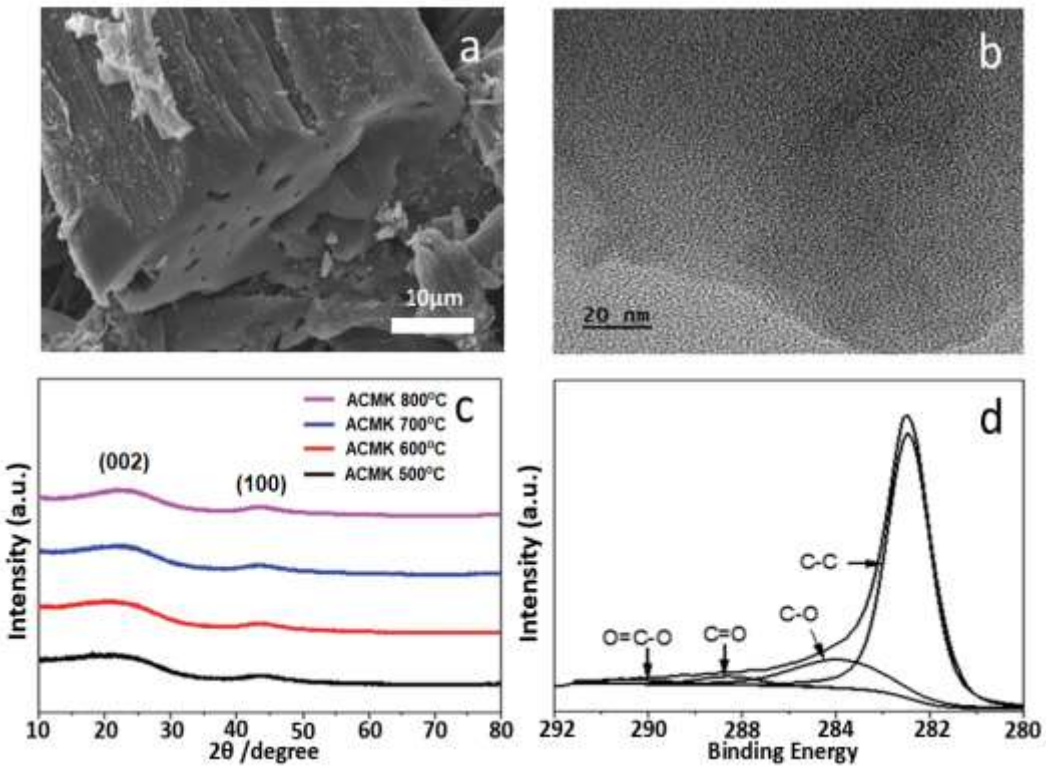


Figure 3.10 kenaf-based activated carbon monoliths (ACMK) with morphological structures (a) SEM image of ACMK, (b) TEM image of ACMK (c) XRD spectra of ACMK, (d) XPS spectra of ACMK. Adapted *Park et al. [86]* from under creative common CC BY license.

Table 3.3 Tables addressing the use and efficiency of the selected materials that are available in the market or under trials validation.

| S. No. | NF | CNM | CC | P | CS | SC | Scale | Ref |
|---|---|---|---|---|---|---|---|---|
| 1. | Coconut shell | | C > 90<br>O ~ 6<br>N < 0.5<br>H ~ 1 | High Ed | CR ~ 98% | | Industry (Kuraray yp-50F) | [97] |
| 2. | Cotton | AC | C > 99% | CD - 1<br>ED - 30.9 | upto 15000 cycles | ~87 | Powering solar lanterns | [98] |
| 3. | Corn cob | C | Hardness >90% | ED ~ 5.3<br>SR - 5<br>PD - 215 | CR ~82 % | ~314 | Lab Scale research | [99] |
| 4. | Kapok fiber | Porous carbon nanoflakes | O - 25.39 | ED – 7.85<br>PD - 1000 | CR - 94 % over 5000 cycles | 430 | Lab Scale research | [100] |
| 5. | Jute | AC | N - 0.17<br>O - 5.55 | ED - 15.44<br>PD - 402 | CR - 93.6 % over 10000 cycles | 346 | Lab Scale research | [101] |

(Note: NF- Natural fiber, CNM- Carbon material, CC - Chemical composition, P- Performance, CS - Cyclic stability, SC – Specific capacitance (F/g), AC – Activated carbon, CR – Capacity retention, CD – Current density ($Ag^{-1}$), ED – Energy density ($Whkg^{-1}$), PD – Power density ($Wkg^{-1}$), cycles – life cycles, O – Oxygen (wt%), N – Nitrogen (wt%), H – Hydrogen (wt%),C – carbon (wt%), SR – scan rate ($mVs^{-1}$))

Carbon materials which are derived from natural fibers are mostly in laboratory scale research, few industries have begun adopting sustainable methods for their production in real world applications. These growing interests highlight sustainable and environmental-friendly alternatives to conventional carbon sources. However, their transition to large scale commercial production remains challenging due to difficulties in generating consistent environmental conditions and maintaining same uniformity in experimental setups. These challenges make it difficult to produce in large amounts with the consistent quality for commercial use. Tables 3.3 addressing the use and efficiency of the selected materials that are available in the market or under trials validation.

## 3.8 Conclusion and future outlook

In this chapter, activating carbon derived from natural fibers precursors has been discussed. It has been displayed that the energy density and specific capacitance of natural fiber derived supercapacitors depend on its obtained carbon structure produced from its precursor. It has been shown that pore size, specific surface area, and heteroatom doping influence the performance of the energy storage device. Different synthesis methods and activation processes, used for making carbon electrodes from natural fiber have been discussed. The chapter highlighted key studies on the improvement of surface area, capacitance, energy density, and conductivity of electrodes derived from natural fiber for energy storage.

It was observed that the physicochemical features such as morphology, surface chemistry, surface area, and pore size distribution, of carbon derived from natural fiber have been tailored and transformed by different synthesis and activation

methods by using various activation agents. It was observed that if there is a slight modification in the active carbon structure then it will transform its electrochemical performance. Future research should be needed to develop device fabrication optimization with the impact of attaining high levels of sustainability. This chapter inspires the development of renewable carbon nanomaterials from natural fiber for next-generation energy storage systems. In supercapacitor systems, natural fiber derived carbon electrodes effectively used in both electric double layer capacitors (EDLCs) and hybrid configuration capacitors. Additionally, the heteroatom doping in natural fiber can improve pseudocapacitive behavior, further improving energy storage performance. These properties make electrodes suitable for various applications including electric vehicles, portable electronic devices, backup power systems, and flexible/wearable devices, thereby contributing to development of sustainable and efficient next generation energy storage technologies. The next chapter use these electrode material properties to predict the performance of supercapacitor using machine learning methodology.

# Chapter 4

# Performance analysis of activated carbon electrode for energy storage devices using AI and Machine learning based approach

*[The best possible storability of a supercapacitor can be achieved by the proper selection of the electrode materials. The optimized result in this chapter can be obtained by applying a Machine Learning (ML) algorithm to predict the best possible electrode material and its storability. In this chapter, ML algorithms are employed to predict the specific capacitance of activated carbon electrode-based supercapacitors, with a focus on optimizing their performance. By integrating various ML algorithms, including regression model, the study identified the most significant parameters influencing supercapacitor performance. Different ML models like Random forest, Decision tree, Linear regression, and XG boost were applied in this study. The result demonstrates that the Random forest model, with a root mean square error (RMSE) of 61.88 and a coefficient of correlation ($R^2$) value of 0.84 shows better results, as compared to other exhibits prediction in comparison to other ML models.]*

## 4.1 Introduction

Energy is an important aspect of our lives in this modern era. Without it, the situation where technological factors come into play in any area cannot be anticipated in the foreseeable future. It is essential to develop an effective energy storage device for enhancing energy efficiency and decreasing reliance on fossil fuels [113]. The scientific communities are actively engaged in developing innovative materials for energy storage solutions. Activated carbon emerged as a promising electrode material for storage devices due to its unique physical and chemical attributes. The production of activated carbon includes various synthesis methods, some of these are physical activation, chemical activation, template method, and carbonization of biomass [114][115]. The objective of these methods is to develop activated carbon that possesses a high specific surface area (SSA), pore size distribution, electrical conductivity and specific capacitance [116].

The utilization of activated carbon as electrode material for supercapacitors has the potential to substantially improve their performance by maintaining their electrochemical characteristics and their energy storage capability [117][7]. Activated carbon is a novel electrode material for energy storage devices due to its unique physical, chemical and electrochemical properties that includes its high surface area upto 3000 $m^2/g$ and adjustable pore size structure which helps in better ion adsorption and efficient storage capacity. Its cost effectiveness and compatibility with different electrolytes make it suitable for energy storage solutions. The SSA and porosity of activated carbon can provide high charge storage capacity and facilitate a better adsorption rate of ions at the electrode-electrolyte interface [118]. The material electrical conductivity provides quick charging/discharging while its physical stability provides a longer operational life

cycle. This results in improved specific capacitance and energy density of supercapacitors [118]. Therefore, activated carbon serves as a promising material for high-performance applications in supercapacitors [119]. Supercapacitors stand out among various energy storage devices due to their capacity for quick charging, good power density, and long-life cycle [120]. The specific capacitance is an important parameter for measuring supercapacitor performance because it can directly affect the energy density of the supercapacitor [121]. The specific capacitance represents the amount of charge stored per unit mass of electrode material. The accurate prediction of specific capacitance is essential for supercapacitor optimization, selection of materials, and increasing overall energy storage performance. The specific capacitance (F/g) selected as output prediction parameter because it is the most reported supercapacitor performance measurement parameter identified during literature surveying process. The consistent parameters and units in the dataset are prioritized to ensure consistency and facilitate reliability. Hence, there is a need to employ advanced techniques like ML as a robust tool for accurately predicting the specific capacitance of supercapacitors, that will enable the efficient optimization of energy storage device performance. ML techniques enable us to determine the meaningful correlation between the device output and their features without providing the physical details [122]. With the use of data-driven models, ML algorithms can predict specific capacitance quite readily without the necessity to follow traditional approaches based on experimental trial and error methods[123]. ML when trained with various datasets can form non-linear relationships between various physicochemical features of activated carbon electrodes used in the study that will influence supercapacitor performance. Thus,

ML will reduce significantly the amount of effort and time needed for a selection of parameters to develop real-life supercapacitors.

In previous studies by the researcher, ML techniques were applied in various physics and chemical-based fields for optimizations. Some of the examples are ML used for the prediction of methane adsorption performance of metal-organic frameworks (MOFS) using combined structural and chemical descriptors[124]. ML also used artificial neural network techniques to predict the specific capacitance of carbon-based supercapacitors [125]. A. G. Saad *et al.* [126] applied ML to a data-driven approach for capacitance prediction of graphene-based supercapacitor electrodes, in which four ML models were applied including K- nearest neighbours' regression (KNN), Decision tree regression, Bayesian ridge regression, and artificial neural networks. The best prediction result was obtained from the artificial neural networks model with RMSE of 60.42 and $R^2$ of 0.88, respectively. X. Yang *et al.* [127] used ML to predict the capacity of biochar through four regression models, which include Decision trees, extreme Gradient boosts, artificial neural networks, and Random forests. The extreme Gradient boost gave the best prediction with RMSE of 12.373 and $R^2$ value of 0.983. H. Su *et al.* [128] applied the four ML models for the prediction of the capacitance of electric double-layer capacitors random tree, multilayer perceptron, support vector regression and Linear regression, the best result obtained from a random tree with RMSE value of 67.72 and correlation coefficient (R) value of 0.76 compared to other models used. S. Dahidi et al. [129] in his study applied data driven approach based on artificial neural network (ANN) technique in the field of power. Y. Zhou et al. [130] in his study applied ML for life prediction of hybrid supercapacitor by using model extreme learning machine algorithm. The study identified aging degree by

monitoring equivalent series resistance and Capacitance to accurately predict the remaining useful life of supercapacitors. The aforementioned models demonstrate that the application of ML in the field of material science and supercapacitor optimization has been successfully implemented with better performance outcomes.

In this chapter, the key focus is to predict the specific capacitance of a supercapacitor made from activated carbon and its optimization by applying different machine-learning models. The method is especially beneficial for complicated device architectures. The chapter covers various aspects such as data gathering, feature engineering, model selection, and evaluation and workflow is shown in Figure 4.1. Considering the complex non-linear relationships among physicochemical features of activated carbon electrodes, ML can address complexity and is more appropriate for achieving accurate predictions than traditional empirical models. ML estimates the performance by building models based on previous experimental data. The prediction of supercapacitor performance used in the chapter involves four regression models with six input physicochemical features such as SSA, pore size, pore volume, potential window, nitrogen doping, $I_d/I_g$ ratio with output feature as specific capacitance. The regression models are Random forest, Decision tree, Linear regression, XG boost and Multilayer perceptron. The Random forest gives a high performance among other models. This technique helps to examine new high-performance electrode materials for supercapacitors. It reduces the time and cost involved in doing experimental testing for developing supercapacitors. It helps in designing supercapacitors that are customized for specific applications, including electric vehicles or grid energy storage.

### 4.1.1 Comparison of ML over traditional approaches

To predict the performance of supercapacitor by using activated carbon electrode in traditional method involves experimental measurement in which material go through synthesis, characterization and electrochemical testing, by doing material characterization in which BET analysis, morphology test such as SEM and TEM, chemical composition using XPS or elemental analysis is performed on the material while electrochemical testing includes Cyclic voltammetry (CV), Galvanostatic charge-discharge (GCD), Electrochemical impedance spectroscopy (EIS) to predict performance of energy storage device. The theoretical approach to find specific capacitance is obtained by using physical models based on material properties: The double layer capacitance model using formula given in Equation 4.1, while materials with doping consider the faradaic process which involves redox reaction to find pseudo capacitance [28], [94], [131].

Machine learning is the best approach if sufficient experimental data is available, ML model can predict the performance of supercapacitor with much better accuracy by analyzing complex, non-linear relationships of physicochemical features. The important features affecting the specific capacitance are identified and their ranges are identified through comprehensive interpretability analysis of ML model; accurate prediction of supercapacitor performance are achieved by doing material screening and analyses. The ML based predictions are systematically evaluated, reliable and robust. The ML approach includes data collection, feature engineering, selecting an algorithm, training the model and evaluating the model to predict the performance of energy storage devices [132], [133].

To predict the specific capacitance involves characterizing the activated carbon electrode material using specific surface area, pore size, pore volume, doping and

defect, after that apply experimental, theoretical or ML based approach for prediction. The process is done for optimization of activated carbon electrode properties based on insights. The combination of experimental validation with ML-driven model prediction enables efficient design of high-performance activated carbon-based supercapacitor [134][135].

## 4.2 Dataset collection

To model supercapacitors to predict performance, initial datasets with more than 500 inputs by surveying and analysing previously published research papers are formed, due to missing values and outlier, final datasets are formed by reducing large data in 100 inputs with 6 important physicochemical features. The missing values and outliers are removed manually from raw data to convert the dataset into more organized form. The data extracted have both physical and chemical features of activated carbon electrodes for supercapacitors. The features are selected according to the scientific understanding in this field. The performance of a supercapacitor depends upon the material used, the electrolyte and its operational condition which includes scan rate, potential window and its life cycle, charge/discharge rate and other factors. Optimize the input variable features for the selection procedure to conduct the analysis. The primary objective is to include all the essential physicochemical features that would impact the performance of supercapacitors, in order to build a model that can accurately predict almost all possible ranges of conditions. The features include pore volume, SSA, heteroatom doping, potential window, and Id//Ig ratio. The specific capacitance is the output feature that will predict the performance of the supercapacitor. The brief description of physicochemical features which is used as independent variables for the prediction of specific capacitance is explained below in the next paragraphs. Each

data entry contains information about the specific capacitance of activated carbon electrode. The specific capacitance is determined by the formula

$$C = \varepsilon A/d \qquad \text{Equation 4.1}$$

where A is the area of electrode, ε is the permittivity of electrode and d is the charge separation distance between the electrodes. Capacitance is increased by increasing the SSA according to above equation [136].

The supercapacitor's specific capacitance works when electrolyte ions move towards the electrode surface and its adsorption on its surface while electrolyte ion accessibility depends on the pore size of electrode material [137]. The presence of micropores/mesopores in the electrode surface provide proper channel for electrolyte ions adsorption that leads to ions diffusion in the supercapacitor [138].

The electrical conductivity depends on the crystallinity of carbon-based electrode which can be detected by Raman spectroscopy. The ($I_d/I_g$) ratio indicates the intensity ratio between the D and G bands, serving as presence of defect presence in the material. The D located at 1360 $cm^{-1}$ bands represent defects, and at G bands located at 1570 $cm^{-1}$ represent $sp^2$-hybridized carbon. The increase in ratio represents presence of defect which signify decrease in electrical conductivity. The low electrical conductivity of electrode directly affects the specific capacitance of supercapacitor [139].

In energy storage devices, higher SSA provides more active sites for electrochemical reactions. The higher SSA increases double layer capacitance by allowing ions to interact with electrode-electrolyte interface. The materials with high SSA such as activated carbon are ideal candidates for making electrode for energy storage devices [140][141][142].

The pore size helps in optimizing ion transport and storage mechanism which plays an important role in increasing specific capacitance of electrode materials. The specific capacitance of electrode material depends on how efficiently it interacts with electrolyte ions which are directly proportional to pore size and its distribution. Pore size contributes to specific capacitance by balancing ion accessibility, surface area, and diffusion efficiency. The hierarchal pore structure included into the ideal material design, by ensuring that micropores in the material help in enhancing surface area, mesopores promotes ion transport, and macropores promotes electrolyte access to improve energy storage performance [143].

The nitrogen doping in the carbon material changes their electronic properties and surface chemistry. The nitrogen group creates active sites and improves electrical conductivity which helps in increasing capacitance. The nitrogen functional groups make material more hydrophilic in nature due to which it improves electrolyte access to the surface of electrode. In the material design process for energy storage, nitrogen doping is incorporated into the carbon material via chemical treatment or by nitrogen rich precursor selection. This combined effect of nitrogen doping considerably improves energy storage performance [144].

The structural features exert direct impact on supercapacitor performance with features including porosity, SSA, dopant and their concentration. Conversely, operational features exert non-linear impact on the supercapacitor performance with features includes as potential window and electrolyte type. Among various other heteroatom doping, nitrogen have significant impact on the specific capacitance. The introduction of nitrogen functional group in the carbon material increases the capacitive properties by adding pseudo-capacitance and quantum-

capacitance. The nitrogen functional group also increases the charge mobility on carbon surface [145].

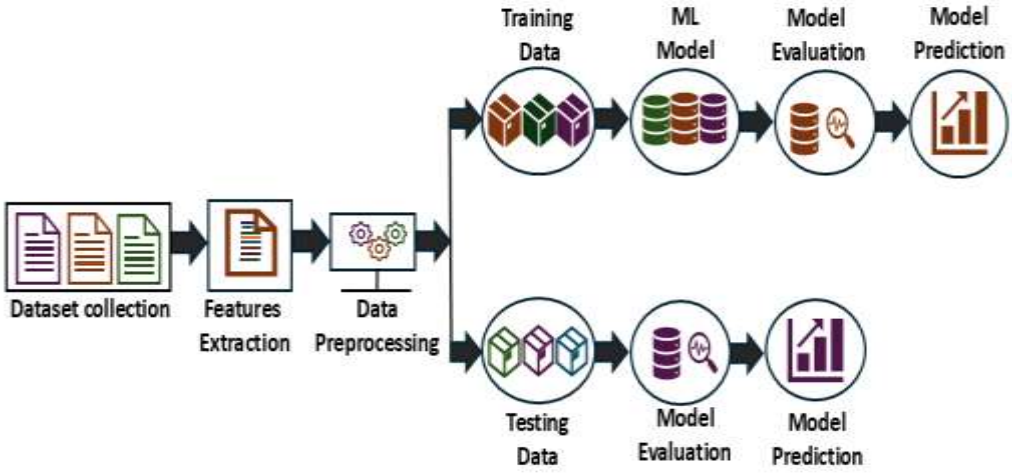


Figure 4.1 Workflow for prediction of Specific Capacitance using machine learning.

It is desired to bring up the material testing system, particularly the electrolyte. The voltage window is the primary effect that different electrolytes introduce in the electrochemical system. According to the energy equation of the capacitor

$$E = CV^2 \qquad \text{Equation 4.2}$$

E is directly proportional to voltage V square. The Equation 4.3 above depicts that the potential window has direct impact on the specific capacitance of supercapacitor. The relation between the specific feature with the specific capacitance is shown in Figure 4.2 and 4.3. The two electrodes' methods of testing for finding out the impact of specific capacitance consist of working and counter electrodes when voltage is applied, and current is collected through it. In three electrodes method of testing, working, counter and reference electrodes are present. The reference electrode is present to measure and adjust the voltage of the working

electrode without any current flow, by the system's requirements. The three-electrode method of testing is considered in the model.

After dataset collection, data preprocessing is conducted to identify missing values or inaccurate data, which must then be replaced or eliminated from the dataset [146]. Feature selection is required for the selection of specific features and then transforming it into desired form or quantitative form that will accurately classify data. It encompasses feature extraction, data dimensionality reduction, and feature selection. Upon completion of data preprocessing, then it can be subsequently utilized for training and testing the model.

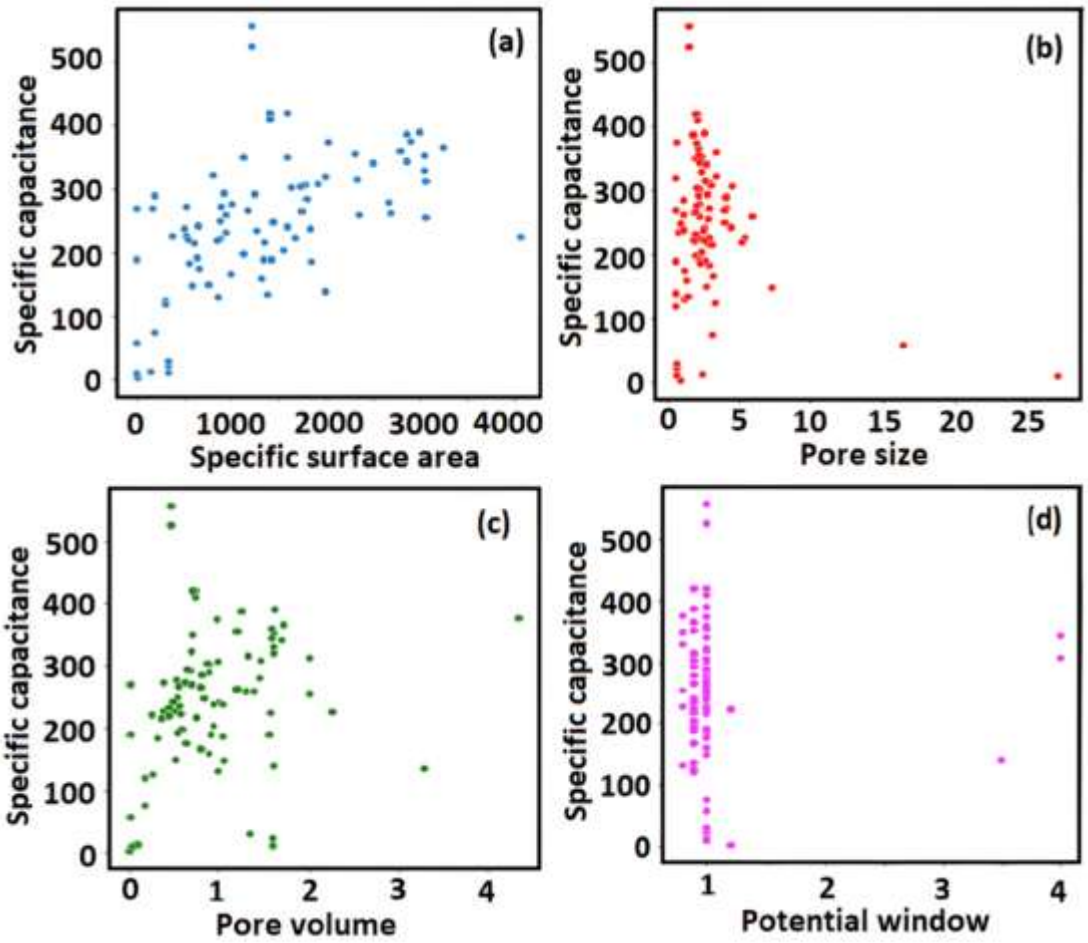


Figure 4.2 Correlation between specific capacitance in the considered dataset with different attributes (a) Specific surface area (m2/g), with Specific capacitance (F/g), (b) Pore size (nm), with Specific capacitance (F/g), (c) Pore volume(cm3/g), with Specific capacitance (F/g), (d) Potential window(V), with Specific capacitance (F/g)

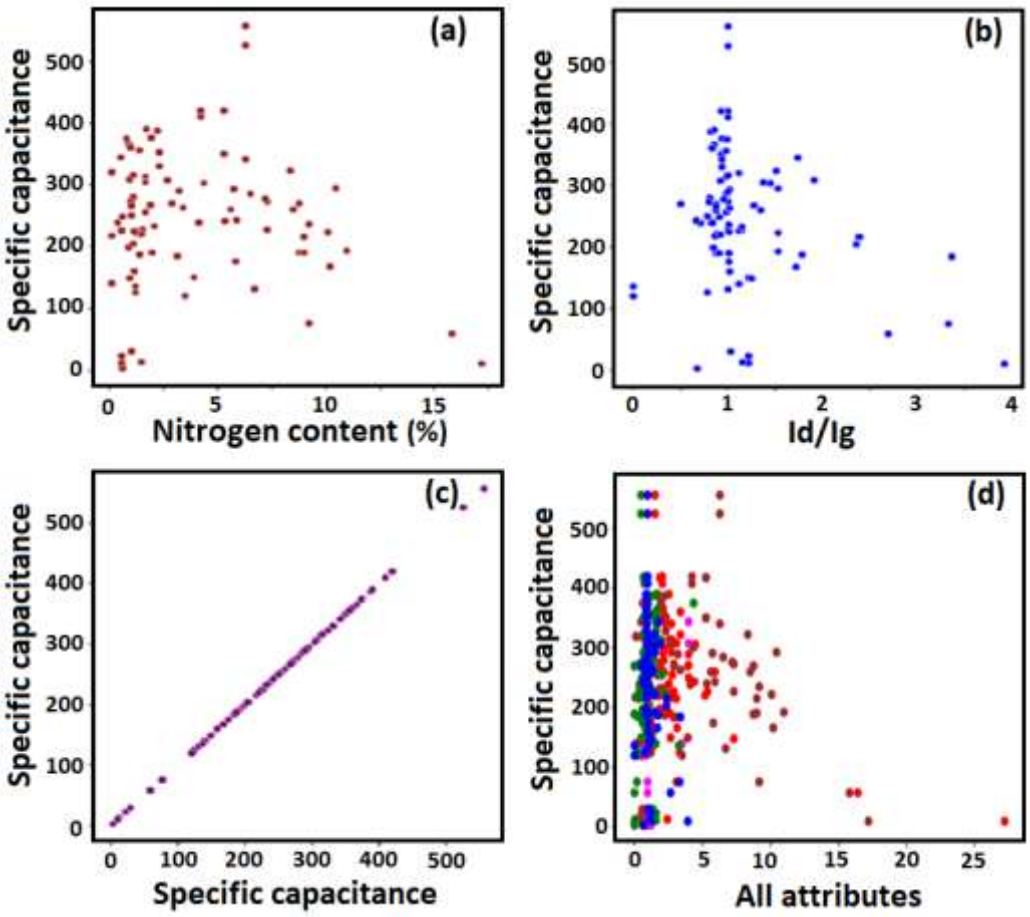


Figure 4.3 Relation between specific capacitance in the considered dataset with different attributes (a) Nitrogen content (%), with Specific capacitance (F/g), (b) $I_d/I_g$, D and G band intensities ratio of Raman spectroscopy with Specific capacitance (F/g), (c) Specific capacitance (F/g), with Specific capacitance (F/g), (d) All attributes with Specific capacitance (F/g).

## 4.3 Machine Learning Models

To conduct a comprehensive analysis of the data using ML techniques, numerous research studies were analysed throughout. The research papers that do not have comprehensive numerical data were excluded from the dataset, resulting in an initial collection of more than 500 data points. Due to some missing feature values, only 100 data points are retained for developing the model. Before fitting into the model, these data points go through preprocessing, which involves removing outliers from the dataset. In order to enable the training and testing of ML models, the prepared dataset was divided into train and test subsets, which accounted for

70% and 30% of the split of data respectively. The input features in train and test subsets are normalised and standardized, before being used for performance evaluation. The analysis of the correlation between input characteristics and the performance of supercapacitor electrodes is conducted using four different ML models. We evaluated all ML models using a ten-fold cross-validation method to assess their performance across several train/test splits. The models applied in the study are Linear regression, Decision tree, Random forest, Xtreme Gradient boost, and random tree. Moreover, hyperparameters also influenced the performance of the model. Hyper-parameters are employed during cross-validation to help prevent the challenges of overfitting and underfitting in the ML model. The hyperparameters were tuned in the RF model using grid search cross validation method with the cv value 10 and n_estimator value 100, in DT model it is tuned using max_depth of tree. The present methodology evaluated the influence of various train/test splits on the model performance. The evaluation of the chosen models is conducted by examining their prediction results using performance measures metrics like $R^2$, RMSE, and mean absolute error (MAE). The brief description of models is explained in the next section.

Linear regression model delineates a linear relation between dependent feature and independent feature or variable. The Linear regression includes the intercept and linear term, as well as the product of independent features from the input dataset [41]. Regression estimated the value of unknown parameters by minimizing the sum of squares of difference between the actual value and the predicted value by linear function of given features.

Multiple Linear regression with several input features, the Equation 4.4 generalizes to:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \qquad \text{Equation 4.5}$$

Where:

$\hat{Y}$ = Predicted value of dependent variables.

$X_1, X_2, \ldots X_n$ = Independent variables.

$\beta_0$ = Intercept.

$\beta_1, \beta_2, \beta_3 \ldots \beta_n$= coefficients for each feature

Decision tree is supervised learning algorithm which is tree-like structure in which each node represents input variable, each branch indicating a feature value, and each leaf node represent outcome or prediction. The node closer to root node has a more significant impact on the target value. An inherent benefit of employing Decision trees is their capacity to efficiently process both numerical and categorical input [147].

The prediction at a leaf node is mean of the prediction values in the node

$$\hat{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad \text{Equation 4.6}$$

Where:

$\hat{Y}$ = Predicted value of dependent variables

$Y_i$ are actual values in the node

Random forest is based on ensemble learning techniques which describe a process of combining multiple Decision tree classifiers to solve a complex algorithm to enhance the model performance. Every tree in Random forest trained on bootstrapped sample of dataset and selected subset randomly for Decision tree split of training data. Every tree made its own predictions. These predictions of each tree are then averaged to determine the output. Random forests avoid overfitting as

compared to in case of Decision tree [43]. It also works on same equation(ii) shown above for Decision tree.

Extreme Gradient boost applies the Decision tree algorithm to the given dataset. This algorithm sequentially applies Decision trees to the dataset. Independent features are assigned weights, which are then fed into the Decision tree to predict the output. The independent features are allocated weights and subsequently inputted into a Decision tree, which subsequently makes predictions on the outcome. The wrong weight predictions will result in the input of these feature variables into the second Decision tree. Next, these Decision tree models are merged to determine the accurate prediction result[148]. It has objective function that comprises two components: the loss function and the regularization term. Loss function measures how perfectly the model fits into the training dataset. The regularization factor introduces to penalize complexity to prevent overfitting.

The equation for objective function (*Obj*) is:

$$Obj = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{k} \Omega\,(f_k) \qquad \text{Equation 4.7}$$

Where:

$L(y_i, \hat{y}_i)$ is loss function that measure difference between actual and predicted value.

$\Omega(f_k)$ is regularization function to penalizes the complexity of the model

$f_k$ indicate individual regression trees

k is total number of trees.

Random trees are used in the ensemble technique, which makes model predictions by averaging the output of independent base models. Random forest builds each Decision tree using a subsample or bootstrapped dataset. The algorithm confirms

that each Decision tree in a forest train on a distinct subset to provide the best split of data and improve the prediction accuracy of the model [149].

## 4.4 Performance Metrics

All the models are optimized using cross validation of value 10 while the performance prediction of model, the prediction results are compared to the specific capacitance output value using these performance metrics given in below equations [150]:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)}{\sum(y_i - \bar{y}_i)} \quad \text{Equation 4.8}$$

$$RMSE = \sqrt{\frac{\sum_1^n y_i - \hat{y}_i)^2}{n}} \quad \text{Equation 4.9}$$

$$MAE = \sqrt{\frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}} \quad \text{Equation 4.10}$$

$\bar{y}_i$ is the mean of $\hat{y}_i$ value is the prediction value and $y_i$ is the actual value. RMSE closer to value zero shows better prediction result. $R^2$ value closer to one represent good predictions.

RMSE is a metric used to measure the accuracy of models by estimating the difference between actual and predicted values. Smaller RMSE values indicate better model performance, and the model better fits the data. RMSE values are scale dependent so they should not be compared across different datasets with different ranges and units. As shown in its formula RMSE squares the residual obtained before averaging the output values, this gives higher weightage to large errors. It gives an idea about how much error value on an average can be expected in future predictions. If the RMSE of prediction model is 65, it means the error between predicted specific capacitance and actual specific capacitance is about 65 F/g.

MAE is the averaging of predicted value and actual value of the regression model. It gives an idea about how far the predicted output is from the actual output. The smaller MAE value is the indication of good model performance. It provides a measure of prediction accuracy. MAE is a reliable and interpretable metric for evaluation of regression ML models, where main concern is about overall error magnitude. MAE metric for prediction is often preferred for real time applications where interpretability of matters as in this case of specific capacitance prediction. It is less sensitive to outliers because it doesn't square the errors. MAE have the same unit as output feature here specific capacitance which is in F/g.

The correlation coefficient $R^2$ is the key metric measure the proportion of variance in the output which is explained by the model predictions. The $R^2$ value when approaches to 1, the model perfectly fits all the variance in the output values which is sign of good performance prediction. The low $R^2$ value signifies the model struggles to determine the variability due to missing physicochemical features, noise in the data or not selecting the best regression model. The $R^2$ is easy to understand by its value in terms of variance, explained like if $R^2 = 0.8$ means model is 80 % variance in the output leaving 20% remained unexplained. The comparison of $R^2$ across different datasets with different variance can be considered misleading, as $R^2$ depends on the variability of output feature.

## 4.5 Results and Discussion

The correlation coefficient is used to quantify the linear correlation between two features, with the value ranging between -1 and 1. As illustrated in Figure 4.4, the numerical value in the square box represents the correlation coefficient value while color reflects the intensity and orientation of the linear correlation between the intersecting variables. Within the matrix, the positive correlation coefficient value

indicates that the values of two features tend to increase or decrease together, whereas the negative correlation coefficient value indicates the increase of one feature is related to the decrease in another feature. A coefficient value near 1 or -1 between specific capacitance and specific features suggests that the feature may be valuable for predicting specific capacitance in a ML model [126].

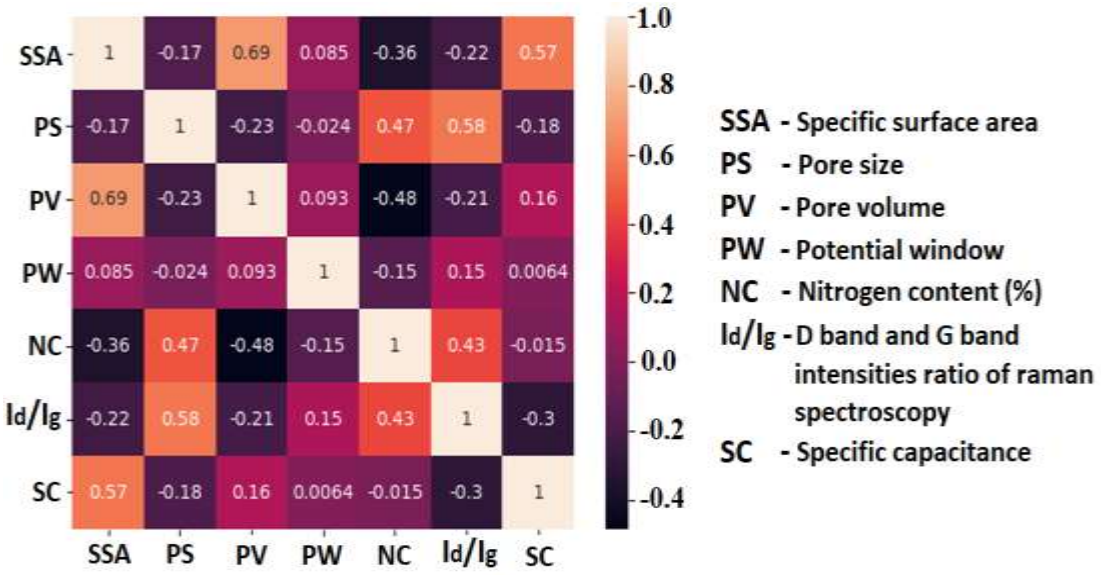


Figure 4.4 Pearson correlation between physicochemical features of activated carbon electrode material of supercapacitor in the dataset to determine feature importance.

Table 4.1 Statistical results of physicochemical features in dataset.

| | SSA | PS | PV | PW | NC | $I_d/I_g$ | SC |
|---|---|---|---|---|---|---|---|
| Count | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| Mean | 1359.82 | 2.581 | 0.9948 | 1.044 | 3.7371 | 1.1373 | 240.006 |
| Std | 989.175 | 3.138 | 0.6908 | 0.502 | 3.5771 | 0.569 | 115.513 |
| Min | 1.4 | 0.54 | 0.0018 | 0.8 | 0.08 | 0 | 2.1 |
| Max | 4070 | 27.2 | 4.337 | 4 | 17.2 | 3.91 | 556 |

The correlation study shows that the specific capacitance increases as the SSA increases, the specific capacitance increases as the pore size decreases, and the specific capacitance increases as the defects decrease. The analysis shows that specific capacitance has a strong correlation with SSA, moderate correlation with pore size, pore volume, nitrogen doping content of activated carbon electrode of supercapacitor and low correlation with remaining features. The analysis depicts that the SSA, pore size and nitrogen doping content are important features for enhancing the specific capacitance of supercapacitors. These findings also depict that these physicochemical features have an independent effect on specific capacitance. The statistical result of physicochemical features as shown in Table 4.1 After completing of the correlation analysis, the dataset divided into two subsets, 70% of the data was allocated in training the model while remaining 30% were reserved in testing the model. In the study, the six physicochemical features of supercapacitor in dataset were considered as independent variable, with specific capacitance being the dependent variable. The Figure 4.8 presents the prediction results of ML models. Upon examining its correlation coefficient and other metrics, it is evident that the XG boost prediction is comparatively lower than other models. Among all the other five ML models as depicted in Table 4.3, XG boost produced the weakest result as depicted by lowest $R^2$ and highest RMSE value. By comparison, the Random forest and Decision tree models exhibited notably superior prediction accuracy, seen from their reduced RMSE and higher $R^2$ values. The Figure 4.5 depicts the actual and prediction value of best ML model. The RF is the best predicted model compared to other model used for evaluation and prediction of specific capacitance of activated carbon-based supercapacitor. The Figure 4.5 depicts that the actual and predicted values of specific capacitance is

near to the best fit line shown by red color in the Figure 4.5 and green dashed line is the perfect prediction line for the model which proves that the error between the real and prediction value by the ML model is low. The Linear regression model predicted specific capacitance formula obtained from Weka software platform depict in Equation 4.11 as

$$SC = 349.7302\ x\ SSA + 134.6803\ x\ NC + (-188.8257 x\ I_d/I_g) + 189.4629$$

…….Equation 4.12

Where SC is specific capacitance

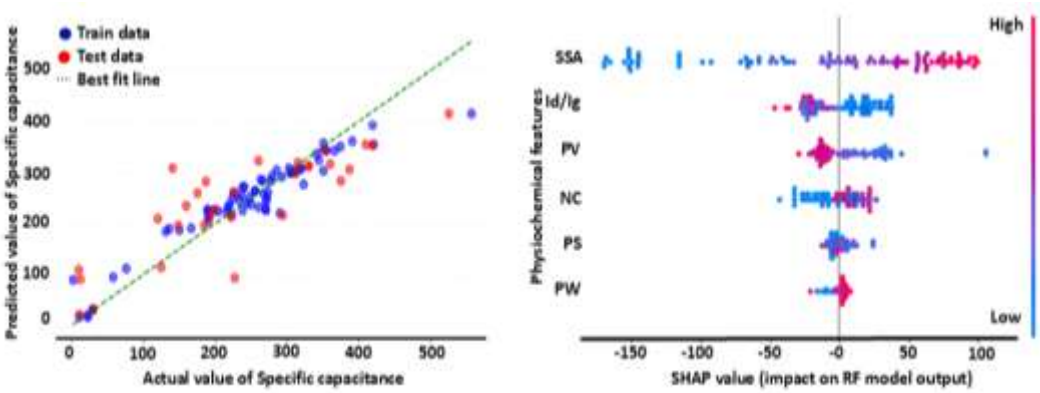


Figure 4.5 Feature analysis of best ML models (a) The plot of predicted specific capacitance from considered dataset compared to actual values of Specific Capacitance of best RF model from machine learning model. (b)The plot of sensitivity analysis (sHAP value) impact on best RF models

Table 4.2 The feature importance score of ML models given below:

| Physicochemical Features | RF | DT | LR | XG boost |
|---|---|---|---|---|
| Specific surface area | 0.56 | 0.588 | 0.100 | 0.313 |
| Pore size | 0.06 | 0.02 | 4.00 | 0.065 |
| Pore volume | 0.126 | 0.162 | 64.25 | 0.198 |
| Potential window | 0.027 | 0.03 | 9.625 | 0.029 |
| Nitrogen content | 0.09 | 0.044 | 8.87 | 0.159 |
| $I_d/I_g$ | 0.127 | 0.146 | 46.87 | 0.233 |

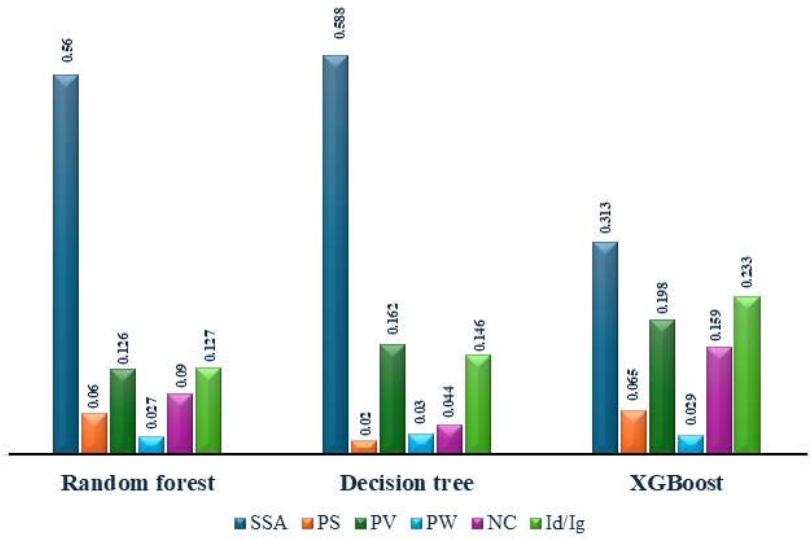


Figure 4.6 Relative contributions of most significant physicochemical features influencing supercapacitor performance in ML model, Random forest feature (RF) importance chart. Decision tree (DT) feature importance plot. Extreme Gradient boost (XGB) feature importance chart. The value of chart is given in Table 4.2.

The Figure 4.6 shows physicochemical features ranked according to their impact on the specific capacitance in the order of the ML regression model. The Table 4.2 shows how different features have different impacts depending on the ML model for prediction of output specific capacitance. The feature importance score shown in Table 4.2 plays an important role in ML model interpretation and optimization of the model. It helps in identifying which physicochemical features are considered most influential in making predictions. The ML model reduces overfitting by removing irrelevant patterns from less important features. In the models like Random forest and XG boost, the most important features will help in decision making, improve model prediction accuracy. Figure 4.7 represents the feature importance selection of each physicochemical features of the RF model. The result shows that the SSA have the most significant feature impacting the specific capacitance of supercapacitor with the value of 56% in RF model followed by the $I_d/I_g$ ratio, accounting to 6% followed by other physicochemical features which is

shown in Table 4.2 of feature importance score of different ML models considered in the study.

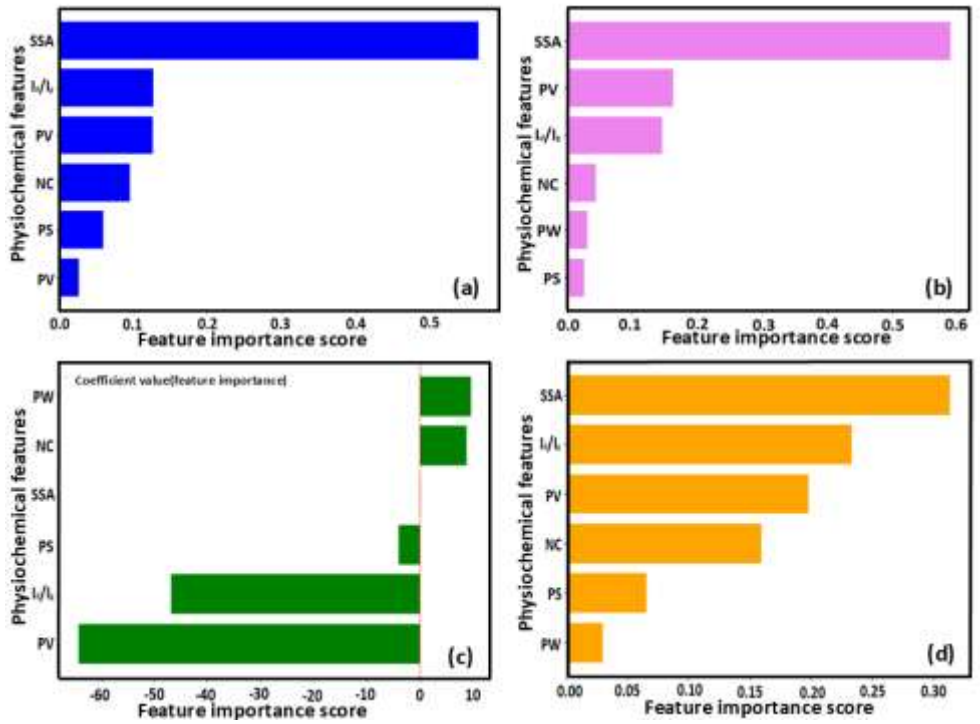


Figure 4.7 Relative contributions of most significant physicochemical features influencing supercapacitor performance in ML model (a). Random forest feature (RF) importance plot (b). Decision tree (DT) feature importance plot (c). Linear regression (LR) feature importance plot (d). Extreme Gradient boost (XGB) feature importance plot

Table 4.3 Regression machine learning model metrics results

| Model | RMSE | MAE | Correlation coefficient | Correlation coefficient % |
|---|---|---|---|---|
| Random forest | 61.8826 | 46.6854 | .8448 | 84.48 |
| Random tree | 72.1946 | 49.855 | .8044 | 80.44 |
| Decision tree | 80.303 | 57.5366 | .8271 | 82.71 |
| Linear regression | 88.4444 | 67.433 | .6451 | 64.51 |
| XG boost | 100.339 | 78.583 | .7018 | 70.18 |
| MLP | 109.580 | 72.6075 | .561 | 56.1 |

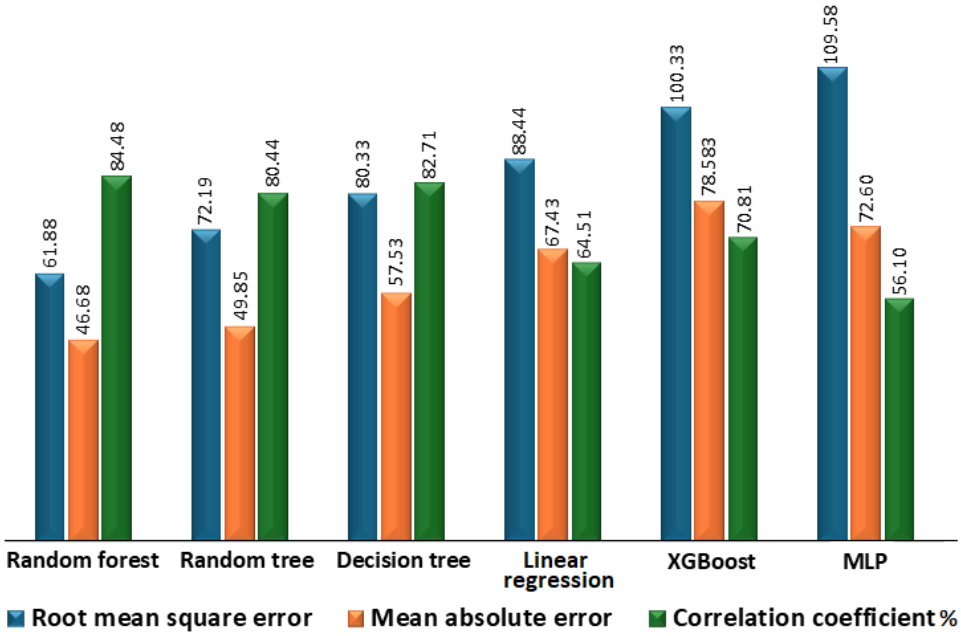


Figure 4.8 Regression Machine Learning Model Results

### 4.5.1 Why did Random forest outperform other models?

Random forest (RF) works on combining prediction of multiple Decision trees to get the resultant output of the model. Each tree on the RF model is trained on separate subsets of dataset acquired from the original dataset and select input features randomly to form diverse set of models to predict the output of each separate model before predicting the final output. This approach reduces overfitting and increases the robustness of models. In RF, if two or more physicochemical features influence the specific capacitance in non-linear manner then it can place appropriate split at different level of Decision trees. It doesn't assume predefined relationships between input features. RF can highlight which physicochemical features have more contribution to non-linear patterns which help in capturing intricate relationships to improve performance. RF achieved smaller RMSE due to its unique nature such as its robustness to overfitting and outliers, ability to handle non-linear relationships among physicochemical features and specific capacitance.

These qualities of RF are effective for minimizing residual errors and reducing its RMSE value.

Table 4.4 The comparison table from other literature for prediction of specific capacitance of supercapacitor by using ML model

| S. No. | ML model | RMSE | MAE | $R^2$ | References |
|---|---|---|---|---|---|
| 1. | Random forest | 38.131 | - | 0.6891 | [151] |
| 2. | Linear regression | 54.90 | - | 0.3555 | [151] |
| 3. | Decision tree | 67.62 | | 0.577 | [152] |
| 4. | XG boost | 25.50 | | 0.79 | [153] |
| 5. | Random forest | 35.715 | 29.268 | - | [153] |
| 6. | Decision tree | 4.25 | 3.3248 | 0.9426 | [154] |
| 7. | MLP | 2.040 | 1.547 | 0.9868 | [154] |
| 8. | Linear regression | 2.75 | 2.11 | 0.9759 | [154] |
| 9. | Decision tree | 15.009 | - | 0.205 | [155] |
| 10. | Random forest | 67.95 | - | 0.562 | [155] |
| 11. | XG boost | 12.373 | - | 0.983 | [155] |
| 12. | Random forest | - | 93.94 | 0.5929 | [156] |
| 13 | MLP | - | 133.2 | 0.577 | [156] |
| 14. | Random forest | 43.96 | - | 0.75 | [157] |
| 15. | Decision tree | 55.63 | - | 0.60 | [157] |
| 16. | XG boost | 40.27 | - | 0.79 | [157] |

The other regression models struggle when there are complex or non-linear dependence of physicochemical features on specific capacitance while RF requires less hyperparameter tunning for optimal performance and is easier to predict the

output. XG boost can easily overfit and be sensitive to outliers if not properly regularized especially when applied to small dataset because it works on boosting which involves sequential building Decision trees where each tree depends on previous Decision tree. It focuses on reducing the residual error which will lead to fitting the noise in the dataset and also leading to large errors and higher RMSE values degrading prediction accuracy during testing of the data. The XG boost requires preprocessing of input features like label encoding which can lead to loss of information and more memory usage.

## 4.6 Conclusion and future work

This chapter examines machine learning models for the prediction of specific capacitance of activated carbon electrodes used in supercapacitors. The experimental data which were employed for model training and evaluation were obtained from the previously published research papers. The physicochemical features considered in the analysis included SSA, potential window, pore size, pore volume, nitrogen content, and Id/Ig ratio with the specific capacitance as the output for the ML regression models. Among all other models, Random forest demonstrated the best predictive accuracy result with the lowest RMSE value of 61.88 and a high $R^2$ value of 0.84 respectively.

These models ranked according to the predicted specific capacitance in decreasing order as Random forest, Decision tree, Linear regression, and extreme Gradient boost. Table 4.3 depicts the performance metrics of different ML models applied for prediction of the specific capacitance of activated carbon-based supercapacitor. Table 4.4 compares the work with other literature work done for prediction of specific capacitance of supercapacitor by using ML model. The study revealed that SSA, pore size and nitrogen content were the most influential features in the ML

model. In conclusion, this chapter presented a data-driven methodology to predict the performance of supercapacitors and highlight the significant impact of physicochemical and electrochemical factors on it. To hope that this chapter provides insight into a comprehensive study of supercapacitor performance. and it would be useful for the design and fabrication of activated carbon electrodes for supercapacitors.

The ML model can be used in optimization of development process of energy storage devices, especially supercapacitors. The model can accelerate innovation, development, reduce costs, improve performance and optimize material and synthesis parameters, considerably reduce the time and effort invested in the trial-and-error process in the laboratory. The ML model trained for energy storage devices can be used for predicting other material properties, like conductivity, power density, energy density, life cycle assessment. The model also helps in identifying different operating conditions like behaviour of electrolyte in varying voltage range and temperature range by introducing new physicochemical and electrochemical input features to improve device reliability. In the next chapter, we will use the ML model which is trained in experimental data to predict the specific capacitance of activated carbon using combinational approach to find the range and value by applying classification and regression models. In conclusion, the application of ML will significantly improve the research and development of energy storge material, enhanced its optimization process.

# Chapter 5

# ML Classification and regression-based technique to predict the specific capacitance of activated carbon electrode for supercapacitors

[*In this chapter, machine learning (ML) based algorithms are employed to establish correlation relationship between electrode material features and specific capacitance of supercapacitors. The comprehensive dataset of more than 500 peers reviewed previous research publications was compiled on activated carbon-based supercapacitors to predict specific capacitance using ML based performance metrics. This chapter uses combinational approach of class prediction and value prediction using ML algorithms. The various ML algorithms used for classification and regression methods are Random forest, XG boost, Decision tree, support vector machine, K nearest neighbors', Ada boost, Gradient boost and others. The prediction-based ML models developed in this chapter offer a pathway to reduce the time and effort required for design and evaluation of electrode material for supercapacitors.]*

## 5.1 Introduction

With the growing need of energy and increasing demand to mitigate climate change problem, the demand to develop efficient energy storage has attracted more attention [158]. Among various energy storage devices, supercapacitors have attracted attention as promising candidates due to their characteristics such as power density, long life cycle, and wide temperature range [11], [159]. The performance and selection of electrode materials play an important in determining the efficiency of supercapacitor [64], [160]. As an important component of supercapacitors architecture, electrode plays an important role in determining key performance parameters, including physicochemical behaviors, power density, energy density, charge/discharge capability and cycling stability [161]. Various carbon-based material architectures such as activated carbon, carbon nanotube, carbon nanofibers, and graphene have been extensively investigated as promising materials for energy storage applications [162], [163], [164].

Owing to supercapacitor's fast-charging capability, they can store and deliver reasonable amounts of energy, making them suitable for applications that have sudden power surges [165], [166]. Supercapacitor can bridge the energy storage gap which persists between conventional capacitors and electrochemical batteries by addressing the discrepancy in energy density and power density [163]. There are two types of adopted way used to achieve the desired performance of supercapacitor. The most used approach involves the increase of specific capacitance by expanding the potential window, thereby getting the highest possible energy density [13], [167]. In supercapacitor, energy storage relies on non-faradic process in which direct storage of electric charge occurs at electrode/electrolyte interface via reversible adsorption of ions on electrode surface, thereby increase the

surface area and porosity [36], [168]. Activated carbon serves as a promising electrode candidate in maintaining specific surface area and porosity in energy storage applications [81]. The specific capacitance of activated carbon-based electrodes can be significantly enhanced by optimizing the porosity and structural physicochemical properties. Furthermore, the performance of supercapacitors can be further enhanced through advanced modification, like introduction of heteroatom doped functional groups, wettability, minimizing defects and incorporation of pseudo-capacitive behaviors in electrode material [90], [169]. The development and application of activated carbon electrode material are important in attaining these performance metrics, as shown in many previous literature [170], [171]. However, increasing supercapacitor performance, particularly in specific capacitance, remains a challenge. This problem is crucial to get resolved for maximizing their energy storage capabilities and developing next generation energy storage systems [7].

Various physicochemical and electrochemical factors affect the performance of activated carbon-based electrodes for supercapacitors. Machine learning (ML) is employed to optimize the key influencing factors and predict the specific capacitance based on input features of activated carbon electrodes, thereby accelerating the optimization process of activated carbon-based electrodes for supercapacitor applications [128]. In recent years, ML has emerged as an important tool in modelling and forecasting complex, nonlinear relationships between electrochemical systems [172]. The various ML models, including both supervised and unsupervised learning approaches, have been increasingly applied to predict the electrochemical and physicochemical performance of energy storage devices with improved accuracy and efficiency [173], [174]. These data-driven methods

facilitate the optimization of key features, and device performance metrics, thereby enabling the fabrication of electrodes for energy storage devices which helps to reduce the time needed for empirical trial and error methods [151]. The previous published research has used the application of ML in forecasting the supercapacitor performance and its characteristics such as surface area, power density, energy density, and capacity retention [175], [176]. Reddy et al. [177] in their study used ANN and back propagation learning algorithms to predict the specific capacitance of carbon-based supercapacitor by using input parameters such as voltage, Id/Ig, nitrogen and oxygen doping content, pore size and surface area, and observed that artificial neural network model has better result in prediction of performance of supercapacitor. Su et al. [128] used ML algorithms for prediction of supercapacitor performance by using four algorithms including Linear regression, support vector regression, random tree, and multi-layer perception in which random tree gave the best results. Zhang et al. [178] extracted input features such as discharge current, terminal voltage and discharge capacity from energy storage device to employed ANN model to predict battery life estimation. Rahimi et al. [179] developed a predictive model for forecasting the specific capacitance of heteroatom rich carbon electrodes in electric double layer capacitors (EDLCs). The model correlated electrochemical performance with important physicochemical features such as surface area, pore size and concentration of nitrogen and oxygen functional groups and operational features like voltage window. Table 5.1 describes other literature using ML for prediction modelling of different electrode materials used for fabricating energy storage devices.

In this chapter, the data-driven methodology using ML based approach to predict the performance of activated carbon-based supercapacitor as shown in Figure 5.1.

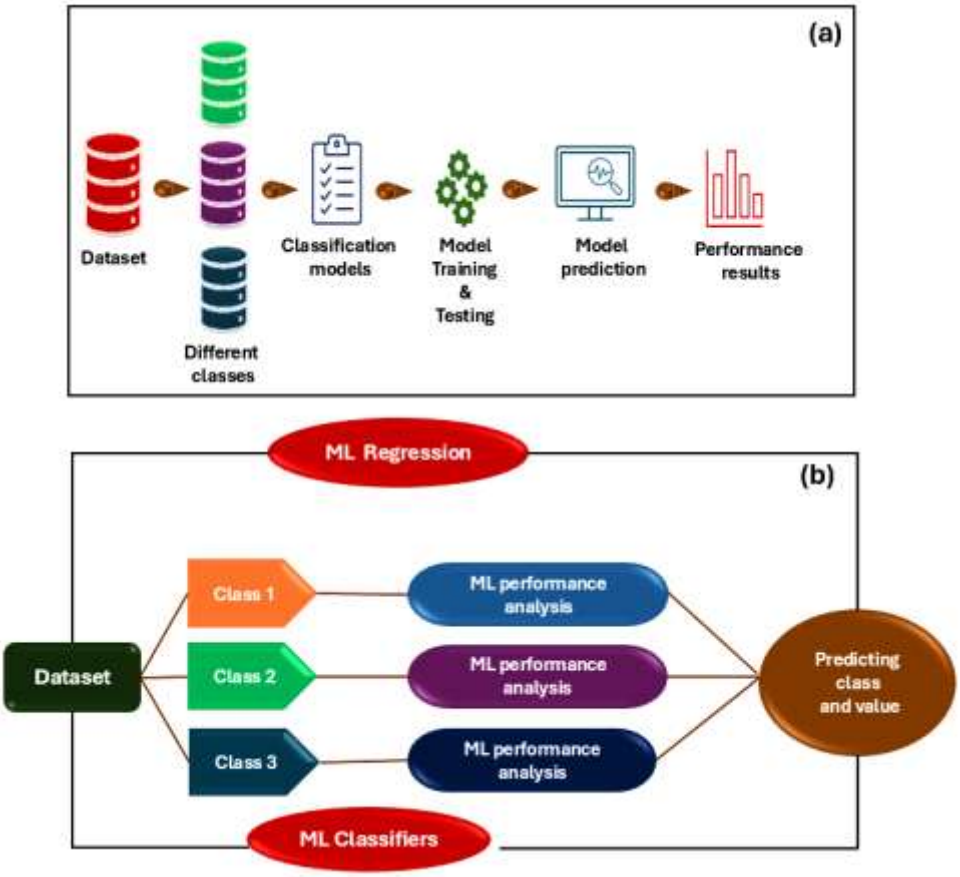


Figure 5.1 Model description for implementation of ML algorithms, (a) Block representation of implementation of ML model by using dataset to model evaluation using performance metrics (b) Schematic representation of method used in this chapter for implementation of classifier method for class predictions and regression method for specific capacitance value predictions.

The comprehensive database was compiled by systematically analyzing over 500 previously published research articles in this field to model the specific capacitance of supercapacitors. After data collection, preprocessing of the dataset is done to handle missing values and outliers. After preprocessing, feature engineering is done to prioritize important features for the study. The final dataset is formed with 558 inputs by considering 6 influential input features and 1 output feature. The important physicochemical features include specific surface area, pore size, pore

volume, Id/Ig, nitrogen content, potential window identified as influencing parameters to affect supercapacitor performance. These features were used as input features for development of ML models, with specific capacitance as output variable. This chapter aimed at developing prediction ML models for identification of specific capacitance within the considerable range. More than six ML models are built for classification and regression analysis to predict supercapacitor performance.

- **Comparison of the work with other literature used ML in the field of supercapacitor**

Table 5.1: Literature review of different AI and ML models for supercapacitor performance predictions.

| S. No. | Material | features | Best ML metrics | Data set | Output feature | Best model | Ref |
|---|---|---|---|---|---|---|---|
| 1 | Zinc-ion hybrid capacitors using Carbon | 9 | RMSE - 4.88 mAh/g<br>MAE -3.62 mAh/g<br>$R^2$ score -0.986 | 624 | C | Light GBM | [180] |
| 2 | Biomass derived carbon | 16 | $R^2$ score - 0.984<br>RMSE - 8550.39<br>MAE - 6036.93 | 119 | PD | XG boost | [181] |
| | | 16 | $R^2$ score - 0.922,<br>RMSE - 9.27<br>MAE - 5.414 | | ED | Light GBM | |
| 3. | Carbon based | 4 | $R^2$ score - 0.997<br>RMSE - 1.55<br>MAE - 6.24 | 130 | C | XGB | [182] |
| 4. | | 22 | $R^2$ score - 0.95 | 324 | SC | | [183] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $MgCo_2O_4$ electrodes | | RMSE - 111.83<br>MAE - 68.25 | | | XGB-RFE_XGB | |
| 5. | CNT- based electrode | 9 | $R^2$ score - 0.82<br>RMSE - 36.31<br>MAE - 68.25 | 146 | SC | GBR | [184] |

Note: XGB-RFE_XGB (XG boost-Recursive Feature Elimination - XG boost), RF – Random forest, XGB – Extreme Gradient boost, Light GBM – Light Gradient boost, GBR – Gradient boost Regression, SC - Specific capacitance (F/g), C - Capacitance (F/g), ED - Energy density (Wh/kg), PD - Power density (W/kg).

## 5.2 Description of dataset collection

The data driven model has been adopted to predict the specific capacitance of supercapacitor. The data driven methods require measurement data for accurate predictions rather than complete formulas. This study gathered essential physicochemical features of activated carbon-based electrode for supercapacitors from earlier published research articles. More than 500 research papers were selected from previously published articles related to activated carbon electrode for supercapacitor to create datasets with an initial of 558 set of datapoints [157]. The goal of this study is to show how physicochemical features influence the specific capacitance of supercapacitors as negative or positive. After gathering all the data from research articles, different sets of input features of activated carbon-based supercapacitors are identified that affect the specific capacitance of supercapacitor. The features selection plays an important role in the application of machine learning models. The data contains informative values for all the features. When more feature values are missing in the set of data, then this row is excluded [133]. The accuracy of models mostly depends on the completeness of input features. The input features include: (i) specific surface area (SSA) in $m^2/g$, (ii) pore size (PS) unit in

nm, (iii) pore volume(PV) in $cm^3/g$, (iv) potential in V, (v) nitrogen doping content (NC)in atm wt%, (vi) $I_d/I_g$ ratio (Intensity ratio of D and G bands peak in Raman analysis) and output features are specific capacitance (SC) in F/g. The relationship between supercapacitors' features and specific capacitance is nonlinear. The supercapacitor's performance is measured in terms of SC due to its significant role in storing charge.

To ensure consistency and reduce variability in model results, this study is restricted to AC electrode-based supercapacitor using aqueous electrolyte. Aqueous electrolyte is widely used due to high ionic conductivity, allowing ions to move quickly, allowing more accessible surface area, and have excellent compatibility with porous carbon structures. By using same type of electrolytes, variation get reduced due to different testing conditions, allowing more accurate investigation of the relationship between material features and performance results. This approach highlights the importance of controlling experimental variables to achieve reliable and interpretable model results.

Following dataset preparation, preprocessing of data is done to remove inconsistent and missing terms, ensuring data reliability and consistency. The important and relevant features are selected and converted into numerical form which is suitable for model training. The feature scaling is applied to minimize redundancy and improve model performance. The final cleaned and structured dataset is then processed for training and testing the ML models.

The specific capacitance prediction is performed in two ways: the first method involves (i) the specific capacitance class prediction, and the second method involves (ii) the specific capacitance value prediction. In the specific capacitance class prediction, specific capacitance is categorized into three classes (i) The

specific capacitance value is less than equal to 138 F/g is classified as Class 1 (ii) The specific capacitance value between 139 to 223 F/g is classified as lass 2 (iii) The specific capacitance value greater than 223 F/g is classified as Class 3. These methodologies provide both range and values for more precise predictions [185].

Table 5.2 Statistical analysis of all features considered in the chapter for performance prediction.

| | SSA | PS | PV | PW | NC | $I_d/I_g$ | SC |
|---|---|---|---|---|---|---|---|
| Count | 558.00 | 558.00 | 558.00 | 558.00 | 558.00 | 558.00 | 558.00 |
| Mean | 1.166 | 1.7148 | 0.8679 | 1.2960 | 1.3293 | 0.5557 | 188.13 |
| Std. | 0.883 | 2.6165 | 0.8618 | 0.7925 | 2.8755 | 0.7981 | 104.47 |
| Min | 0.00 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 25% | 0.461 | 0.0000 | 0.2380 | 1.0000 | 0.0000 | 0.0000 | 120.00 |
| 50% | 1.0167 | 0.9000 | 0.6900 | 1.0000 | 0.0000 | 0.0000 | 176.50 |
| 75% | 1.6410 | 2.5225 | 1.2950 | 1.0000 | 1.1375 | 0.9975 | 247.75 |
| Max | 4.0730 | 27.20 | 5.9100 | 4.0000 | 19.800 | 4.5400 | 587.00 |

The class prediction of specific capacitance is divided into three classes and their counts on each feature value is shown in Figure 5.2. The ML models used for value prediction features with reference to the output feature in which range the specific capacitance values lies is described in Figure 5.3. All features considered in this research article are listed in Table 5.2 along with their statistical analysis.

## 5.3 Model Development

Several machine learning classification and regression algorithms applied in class prediction and value prediction models. These algorithms are used to improve

accuracy and give precise predictions. The ML models applied in this chapter are described below for both classification and regression analysis.

### 5.3.1 Decision tree

It is a supervised learning algorithm which is used for both classification and regression problems models. The model works on taking decisions and their possible outcomes as tree-like structure where every node of tree represents test dataset feature, and each branch of tree represents its decision result, and each leaf node represents a class or regression value. Each tree divides the data into subsets depending on feature values to increase the similarity of output feature in the final subsets.

The goal of the Decision tree is maximizing its class purity and reducing the error in predicting the class and regression value.

**For classification:**

The class prediction at each node using measures such as Gini impurity and information gain.

$$Gini = 1 - \sum_{k=1}^{K} p_k^2 \qquad \text{Equation 5.1}$$

For node with $K$ classes and class with $p_1, p_2, \ldots . p_K$.

Information gain $(Ig)$ is determined by the function:

$$Entropy = -\sum_{k=1}^{K} p_k \log_2 p_k \qquad \text{Equation 5.2}$$

$$Ig = Entropy(root) - \sum_{i=1}^{n} \frac{N_i}{N} * Entropy(child) \qquad \text{Equation 5.3}$$

$N$ = total sample at root, $N_i$ is sample at child $i$, $n$ is the number of child nodes.

**For regression tree:**

The value prediction is determined by errors such mean square error and mean absolute error. At node, with $N$ number of samples $y_1, y_2 \ldots . y_N$

$$Mean\ square\ error = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2 \quad \text{Equation 5.4}$$

$$Mean\ absolute\ error = \frac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}| \quad \text{Equation 5.5}$$

$$\bar{y} = \frac{1}{N}\sum y_i \quad \text{Equation 5.6}$$

### 5.3.2 Random forest

Random forest is an ensemble learning method that works on building different number of decisions trees and combines their results or average to predict the class or continuous value. Each Decision tree is trained on random subsets of training data. At every Decision tree, features are selected randomly. For classification, it predicts the output on majority votes from different trees. For regression, it takes the average prediction value. The mathematical computation of Random forest is the same as the Decision tree. It reduces overfitting compared to Decision tree. It works well on high dimensional data, outliers and noise [186].

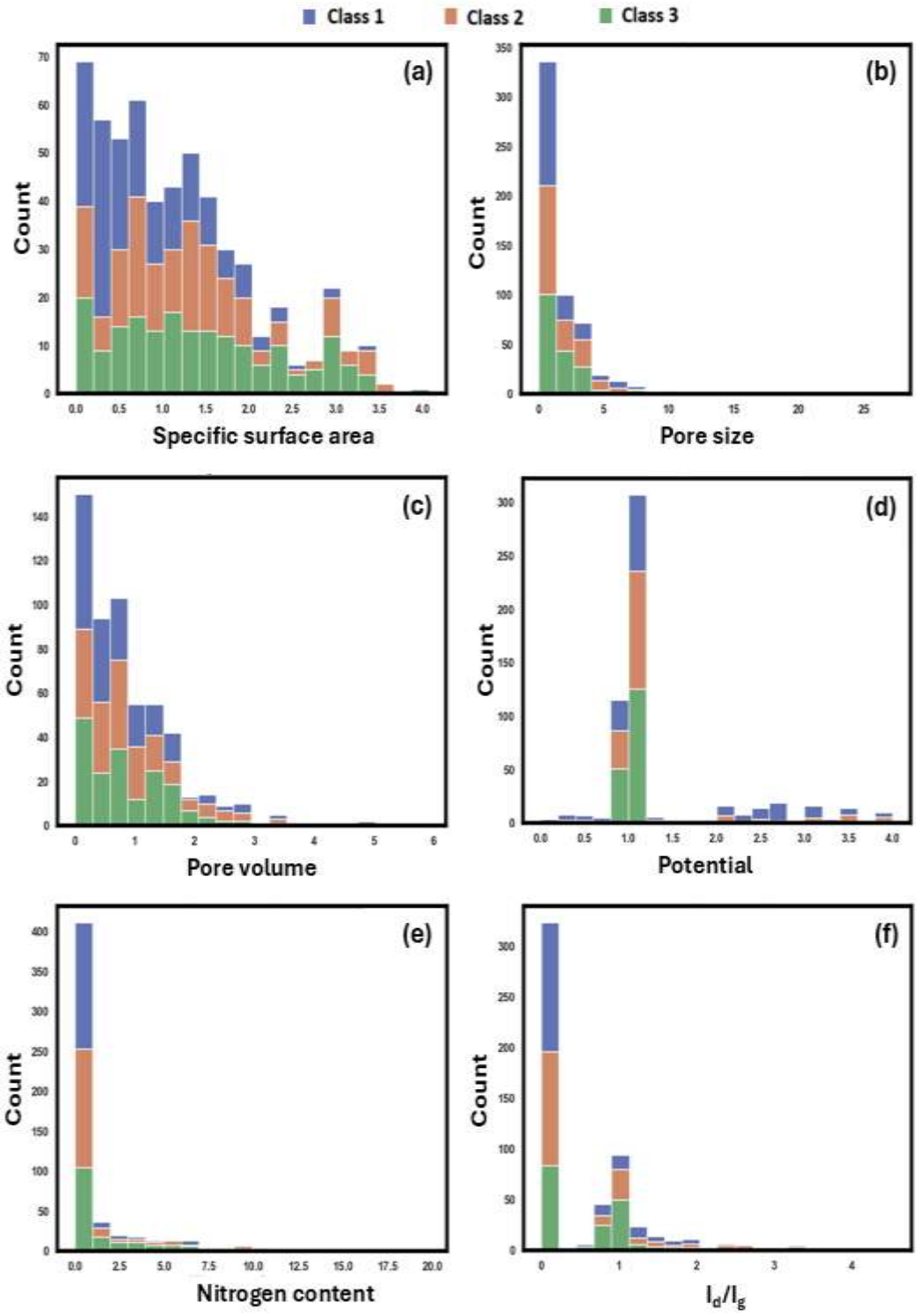


Figure 5.2 Features used in the study for class prediction analysis: a) specific surface area count with reference to the classes, b) Pore size count with reference to the classes, c) Pore volume count with reference to the classes, d) Potential window count with reference to the classes, e) Nitrogen content count with reference to the classes, f) Raman analysis ratio of D and G band ($I_d/I_g$) count with reference to the classes.

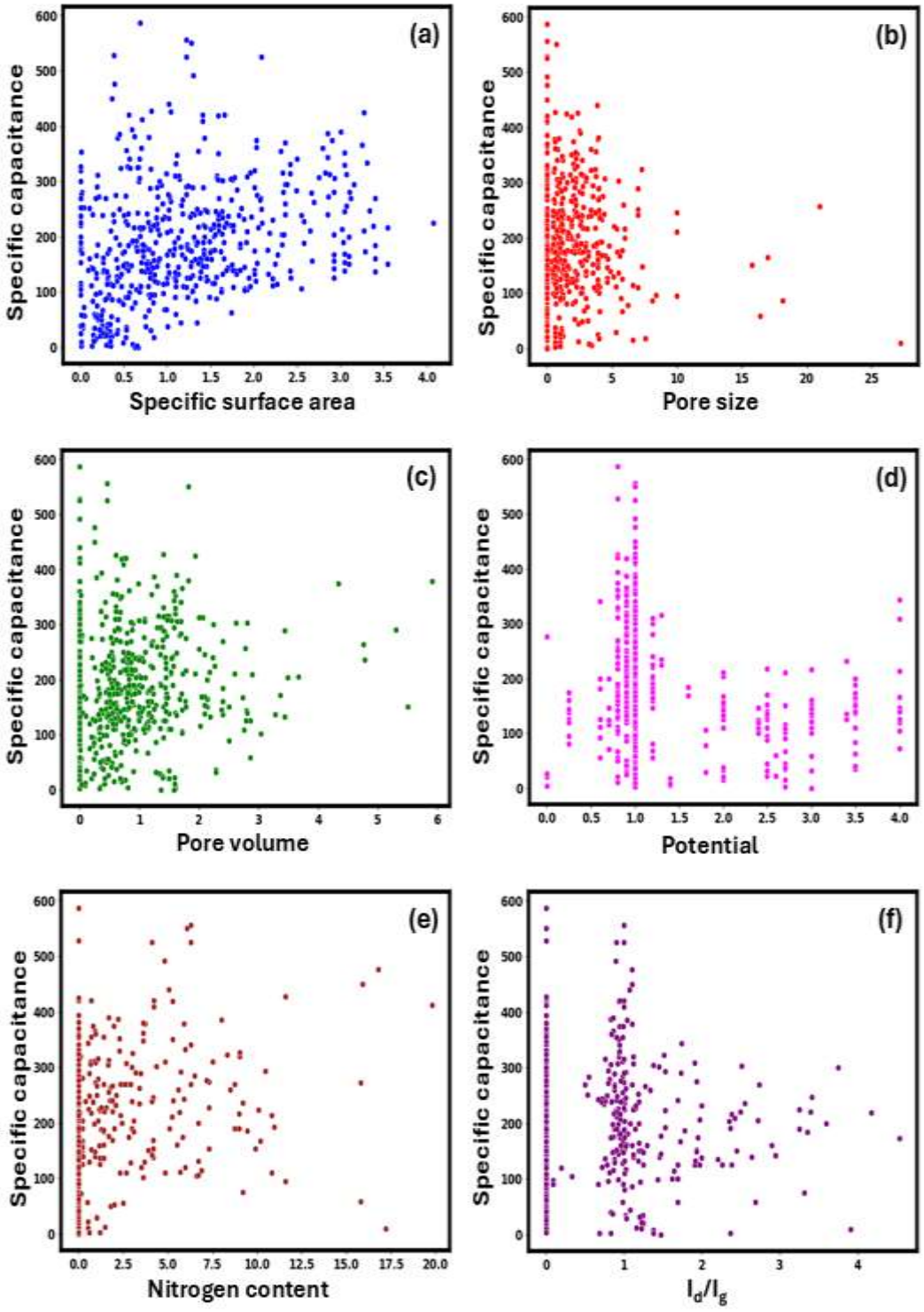


Figure 5.3 Features used in the study for value prediction analysis: a) specific surface area with reference to the specific capacitance, b) Pore size with reference to the specific capacitance, c) Pore volume count with reference to the specific capacitance, d) Potential window count with reference to the specific capacitance, e) Nitrogen content count with reference to the specific capacitance, f) Raman analysis ratio of D and G band ($I_d/I_g$) count with reference to the specific capacitance.

### 5.3.3 Ada boost

AdaBoost is a boosting algorithm that works by combining multiple weak learners to create a strong model which is classified as strong learner. It is called adaptive boosting because the weights are assigned many times to each point (features) until higher weights are assigned to incorrectly classified points. Now the higher weight points are given more importance in the next model, it will keep running and training the model until error is reduced [187]

For classification:

The training data is defined as: $P = \{(x_i, y_i)\}_{i=1}^{n}$

Let initialize weights: $w_i^{(1)} = {}^{1}/_{n}$ , $i = 1,2, \ldots . n$

For boosting $b = 1,2, \ldots . , B$

Weak learner as $h_b(x)$ using weight $w_i^{(b)}$

Find error using function

$$\varepsilon_b = \frac{\sum_{i=1}^{n} w_i^{(b)} . II(h_b(x_i) \neq y_i)}{\sum w_i^{(b)}} \qquad \text{Equation 5.7}$$

Model weight equation is:

$$\alpha_b = \frac{1}{2} \ln\left(\frac{1-\varepsilon_b}{\varepsilon_b}\right) \qquad \text{Equation 5.8}$$

Sample weight is calculated using:

$$w_i^{(b+1)} = w_i^{(b)} . \exp\left(-\alpha_b y_i h_b(x_i)\right) \qquad \text{Equation 5.9}$$

Normalize the $\sum w_i^{(b+1)} = 1$

The final output prediction function is:

$$H(x) = sign\left(\sum_{b=1}^{B} \alpha_b h_b(x)\right) \qquad \text{Equation 5.10}$$

For regression:

P = $\{(x_1, y_1), (x_2, y_2), \ldots . (x_n, y_n)\}$

The weights are assigned as: $w_i^{(1)} = {}^1/_n$

For boosting $b = 1,2, \ldots ., B$

The Weak Regressor as $h_b(x)$ using weight data. The error is computed for each feature as:

$$\varepsilon_i = |y_i - h_b(x_i)| \qquad \text{Equation 5.11}$$

The error normalization is done using:

$$\widetilde{\varepsilon_i} = \frac{\varepsilon_i}{max_j \varepsilon_j} \qquad \text{Equation 5.12}$$

The model error is computed using function as:

$$\varepsilon_b = \sum_{i=1}^{n} w_i^b . \widetilde{\varepsilon_i} \qquad \text{Equation 5.13}$$

Model weight function:

$$\beta_b = \frac{\varepsilon_b}{1-\varepsilon_b}, \text{After that } \alpha_b = \ln\left(\frac{1}{\beta_b}\right) \qquad \text{Equation 5.14}$$

Sample weight is normalized using:

$$w_i^{(b+1)} = w_i^{(b)} . \beta_b^{1-\widetilde{\varepsilon_i}} \qquad \text{Equation 5.15}$$

The final output is predicted using:

$$H(x) = \frac{\sum_{b=1}^{B} \alpha_b h_b(x)}{\sum_{b=1}^{B} \alpha_b} \qquad \text{Equation 5.16}$$

### 5.3.4 Gradient boost

The Gradient boost algorithm works in an iterative manner where model is trained successively to resolve the residual errors made by the previous model. The final prediction is obtained by summing the contribution of each prediction output of all the models [188].

The training data is defined as: $P = \{(x_i, y_i)\}_{i=1}^{n}$

The objective here is to find a function that helps to minimize the loss

$$Loss = \sum_{i=1}^{n} L(y_i, F(x_i)) \qquad \text{Equation 5.17}$$

Let initialize the model with constant:

$$F_0(x) = \arg\min_c \sum_{i=1}^{n} L(y_i, c) \qquad \text{Equation 5.18}$$

For boosting let assign boosting round as $m = 1, 2, \ldots M$

The gradient equation is:

$$g_i^{(m)} = -\left[\frac{\partial\, L(y_i, F(x_i))}{\partial\, F(x_i)}\right]_{F=F_{m-1}} \qquad \text{Equation 5.19}$$

The base learner $b_m(x)$ to the residuals $g_i^{(m)}$

The multiplier finds by equation:

$$y_m = \arg min_\gamma \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma\, b_m(x) \qquad \text{Equation 5.20}$$

The model is evaluated using functions

$$F_m(x) = F_{m-1}(x) - \delta y_m b_m(x) \qquad \text{Equation 5.21}$$

Where $\delta$ is the learning rate $\delta \in (0,1]$

## 5.4 Model Evaluation

The accuracy and performance of machine learning models are evaluated and compared through statistical performance metrics which are root mean square error (RMSE), mean absolute error (MAE), kappa statistics, and correlation coefficient ($R^2$). These performance metrics perform as an indicator for assessing model performance by quantifying the difference between actual and predicted values obtained from the ML models [189]. Within a considered dataset, each actual value denoted as $a_i$ and predicted value as $b_i$.

Among these metrics, RMSE is informative as it highlights large errors due to deviation square, thereby serves as important indicator for model accuracy. It is defined as the square root of average square difference between actual and predicted values. The RMSE Equation 5.22 is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - b_i)^2} \qquad \text{Equation 5.22}$$

Where n represents total number of observations in considered dataset. The lower value of RMSE indicates a high degree of precision of the model, reflecting minimal deviation between predicted and actual value of specific capacitance.

In addition to RMSE, MAE is also employed as performance parameter metrics. MAE provides a linear and interpretable measure of absolute deviation from actual values. MAE is defined as arithmetic means of absolute difference between each actual value $a_i$ and predicted value $b_i$. The expression for MAE is given in Equation 5.23:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|a_i - b_i| \qquad \text{Equation 5.23}$$

A low value of MAE indicates good performance of the model. It is widely used in regression analysis for comparative assessment of ML models.

The third statistical metric is correlation coefficient ($R^2$) measure the proportion of variance in output feature predicted by using correlation with the independent features. The value close to 1 suggests good performance. The value is important in determining how well the model fits with unseen data. The mathematical Equation 5.24 is:

$$R^2 = 1 - \frac{\sum_{1=1}^{n}(a_i - b_i)^2}{\sum_{1=1}^{n}(a_i - \overline{a_i})^2} \qquad \text{Equation 5.24}$$

Where $\overline{a_i}$ is the mean of actual value of specific capacitance.

The kappa statistics metric is employed to evaluate the classification performance of ML models. Kappa statistics give measure of consistency between the predicted and actual class obtained by the model. Kappa statistics adjusts for random agreement, making it informative and reliable metric for evaluating classification ML models [190]. The mathematical Equation 5.25 is:

$$k_s = \frac{p_o - p_e}{1 - p_e} \qquad \text{Equation 5.25}$$

Where, $p_o$ is observed agreement between predicted and actual classification

$p_e$ is expected agreement by chance, which obtained from marginal distribution of each class in confusion matrix.

## 5.5 Results and discussion

### 5.5.1 Correlation coefficient analysis

Correlation analysis is an important tool for determining the relationship between features with other features, helping to understand patterns and correlations within the dataset [191]. Figure 5.4 shows the correlation coefficient heatmap for

independent input features with other features and with output feature which is the specific capacitance. Figure 5.3 offers a clear representation of the relationship between various features and specific capacitance of the supercapacitor. By analyzing the correlation coefficient values, which features have high correlation with specific capacitance and which feature have low correlation with the specific capacitance. The correlation coefficient value lies between -1 and 1.

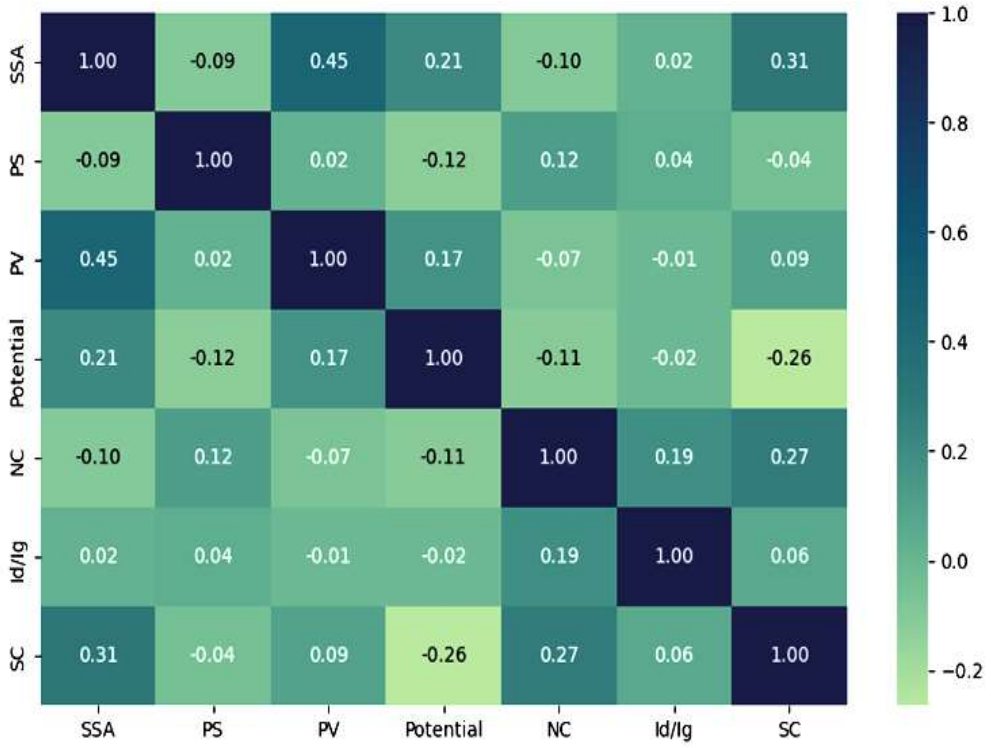


Figure 5.4 Correlation coefficient analysis between two physicochemical features of supercapacitor.

Above Figure 5.4, shows that the specific surface area (SSA) has positive correlation with the specific capacitance (SC), it means these SC tends to increase with SSA, its coefficient value is 0.31 which show the strong correlation with the SSA. Similarly, specific capacitance has a strong correlation with nitrogen content % (NC) and it has a low correlation with pore volume (PV) and Id/Ig. The distribution of NC in the considered dataset is mostly concentrated at lower values

as depicted in Fig 3, reflecting real world experimental data from published literature where high doping of NC in carbon matrix is not frequently reported. As this study focused on the dataset compiled from published experimental works, it only captures the current research pattern. The correlation analysis shows that it has a negative correlation with potential and pore size (PS), which means specific capacitance values increase with decrease in potential and pore size. The study shows that the SSA, potential and NC are important physicochemical features for determining specific capacitance of supercapacitor. The analysis successfully identifies NC as important features, contributing to estimate specific capacitance along with other physicochemical features such as SSA.

### 5.5.2 The class prediction model

The class prediction model used classification algorithm technique to predict specific capacitance class [156]. In this model, classification method is used to divide specific capacitance of supercapacitor into three classes, class 1, class 2, and class 3 as also described in above section. Figure 5.2 shows all the relevant features that affect the specific capacitance of supercapacitor and their separation into three major classes. For class prediction, various machine learning classification models are used to predict the class in which SC lies, models are Decision tree (DT), Random forest (RF), Ada boost (ADB), SVM, and Gradient boost (GB). The performance of classification models as depicted in Figure 5.5 are evaluated based on MAE, RMSE, and Kappa statistics using hyperparameter optimization technique with cross validation value of 10. The kappa statistics value should be high, the RMSE value and MAE value should be low for better performance of the models. Table 5.3 depicts the performance metrics of the class prediction models. By analysis of its metrics, it is depicted that the ADB shows the best result among all

the ML models because of low MAE and RMSE value. After this, RF and support vector machine (SVM) performances are better by analyzing its performance parameters. All the model's performance parameters are close to each other, signifying that all models work almost similarly. ADB works best for the activated carbon electrode material for supercapacitor dataset due clean well processed dataset without noise, and automatic feature selection characteristics through weighting process.

Table 5.3 Machine learning data driven classification models for class prediction.

| S. No. | ML Models | MAE | RMSE | Kappa statistics |
|---|---|---|---|---|
| 1. | Random forest (RF) | 0.46 | 0.74 | 0.370 |
| 2. | Decision tree (DT) | 0.535 | 0.855 | 0.345 |
| 3. | Gradient boost (GB) | 0.508 | 0.818 | 0.360 |
| 4. | Ada boost (ADB) | 0.419 | 0.687 | 0.407 |
| 5. | Support vector machine (SVM) | 0.464 | 0.744 | 0.376 |

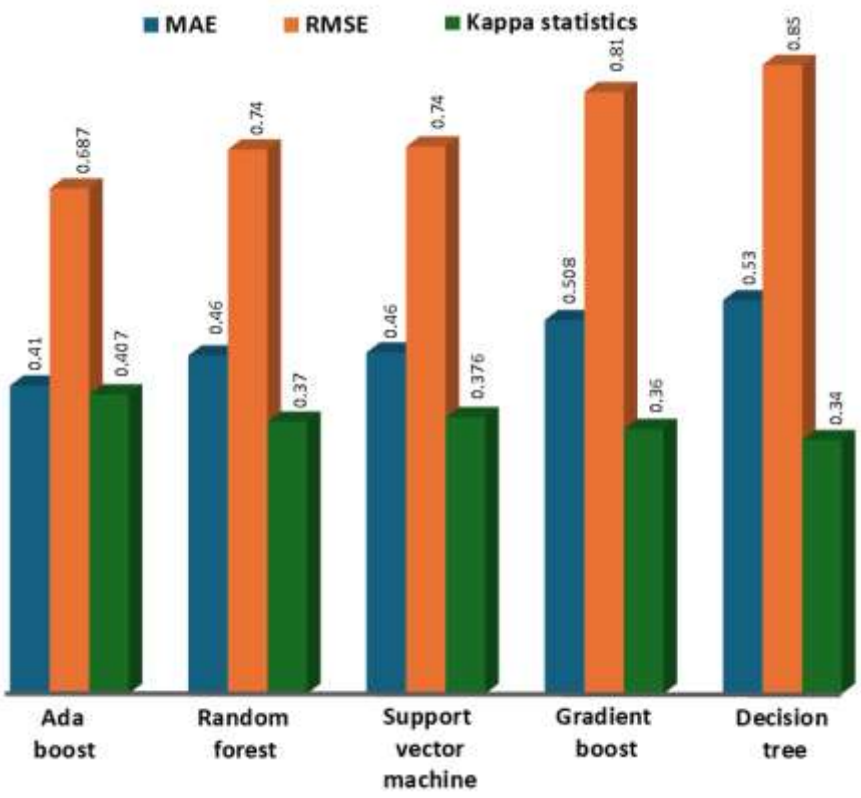


Figure 5.5 ML classifier models for class predictions with its performance metrics results

Subsequently, the models are evaluated from different class perspectives to analyze which ML classification model performs best on each class basis. The most suitable model is selected through more statistical performance metrics such as precision, recall, and F1-score to ensure high reliability and accuracy to assess the effectiveness of models as shown in Table 5.4 below.

Precision used to measure the ratio of correctly predicted positive occurrences out of all occurrences that were predicted as positive by the ML model. TP is a true positive occurrence where model correctly predicts the class. The FP is the occurrence where model is not correctly predicting the positive class. Precision influences the accuracy of positive prediction of classifier model for the output class.

$$Precision = \frac{TP}{TP+FP}$$ Equation 5.26

Recall also known as true positive rate or sensitivity. It measures the ability of the model to correctly identify all relevant occurrence of specific class. It is measured as the proportion of true positive prediction to total of actual positive occurrences. Thus, recall quantifies the ability of model to identify all relevant occurrences of given class. FN is the false negative occurrence where models failed to identify the positive class.

$$Recall = \frac{TP}{TP+FN}$$ Equation 5.27

F1-score is the means of precision and recall that provides a single metric that is used to balance both metrics. It is beneficial in cases when there is not even distribution of class or when a balance between false positives and false negative is at a priority.

$$F1 - score = 2.\frac{Precision*Recall}{Precision+Recall}$$ Equation 5.28

Table 5.4 Machine learning classification models weighted average of statistical metrics for class performance.

| Models | Precision | Recall | F1- score |
|---|---|---|---|
| ADB | 0.61 | 0.61 | 0.61 |
| RF | 0.58 | 0.58 | 0.58 |
| SVM | 0.590 | 0.580 | 0.570 |
| GB | 0.571 | 0.571 | 0.565 |
| DT | 0.563 | 0.562 | 0.569 |

As for class prediction, the next task is to evaluate the ML models for each class individually. Accordingly, ML classification models are tested from different class perspectives. In this phase, the performance and effectiveness of classification models are evaluated for each class, and most suitable model is identified depending on its statistical metrics like precision, recall and f1-score.

As shown in Table 5.4, weighted average of all statistical classification metrics highlights ADB as superior models as compared to other models. The result for individual classes mostly aligned with this observation, except recall for class 1 as shown in Table 5.5. The increase recall for class 1 in the RF model suggest it is more efficient in identifying values lying to this class when SC lies within its range. Thus, for class 1 prediction, RF is suitable choice to predict correctly classified instances. Class 3 exhibit high precision, recall and f1-score, indicating high occurrences of true positives and very low occurrences of false positives. These results demonstrated that these ML classifiers models are effective in identifying class 3, thereby ensuring high reliability in predicting the specific capacitance of supercapacitor at higher value which is greater than 223 F/g in class 3 which also proves that the supercapacitor performs better at higher specific capacitance. In last, considering overall class prediction performance, ADB based model demonstrate best efficiency in class prediction.

Table 5.5 Machine learning precision, recall and f1-score of each class for best models among all models.

| Models | ADB | | | RF | | |
|---|---|---|---|---|---|---|
| | Precision | Recall | F1 score | Precision | Recall | F1 score |
| Class 1 | 0.60 | 0.62 | 0.61 | 0.55 | 0.65 | 0.59 |
| Class 2 | 0.52 | 0.51 | 0.52 | 0.51 | 0.44 | 0.48 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Class 3 | 0.71 | 0.71 | 0.71 | 0.69 | 0.69 | 0.69 |
| Weighted Avg | 0.61 | 0.61 | 0.61 | 0.58 | 0.58 | 0.58 |

### 5.5.3 The value prediction regression ML models

The value prediction used machine learning regression models to predict the specific capacitance value [192]. Again, various machine learning regression models are used, which are Decision tree, Random forest, SVM, Gradient boost and XG boost (XGB). XG boost is selected as a value prediction model in place of Ada boost for class prediction because of its effectiveness in the regression-based task. Ada boost is suitable for classifiers, and XG boost is suitable for regression. The two methodologies are mostly employed in value prediction models for finding the value of specific capacitance. The first methodology involves implementation of regression model directly to the overall dataset to directly predict the specific capacitance value. The second method is a classification-based two-step process: the model identifies the class in which the specific capacitance value is situated. Once the class is determined, the regression model is applied in that class to predict the specific capacitance value. In the second method, the regression model is applied independently to each class rather than to entire dataset, allowing more accurate class based specific capacitance prediction process. These approaches allow for both category wise predictions and analytical predictions of specific capacitance value of supercapacitors. The first methodology is described in this study to predict the SC value.

In value prediction model, the SC of supercapacitor dataset split into two parts for training and testing the model for regression analysis. 80 % of the entire dataset is used for training the ML model and the remaining 20 % of the dataset is used for

testing and evaluating the performance of ML model using metrics results. Hyperparameters are also employed during training to reduce the challenges of overfitting, to ensure the better performance of the models. The cross validation (cv) technique is used during training the model, the cv value of 10 is used in the regression models for optimization using random search and grid search hyperparameter optimization techniques. The regression model performance is evaluated by using three performance metrics which are MAE, RMSE, and correlation coefficient ($R^2$). The description of performance metrics is explained in the model evaluation section. The result of five regression models is depicted in Figure 5.6 and metrics results are shown in Table 5.6. For showing the better performance of models, metrics should have higher value of $R^2$ and lower value of RMSE and MAE.

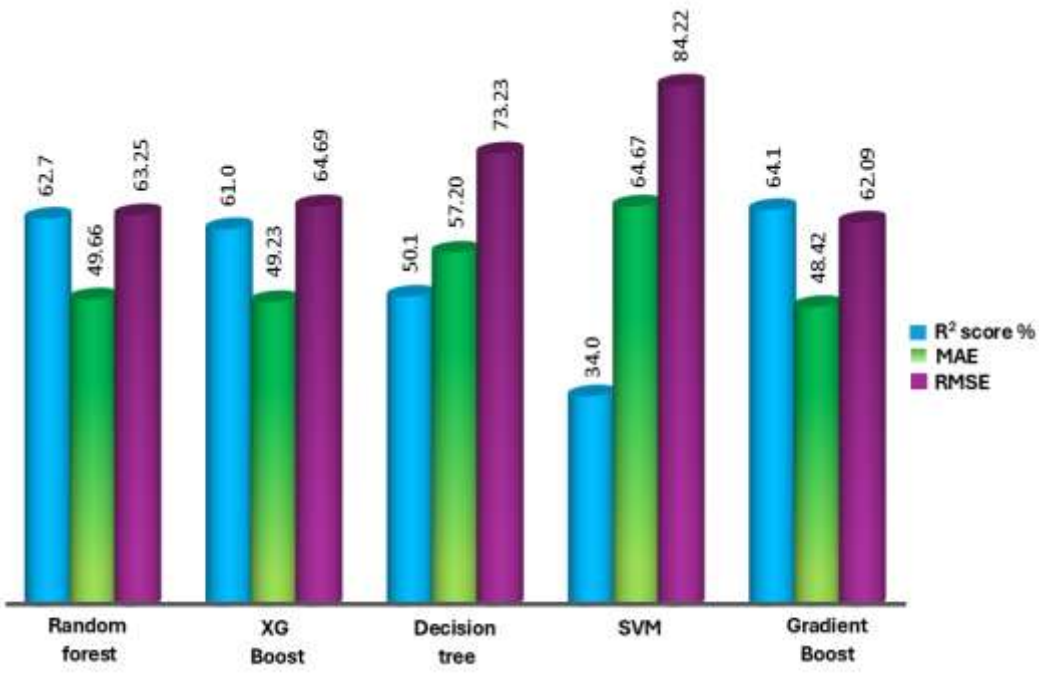


Figure 5.6 ML regression model result for specific capacitance value prediction with its performance metrics results. The best model performance

Random forest (RF) and Gradient boost (GB) models work best among various other algorithms by seeing its RMSE and correlation coefficient metrics in Table 5.6. RF models with random search hyperparameter tunning give $R^2$ score value of 0.627 and GB model with random search hyperparameter tuning give the best $R^2$ score value of 0.641 can be considered as moderate to strong correlation. GB outperforms other models of its low error metrics result such as low MAE and RMSE value of 48.423 and 62.09 as shown in Table 5.6. The dataset consists of activated carbon-based supercapacitor dataset that is the reason for getting moderate correlation coefficient value. GB models work best because they excel in nonlinear data because of its ensemble technique, which connects all Decision trees to process complex patterns of the considered dataset to reduce the problem of overfitting in the data. GB sequentially reduces errors, emphasizes important features and adapts relations better than RF, or linear models. The model has inherent feature importance ranking prioritization which helps in increasing specific capacitance prediction accuracy. Unlike SVM, it requires very little tunning of hyperparameter, also handle missing data better than other algorithms. All these properties make GB model work best for our activated carbon-based supercapacitor dataset for predicting its output specific capacitance value. XG boost have MAE and RMSE values of 49.233 and 64.697, which shows that it effectively captures the important features and patterns of our dataset of activated carbon-based supercapacitor. The GB model ability to model complex relationships outperforms other models which fail to solve nonlinearity in dataset. These two models, RF and GB, work best compared to other model performances described in Table 5.6. The results prove that the selection of appropriate features plays an important role in improving the performance of the models.

Table 5.6 The various ML regression model metrics results.

| S. No. | Models | $R^2$ score | MAE | RMSE |
|---|---|---|---|---|
| 1. | Random forest (random search) | 0.627 | 49.662 | 63.254 |
| 2. | XG boost (random search) | 0.610 | 49.233 | 64.697 |
| 3. | Decision tree (random search) | 0.501 | 57.204 | 73.234 |
| 4. | SVM (grid search) | 0.340 | 64.677 | 84.219 |
| 5. | Gradient boost (random search) | 0.641 | 48.423 | 62.097 |

Figure 5.7 illustrates the performance of GB and RF models on training and test set for predicted specific capacitance of supercapacitor. In Figure 5.7, the x axis denotes the actual value of specific capacitance and y-axis is the prediction value of specific capacitance of supercapacitor, where red data points refer to test set points and blue data points refer to train set points. Among the evaluated models, GB and RF demonstrated superior performance and therefore selected comprehensive analysis.

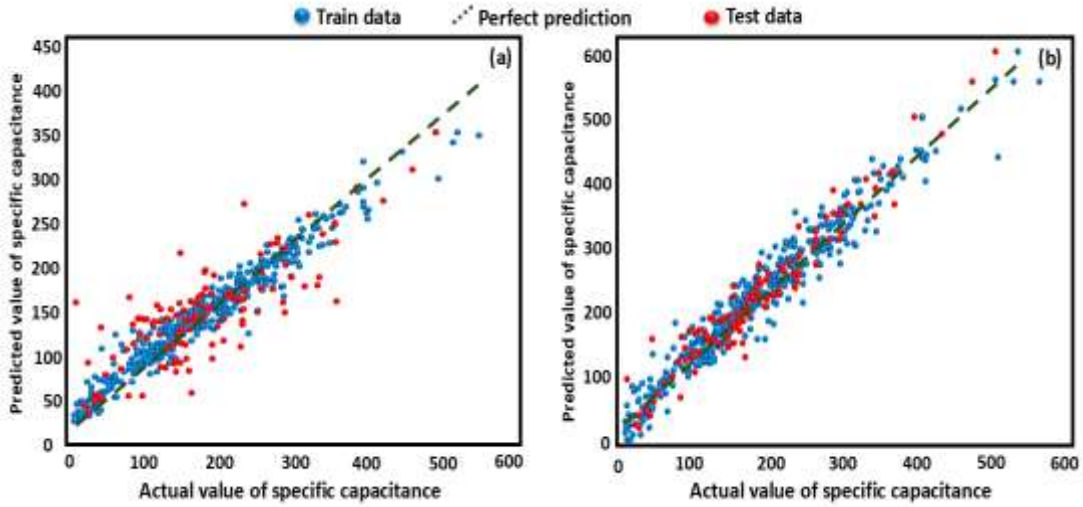


Figure 5.7 Actual and predicted specific capacitance of supercapacitor (a) Random forest test result of specific capacitance prediction. (b) XG boost test result of specific capacitance prediction.

As illustrated in Figure 5.7 and Figure 5.8, these models outperformed others based on their various performance measures like actual vs predicted specific capacitance results and metrics results. It is observed that the prediction for test points is better for GB compared to RF. Based on this comprehensive performance evaluation, the GB model is identified as the most reliable for value predictions. Figure 5.8 presents the comparative analysis of key performance metrics, including highest $R^2$ score %, along with lowest RMSE and MAE values of both models for better analysis. Based on performance assessment of both models, the GB model can also be suitable for identifying the specific capacitance value after class prediction in which class the specific capacitance lies. Following class prediction, the GB model is further employed for accurate value prediction to ensure best performance accuracy.

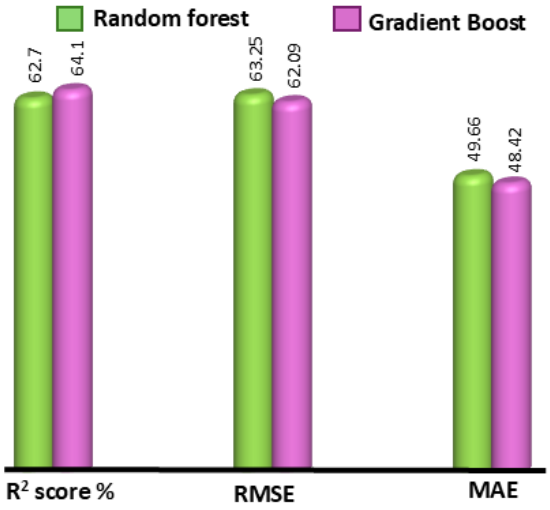


Figure 5.8 The performance result of two best models i.e. (i) the Random forest (RF) model and (ii) the Gradient boost model (GB).

The GB model depicts the best result by achieving RMSE of 62.09 as shown in Fig 8. To evaluate whether model captures actual values rather than only memorizing the trend, we established baseline model using a dummy regressor by using strategy

as mean. The mean specific capacitance is 188.13 F/g, with standard deviation of 104.9 F/g. The model predicting this mean value for tests set yield RMSE of 103.67. Optimized GB model achieved RMSE of 62.09 with 40.1 % reduction in prediction error compared to dummy baseline. Furthermore, $R^2$ calculated from these values is approximately 0.65, showing that the model estimates approximately 65% of variance in specific capacitance. The metrics in Table 7 prove that model successfully identified meaningful results rather than merely predicting average value of model.

**Table 7**: Baseline comparison of model for SC prediction

| S. No. | Model/Metrics | Value | Improvement |
|---|---|---|---|
| 1. | Mean of SC | 188.13 | |
| 2. | Standard deviation of test set | 103.66 | |
| 3. | Standard deviation of dataset | 104.9 | |
| 4. | Dummy regressor RMSE | 103.67 | |
| 5. | GBR model RMSE | 62.09 | 40.1% |

- **Features Prioritization**

It is mostly observed that multiple numbers of features increase the complexity of the models which result in poor performance. Hence, feature importance score is important to determine which features have a high impact on the ML model performance [193]. By analyzing the score, it helps to determine which features are important for perfect performance prediction. Along with performance, it improves interpretability and reliability of the ML model. Figure 5.9 shows the order in which features have importance in specific capacitance, prediction and its sensitivity. Table 5.7 described the importance of features which have significance on model performance. The most important features help in making decisions and help in

improving accuracy of the model. From the results of feature importance analysis, the SSA has high impact on GB model, with the importance score value of 0.322, after that potential have importance score of 0.190, and the least influence feature is Id/Ig in the GB model. In the RF Model, SSA and potential have high impact on model performance. The SSA has an important score value of 0.304, then potential with the value of 0.185 and the least influential feature is Id/Ig in RF model with the importance score value of 0.094.

SHAP analysis as shown in Figure 5.9(c) and 5.9(d) provides insights into how each feature contributes to the output of the model. It enhances model reliability and helps in explaining model behaviors on important features. In Figure 5.9, it is depicted that both RF and GB models work similarly but by seeing its feature importance score it is proved that GB model is more accurate and reliable.

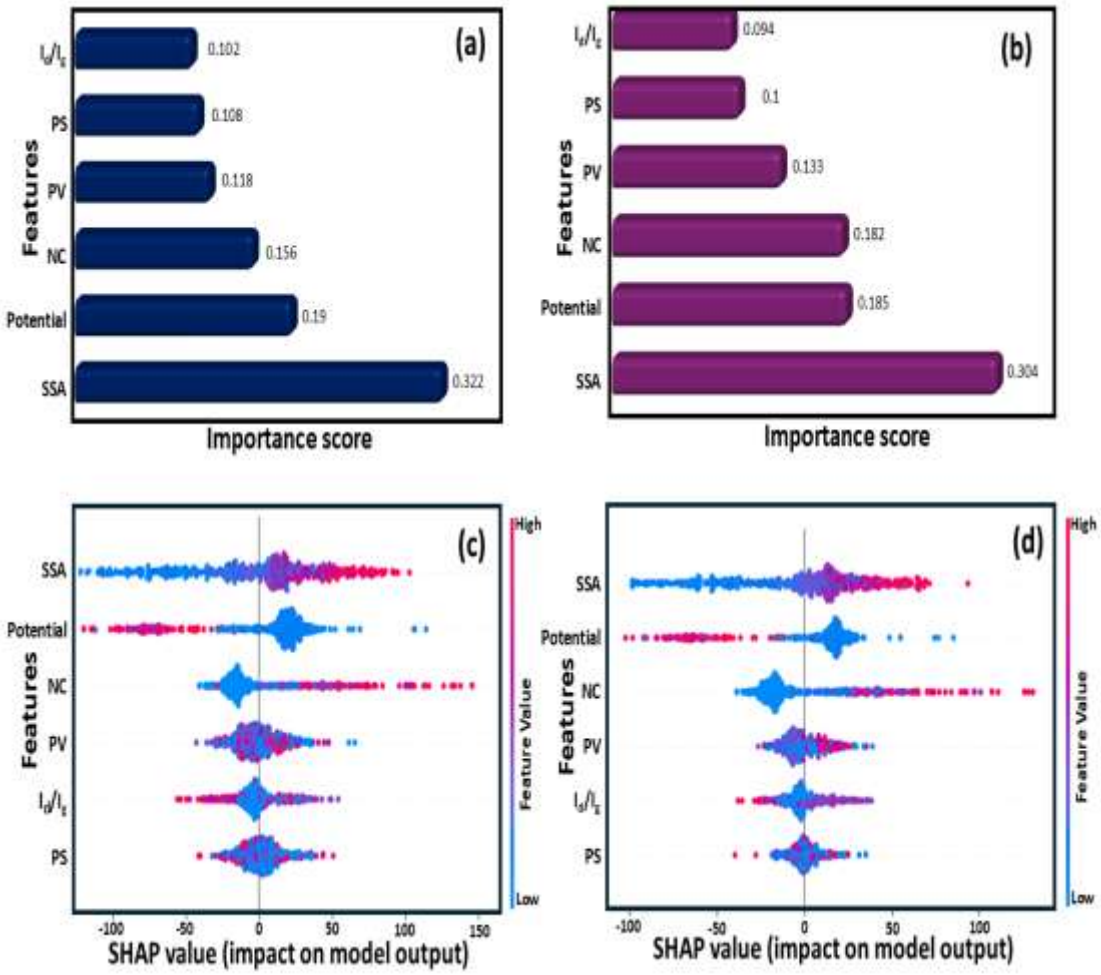


Figure 5.9 The impact of different features in prediction of specific capacitance using different ML models. (a) GB feature prioritization plot, (b) RF feature prioritization plot, (c) sensitivity analysis impact on GB model (SHAP analysis), (d) sensitivity analysis impact on RF model (SHAP analysis).

- **Correlation of the regression-based prediction results with electrochemical performance**

The relationship between physical properties and electrochemical performance was properly analyzed by correlations with its input features with predicted specific capacitance. A shown in Fig 9 there is a positive correlation between SSA and predictive SC. This observation is consistent with prior literature studies, which shows that that increased SSA provides more active sites for double layer formation through electrode-electrolyte interface. The increase in SSA enhances adsorption of ions and charge accumulation, thereby improving specific capacitance [194].

sHAP analysis reveals that SSA is one of the most important contributors to model predictions and analysis, with higher SSA values yield positive correlation contributions toward specific capacitance. However, the wide range of sHAP values at high SSA indicate non-uniform performance, depicting that other additional structural features like pore size distribution and how ions access the pore affect the utilization of surface area. This confirms that using only SSA is insufficient to describe the complexity of electrochemical behaviors [155].

The potential window acts as second most influential feature, revealing the importance of electrochemical operating conditions. By seeing the formula of energy density ($E \propto CV^2$), even the smallest change in potential window significantly affects the specific capacitance. This depicts that the electrochemical performance of AC based supercapacitor is not only dependent on material properties but also affected by testing conditions. NC introduced pseudocapacitive contribution and improve surface wettability of material. The nitrogen functionalities in carbon matrix improves the charge transfer characteristics. The sHAP results denote nonlinear contribution of NC, demonstrates the complex dependence of the distribution of type of nitrogen within the carbon matrix. The contribution of PS and PV effectively regulate the utilization of SSA, which explains why materials with similar SSA can have different specific capacitance. The low importance score of $I_d/I_g$ revealed that defects and disorder have low influence of specific capacitance in the considered dataset. Although defects can improve conductivity, their contribution has less impact within the dataset.

In result, the GB model capture the nonlinear relationship between physicochemical and electrochemical features. The GB model reveals that the specific capacitance is governed by complex relationship between SSA, pore architecture, and operating

conditions. The finding highlights the limitations and necessity of incorporating electrochemical features like electrolyte type and current density to further improve the model prediction accuracy and material screening. Overall, GB model demonstrates the potential window also acts as important features as other material structural features. For accurate specific capacitance prediction requires determining the combined nonlinear effect of both structural and electrochemical behavior properties.

Table 5.7 The feature pyritization score of two best ML regression models.

| S. No. | Features | Gradient boost | Random forest |
|---|---|---|---|
| 1. | Specific surface area (SSA) | 0.322 | 0.304 |
| 2. | Pore size (PS) | 0.108 | 0.100 |
| 3. | Pore volume (PV) | 0.118 | 0.133 |
| 4. | Potential | 0.190 | 0.185 |
| 5. | Nitrogen content (NC) | 0.156 | 0.182 |
| 6. | $I_d/I_g$ | 0.102 | 0.094 |

## 5.6 Conclusion and future outlook

This chapter presented a ML model based on classification and regression analysis to predict supercapacitor performance. The dataset is divided into 80:20 split which means 80% of data was used to train ML models whereas 20% of data is used to test the ML models. The class prediction model was used to determine the range of specific capacitance value lies. The five ML classification models were applied for class prediction in which Ada boost give the best result with RMSE value of 0.687. The best class prediction with precision, recall and accuracy is class 3 which shows that high specific capacitance prediction is more accurate. After applying

classification, the value prediction is done to find out the value of specific capacitance through ML regression models. The five major regression models were applied to predict the specific capacitance values. The GB and RF model have the best results among all other models with GB model RMSE value of 62.09 and MAE value of 48.43 which is very low compared to other regression models. The best regression models analyzed the input features using feature importance score indicated specific surface area, and potential were the most impactful features whereas pore size and Id/Ig were the least influential features of GB model.

The integration of ML into the development process of energy storage devices, particularly supercapacitors. This presents a data-driven approach to accelerate innovation and fabrication to optimize material synthesis and performance. This significantly reduces dependence on conventional trial and error methods, thereby minimizing experimental cost, time and effort. The limitation of the work is that the model $R^2$ values indicate that considerable amount of variance comes from factors and features which are not consistently reported in the analysis such as electrolyte, operating conditions, and scan rate. The future research will incorporate detailed physicochemical and electrochemical features that will significantly improve model accuracy and predictive performance. In further future research, ML models can be trained to predict capacity, power density, energy density and electrolyte behaviors of other electrode materials such as graphene, transitional metal-based energy storage devices or natural fiber derived electrodes materials for synthesis of energy storage. This approach not only improves the efficiency of development cycle of material but also promotes the use of renewable and sustainable resources. In further chapter, boosting ML models are used using hyperparameter tunning techniques to optimize the performance of supercapacitor electrodes. In conclusion,

the application of ML in the energy storage field offers powerful techniques for enhancing design, optimization and performance of supercapacitors. It supports data driven research methodologies that help faster development and cost-effective innovation in energy storage.

# Chapter 6

# Sustainable Data-Driven ML based energy storage devices performance analysis using its physicochemical and electrochemical characteristics

*[Activated carbons have emerged as most promising materials for energy storage devices, especially supercapacitors owing to their exceptional properties and excellent electrical conductivity. These carbon matrix structure present various advantages for applications in supercapacitor. The electrochemical characteristics of carbon electrodes are significantly influenced by its properties such as surface area, pore architecture, Raman spectroscopy and doping. Nevertheless, the absence of physical model that is capable of precisely predicting the performance of supercapacitor. This chapter introduces the data driven approach utilizing machine learning (ML) boosting models with hyperparameter tunning cross validation techniques to accurately estimating supercapacitor performance based on its microstructural and electrochemical characteristics of activated carbon material-based supercapacitor.]*

## 6.1 Introduction

Supercapacitors represent excellent advancements in the domain of energy storage, because it fulfills the gap between batteries and capacitors. Due to their high cycle life, increased power density and rapid response capabilities. Now a days, supercapacitors are widely used in smart power grids, energy storage systems, electrical infrastructures and aviation industries. The durability of supercapacitor is influenced by many factors, that include fabrication methods, choice of electrode materials, electrode design and their operating conditions [13]. These features play important roles in determining the supercapacitor's efficiency and longevity, highlighting the necessity for proper attention in both fabrication and designing process. Moreover, supercapacitor exhibits long cycle lifetime, enduring large number of charge-discharge cycles without degrading its functionality. This high durability ensures that it reduces replacement frequency overtime, emphasizing their economic ability and environment sustainability. Nonetheless, improving the efficiency of supercapacitor, particularly specific capacitance (SC) in F/g, remains a considerable challenge [170] [96].

Activated carbon (AC) has been highly investigated in nanomaterials, catalysis, and energy storage and conversion. AC is considered as excellent electrode material for supercapacitors due to high surface area, tunable porous structure, and excellent electrical conductivity, which facilitates energy storage. The specific capacitance of activated carbon is important factor in determining supercapacitor performance as it directly relates to material ability for charge storage. The progress observed in supercapacitor and energy storage fields is largely influenced by the innovative alteration of electrode materials and substrates through the introduction of nanomaterials, which have led to the significant increment in life cycle, energy

storage capacity, and power density. This paradigm has set new benchmarks for energy storage technologies. This chapter use boosting based ML models for evaluating the physicochemical and electrochemical characteristics of activated carbon electrode through the utilization of machine learning (ML) models as shown in Figure 6.1. This technique promotes the reliability of assessing supercapacitor performance, thereby facilitating the possibilities for adjusting electrode configurations and improving the effectiveness of energy storage devices. Despite these modifications, there is critical need to enhance the accuracy of predictions. The incorporation of AI and ML into supercapacitor study is imperative. The performance is evaluated by metrics such as correlation coefficient ($R^2$); root mean squared error (RMSE) and mean absolute error (MAE). These metrics provide quantitative framework for assessing and improving these ML models. The aim of this chapter is to leverage the recent development in ML to enhance performance of supercapacitor using the specific capacitance predictions. This represents the pivotal advancement in enhancement of energy storage technologies [195].

ML can be used at different stages in optimizing performance: predicting material properties, guiding synthesis route, modelling electrochemical responses and optimizing fabrication parameters. Recent applications combine regression and classification methods, and neural networks to either predict features performance or to propose processing methods [196][197][198][199]. In material discovery and activation process, ML was used to identify important precursors to predict sustainable oxygen-rich porous carbon, ML informed activation pathway produced carbon with tunable porosity and significantly improved specific capacitance in acidic environments thereby validating data driven selection [196]. In features performance prediction, regression and ANN models quantifying how pore size and

surface chemistry correlate with capacitance and energy density, facilitating quick screening without the need of repeating experiments [197]. In device and process optimization, evolutionary and swarm intelligence algorithms, together with ML based parameters are employed to optimize storage efficiency and tune fabrication parameters to achieve desired electrical characteristics [198].

### 6.1.1 Literature review

The application of machine learning in materials science and energy storage has gained interest rapidly due to its capacity to process complex high-dimensional datasets, non-linear correlations between material properties and performance metrics results. Within the energy storage domain, ML facilitates the optimization of materials by enabling the prediction of electrochemical behaviour of electrode materials based on its input parameters such as morphology, cyclic voltammetry, compositions and fabrication conditions. These ML based prediction models allow researchers to reduce experimental iteration, reduce resources, and achieve accurate predictions of key performance parameters including specific capacitance, energy density, power density and charge-discharge efficiency [200]. As illustrated in Fig. 6.1., ML models employed for energy storage device performance prediction are mainly used supervised learning approaches. Supervised learning is ML algorithm that depends on a mapping function from labelled datasets, where both input features and corresponding outcome features are known. The primary objective is to reduce prediction error to unknown data using statistical and optimization techniques. The supervised learning-based ML models are Decision tree, Random forest, Gradient boost, Ada boost, and XG boost [132], [201]. These models effectively predict electrochemical properties that include current, energy density and specific capacitance. Among these models, tree-based models are preferred for

their capability to manage complex datasets while providing insights into feature importance, thereby aiding in the determination of important key factors affecting the performance of the device.

Previous research demonstrated ML can act as powerful tool in advancing supercapacitor technology. For example, Saad et al. [126] used multiple ML regression models to predict the performance of graphene-based electrode for supercapacitor. Their analysis depicted that ANN model works best and achieved high accuracy by showing lowest error in the actual and predicted value (RMSE = 60.42 (F/g) effectively spotting complex, nonlinear data patterns This highlights ANN model's potential for accelerating the discovery of high-performance materials.

The author demonstrates a hybrid deep learning model (CDTCN-GRU) for predicting remaining useful life for supercapacitors. The model demonstrates high accuracy with low RMSE and MAE value of 1.31 and 0.94 in charging phase. In discharging phase, it obtained RMSE value of 2.35 and MAE value of 1.46. It effectively explains complex degradation patterns and works best under different conditions. It uses explainable AI (XAI) to explain its interpretability, enhancing the transparency. This makes models powerful and transparent tools for effectively predicting the lifespan of supercapacitors [202].

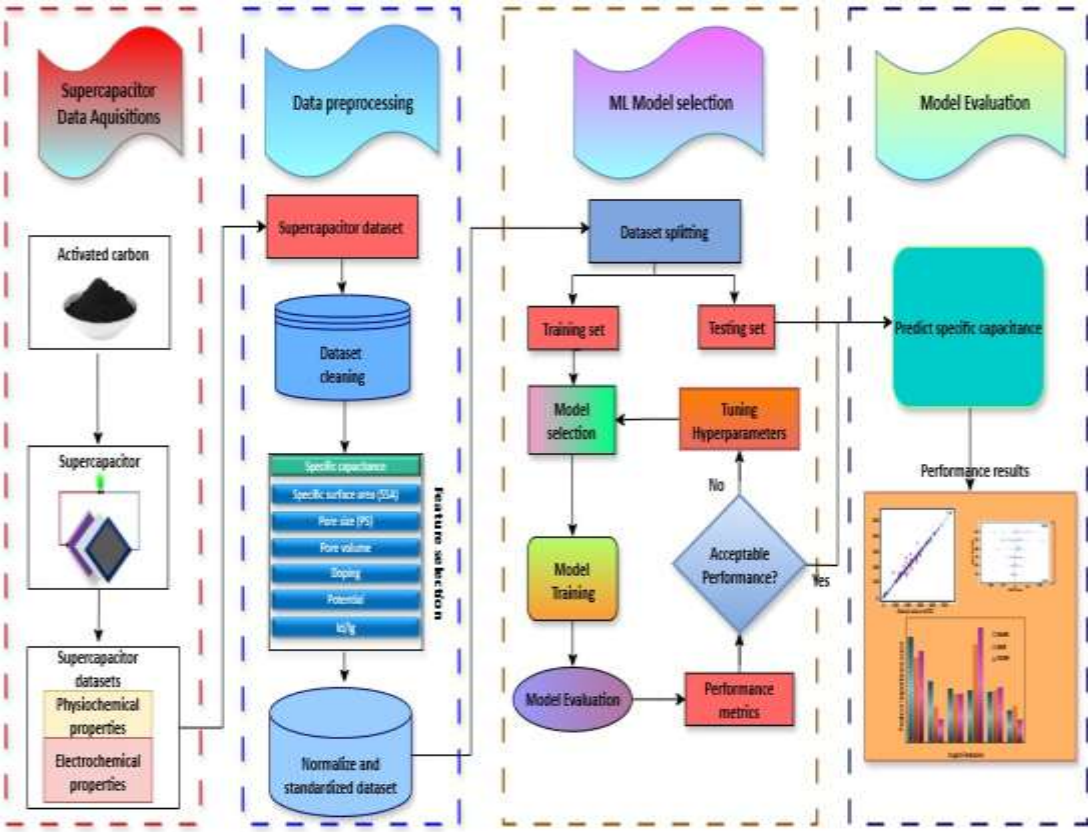


Figure 6.1 Proposed Machine learning regression model for specific capacitance prediction for AC electrode-based supercapacitor

M. Tang et al. [203] in their research study employed Random forest (RF), XG boost (XGB), and regression tree (RT) for regression analysis and feature selections of $MgCo_2O_4$ electrode for capacitance prediction. Author constructed nine ML models sourced from published data, input features of the electrode included structural, and electrochemical parameters. The research introduces recursive feature elimination (RFE) methods with regression models. It repeatedly builds models and then removes the least important feature. This process continues until the best and most important set of features is left. The XGB-RFE-XGB model demonstrates best performance with $R^2$ of 0.95 and low error with RMSE of 111.86 F/g, establishing as reliable technique for optimizing experimental design and improving material development.

## 6.2 Proposed work

This chapter has incorporated a hyperparameter-based grid-search optimization technique in supervised regression ML models with the framework for specific capacitance prediction based on physicochemical and electrochemical features of activated carbon electrode to enhance its overall accuracy. This specific optimization base technique reduces the risk of overfitting, as accurate predictions and productively balance the ML model algorithms. Integrating ML regression model in specific capacitance prediction for energy storage system improve reliability, transparency, and usability before moving towards the designing process. These models provide decision making procedures, allowing fabricating designers to understand how electrode material properties impact energy storage device overall performance.

## 6.3 Methods and materials

This chapter entailed the formulation of ML models utilizing data procured from surveys of previously published research papers [185]. The input features employed in the chapter consist of indicators pertaining to both structural attributes of the material and the properties derived from physical and electrochemical testing. The structural attributes are represented by elements such as pore size distribution, surface area, Id/Ig ratio (Raman spectroscopy), and the heteroatom doping. Conversely, the electrochemical testing properties include electrolyte type, concentration, conductivity, current density, and potential window. The efficiency of activated carbon-based electrode material is assessed by specific capacitance. Only very limited research articles provide numerical data on all desired features, despite knowing that the large information was obtained from the previous research papers examined. For developing the dataset, total data of 558×7 rows and columns

were employed, which were subsequently partitioned into 90% training subset and 10% testing subset to appraise the performance of three supervised ML regression models. By introducing the previously unexplored features of AC electrode material, our objective is to provide comprehensive modelling strategy that summarizes the inherent complexity found in experimental research on supercapacitors[172] [132].

In the present study, the key parameters that directly and indirectly influence the specific capacitance and device performance are discussed prior to model development and algorithm design. The summary of all input features used for analysis are specific surface area (SSA) measured in $m^2/g$, pore size distribution (PS) measured in nm, pore volume (PV) in $cm^3/g$, potential window (PW) in V, nitrogen doping content (NC) measured in atm %, Raman spectroscopy which is D and G band ratio ($I_d/I_g$) form, and specific capacitance (SC) in F/g.

SSA of AC plays important role in deciding the electrochemical performance of supercapacitor, as it directly affects the electrode-electrolyte interface available for charge accumulation. Although AC exhibits high SSA values ranging between 1000-3000 $m^2/g$ but affect the electrochemically accessible surface area of electrode to contribute to specific capacitance. The micropores less than 2 nm serve as primary charge accumulation sites, while mesopores facilitate ion transport, reducing diffusion resistance and improving rate capability. The macropores acts as ion-buffering reservoirs. Therefore, achieving high specific capacitance depends on well-designed hierarchical pore structure than only enhancing the surface area. Recent literature has therefore emphasized more on the factor that specific capacitance correlates with electrochemically accessible surface area and pore size distribution than only on total SSA [204].

The operation potential window is important factor that governs the electrochemical performance of AC based supercapacitor as it directly affects the energy output and charge storage behaviour of device by accumulating more electrolyte ion on electrode surface, leading to higher charge storage and specific capacitance. At very high potential when limit exceed, electrolyte decomposition and gas formation can occur, which reduce efficiency and compromise electrode integrity. Therefore, optimization of operating potential window is important to enhance specific capacitance while maintaining electrochemical stability and life cycle performance. Nitrogen doping improves the specific capacitance by incorporation of nitrogen functional groups which introduces pseudocapacitive redox sites, that directly influence the electrical conductivity, and surface wettability of the device. However, the increase of specific capacitance depends more on the type of nitrogen doping and distribution of nitrogen content than total nitrogen content, as excessive doping also affect carbon structure and reduces conductivity, which negatively impacts specific capacitance and affect performance [205].

The intensity ratio of D and G band is key factor for evaluating structural disorder in carbon-based electrode materials. The high ratio of Id// Ig ratio indicates higher degree of defects within the carbon framework. Such structural imperfection results in electrical conductivity of electrode material that will affect the specific capacitance and performance of the device. Consequently, higher the ratio, higher the defects can lead to reduction in overall capacitive performance of the energy storage device [139].

After dataset formation, missing values, and outliers are handled properly [206]. Continuous features were properly normalized and standardized to reduce disparities across different dimensions. This process forms the final dataset of 221

records of AC based electrodes. Each data entry contains valid numerical values for all considered features.

The output feature is predicted using distinct ML regression models which are Gradient boost regression model (GBR), Extreme Gradient boost regression model (XGBR), Ada boost regression model (ADBR)[207]. GBR is boosting technique that builds groups of weak models, mainly Decision tree, in sequential stage manner. It improves the performance of models by reducing errors of previous models, minimizing their loss function using gradient descent. This model obtained accuracy and reduced overfitting complexity, suitable for real world non-linear complex data. XGBR is advanced version of Gradient boosting through parallel processing, Decision tree pruning, and loss function regularization to reduce overfitting problems. It has been built to automatically handle missing values during training. ADBR operates by sequentially fitting weak learners to training data after one another. The final prediction is weighted median of all previous weak learners improving accuracy by concentrating on difficulty predicting data samples by minimizing bias of the model.

For ML model development, using cross validation (cv) of 5 for grid-search hyperparameter tuning for optimizing the ML models. Subsequently, the models are trained using optimum features and evaluated through performance metrics that include correlation coefficient ($R^2$) and root mean square error (RMSE), thereby improving the robustness of model performance across different conditions. $R^2$ acts as important metric parameter for evaluating the fit of ML models by computing the variance by independent features relative to dependent features. $R^2$ values lie between 0 to 1. The value close to 1 depicts strong fit of model. RMSE used to measure prediction errors between actual and predicted output feature value. It is

used to amplify large errors, ideal for outliers and high deviations. RMSE helps to find bias by depicting non-uniform model performance. The value close to 0 depicts strong model with no error [208].

The mathematical representation of performance metrics are described below where *n* is total number of data points, $C_i$ is actual value of output, which is specific capacitance, $\hat{C}_i$ is the predicted value of output, $\bar{C}_i$ is mean of actual value of output, and $\bar{\hat{C}}_i$.

Coefficient of determination: $R^2 = 1 - \frac{\sum_{i=1}^{n}(C_i - \hat{C}_i)^2}{\sum_{i=1}^{n}(C_i - \bar{C}_i)^2}$ Equation 6.1

Mean absolute error: $MAE = \frac{1}{n}\sum_{i=1}^{n}\left|C_i - \hat{C}_i\right|$ Equation 6.2

Root means square error: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(C_i - \hat{C}_i)^2}$ Equation 6.3

## 6.4 Mathematical modelling of ML models

The ML regression models algorithms are explained in detail in this section as follows. The three algorithms are GBR, XGBR, and ADBR.

### 6.4.1 Gradient boost algorithm

Training dataset $P = \{(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$, where $x_i$ is input feature vector and $y_i \in R$ is output feature vector. Loss function: $L(y, F(x))$, MSE : $L\big(y, F(x)\big) = \frac{1}{2}(y - F(x))^2$. Number of iterations for trees: M. Learning rate of model: $\eta$. Weak learner: Tree $h(x)$ with $T$ terminal nodes. Output value vector $\gamma$ for each leaf. The algorithm steps are below:

#1: Initialize the model with constant prediction value $\gamma$ that minimizes the loss.

$$F_0(x) = argmin_\gamma \sum_{i=1}^{n} L(y_i, \gamma)$$ Equation 6.4

For Mean square loss (MSE), the aim is to obtain the mean of output feature,

$$F_0(x) = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad \text{Equation 6.5}$$

#2: Build each tree sequentially from $m = 1\ to\ M$. Compute its negative gradients (pseudo residuals)

$$r_i^m = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}, \text{ for } i = 1,2, \ldots . n \qquad \text{Equation 6.6}$$

For the square error loss (MSE):

$$r_i^m = -\left[\frac{\partial}{\partial F}\left(\frac{1}{2}\left(y_i - F(x)\right)^2\right)\right]_{F(x)=F_{m-1}(x)} = y_i - F_{m-1}(x_i) \qquad \text{Equation 6.7}$$

Fit weak learner to the pseudo residuals; train the tree $h_m(x_i)$ to current data $(x_i, r_i^m) for\ i = 1,2, \ldots n$. This tree partitions the inputs into *T* leaves $(R_1^m, R_2^m, \ldots . . R_T^m)$. For each leaf $t = 1,2, \ldots T$ in tree $h_m$. Find out the output value for $\gamma_t^m$ for each leaf region $R_t^m$ that minimizes the loss for all samples of that leaf.

$$\gamma_t^m = argmin_\gamma \sum_{x_i \epsilon R_t^m} L(y_i,\ F_{m-1}(x) + \gamma. \qquad \text{Equation 6.8}$$

$$\text{For MSE loss } \gamma_t^m = mean_{x_i \epsilon R_t^m} r_i^m. \qquad \text{Equation 6.9}$$

The new model $F_m(x)$ is combination of previous model plus the new tree

$$F_m(x) = F_{m-1}(x) + \eta h_m(x). \qquad \text{Equation 6.10}$$

#3: After *M* iterations, the final model is

$$F(x) = F_0(x) + \eta \sum_{m=1}^{M} h_m(x) \qquad \text{Equation 6.11}$$

### 6.4.2 XG boost algorithm

This ML regression algorithm has an additive $F(x)$ that minimizes the regularized objective function.

Training data: D = $\{(x_1, y_1), (x_2, y_2), \ldots., (x_n, y_n)\}$ where $x_i$ is input feature vector and $y_i \in R$ is output feature vector. The loss function is $L(y, F(x))$. Boosting iterations: M. Learning rate of model is denoted by $\eta$. The regularization parameter is $L_2$ regularization on leaf weights denoted by $\lambda$. The weak learner regression tree $f_m(x) = w_{q(x)}, q(x)$ maps sample $x$ to one of $T$ leaves and $w$ is weights.

The objective function includes both loss and regularization term to penalize model complexity.

$$L^m = \sum_{i=1}^{n} L\big(y_i, F_{m-1}(x_i) + f_m(x_i)\big) + \Omega(f_m) \qquad \text{Equation 6.12}$$

where regularization term $\Omega(f_m)$ for tree $f_m$ is defined as

$$\Omega(f_m) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \qquad \text{Equation 6.13}$$

where $T$ is number of leaves in tree, $\gamma$ is cost factor for adding new leaf.

#1: Initializes the model using function

$$F_0(x) = \arg_w \min \sum_{i=1}^{n} L(y_i, w) + \frac{1}{2}\lambda w^2 \qquad \text{Equation 6.14}$$

for MSE loss

$$F_0(x) = w_0 = \frac{\sum_{i=1}^{n} y_i}{n+\lambda} \qquad \text{Equation 6.15}$$

#2: Compute pseudo residuals and hessian functions for each instance $i = 1,2, \ldots., n$, calculate $g_{im} = \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\Big|_{F(x)=F_{m-1}(x)}$, $h_{im} = \frac{\partial^2 L(y_i, F(x_i))}{\partial F(x_i)^2}\Big|_{F(x)=F_{m-1}(x)}$.The goal is to find tree structure $q(x)$ and weight $w$ that minimize *II*$^{nd}$ order approximation of objective function.

$$L^{(m)} = \sum_{i=1}^{n}[\, g_{im} f_m(x_i) + \frac{1}{2} h_{im} f_m(x_i)^2] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \,. \qquad \text{Equation 6.16}$$

For fixed tree structure $q(x)$, find optimal weight $w_j^*$ for leaf $j$ by taking derivation of objective and fix it to zero. All samples in $x_i$ in leaf $j$ denoted by

$$I_j = \{i|\, q(x_i) = j\},\ w_j^* = -\frac{\sum_{i\in I_j} g_{im}}{\sum_{i\in I_j} h_{im}+\lambda} \qquad \text{Equation 6.17}$$

Calculate optimal objective reduction for split which is calculated by gain formula. The gain is reduction in the objective achieved by making split compared to not splitting like having single leaf.

$$Gain = \frac{1}{2}\left[\frac{(\sum_{i\in I_L} g_i)^2}{\sum_{i\in I_L} h_i+\lambda} + \frac{(\sum_{i\in I_R} g_i)^2}{\sum_{i\in I_R} h_i+\lambda} - \frac{(\sum_{i\in I} g_i)^2}{\sum_{i\in I} h_i+\lambda}\right] - \gamma \qquad \text{Equation 6.18}$$

where $I_L$ and $I_R$ are instances of left and right children after split. $I = I_L \cup I_R$ split is only possible if gain is positive. After tree structure $q(x)$ is made and leaf weight is calculated. Update the model by using function

$F_m(x) = F_{m-1}(x) + \eta . f_m(x)$, for sample $x$ that ends up in leaf $j$:

$$F_m(x) = F_{m-1}(x) + \eta . w_j^*. \qquad \text{Equation 6.19}$$

#3: The output of final model

$$F(x) = F_M(x) = F_0(x) + \eta \sum_{m=1}^{M} F_m(x) \qquad \text{Equation 6.20}$$

### 6.4.3 Ada boost algorithm

Ada boost works by sequentially fit model to re-weight version of data, its focus on instances where previous models did not perform well.

Training data: D = $\{(x_1, y_1), (x_2, y_2), \ldots., (x_n, y_n)\}$ where $x_i$ is input feature vector and $y_i \in R$ is output feature vector. The Ada boost algorithm weak learner can be

trained on weighted data of shallow Decision tree called decision stump. Boosting iteration are $M$. Loss function of algorithm is $L(y, F(x))$. The weight is $w_i$

#1: Initializes the weights uniformly for first model. $w_1(i) = \frac{1}{n}\ for\ i = 1,2,3, \ldots . n$. initialize the prediction to 0 $F_0(x_i) = 0\ for\ i = 1,2, \ldots n$.

#2: For $m = 1\ to\ M$, fit weak learners $f_m(x)$using training dataset with weighted data $w_m(i)$.

$$f_m(x) = \ arg_f min \sum_{i=1}^{n} w_m(i).\ L(y_i, f(x_i)) \qquad \text{Equation 6.21}$$

Calculate loss for each data point based on new weak learner's prediction. The loss normalized between 0 and 1. $L_m(i) = \frac{|y_i - f_m(x_i)|}{D_m}$, $L_m(i) \epsilon [0,1]$ where $D_m = max_{i\epsilon\{1,2,\ldots,n\}} |y_i - f_m(x_i)|$. Calculate total weighted average loss for weak learner $f_m$. This measures overall performance of this learner

$$\bar{L}_m = \sum_{i=1}^{n} w_m(i).L_m(i) \qquad \text{Equation 6.22}$$

Calculate weight $p_m$ for weak learner $f_m$. A low average loss gives higher confidence $p_m$, $p_m = \frac{L_m}{1-L_m}$ ,if $\bar{L}_m \geq 1\ or\ \bar{L}_m \leq 0$ algorithm should be stopped if condition is fulfilled because of weak learner is perfect. Add new weak learner prediction, scaled by confidence $p_m$ to the ensemble function

$$F_m(x) = F_{m-1}(x) + p_m.f_m(x) \qquad \text{Equation 6.23}$$

Update the weight of instances that were predicted poorly by $f_m$ so next learner mainly focus on them $w_{m+1}(i) = \frac{w_m(i)}{Z_m}.{p_m}^{1-L_m(i)}$, where $Z_m$ is normalization factor.

#3: The final model $F(x)$ is median of all weak learners. For $x$ each learner $m$ gives prediction $f_m(x)$ with weight $w_m$.

$$F(x) = F_M(x) = \sum_{m=1}^{M} p_m . f_m(x) \quad \text{Equation 6.24}$$

## 6.5 Results and discussion

In result, we will discuss about the statistical analysis, correlation matrix in which the correlation between features is explained, how they are correlated. After that, different ML models performance metrics results will be discussed, which model is best suited for analyzing the performance of supercapacitor. Afterward, training and testing result of actual and predicted SC is shown. Finally, feature prioritization graph with sHAP analysis of best model will be discussed to show which features are important for model result predictions [133], [209].

### 6.5.1 Statistical analysis

The dataset has 558 samples datapoint with 7 key features which are already described above. Table 6.1 highlights the important statistical tendency and variability of each key feature. The mean of SSA is 1166 $m^2/g$ with standard deviation of 883.2 $m^2/g$ whereas PS and NC show wide ranges and relatively high variance, showing the presence of outliers. The SC display high variability by mean of 188.13 F/g, and maximum of 587 F/g, reflecting varied electrode performance. For ML model, it highlights the need to introduce normalization to balance feature scales, and outlier handling to improve model performance.

Table 6.1 AC based electrode features statistical analysis

| | SSA | PS | PV | PW | NC | $I_d/I_g$ | SC |
|---|---|---|---|---|---|---|---|
| Count | 558 | 558 | 558 | 558 | 558 | 558 | 558 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Mean | 1166.8 | 1.714 | 0.867 | 1.296 | 1.329 | 0.555 | 188.130 |
| Std | 883.20 | 2.616 | 0.861 | 0.792 | 2.875 | 0.798 | 104.47 |
| 50% | 1016.7 | 0.900 | 0.690 | 1.00 | 0.00 | 0.00 | 176.50 |
| max | 4073 | 27.20 | 5.91 | 4.0 | 19.8 | 4.54 | 587 |

### 6.5.2 Pearson correlation coefficient analysis (PCC) analysis

Figure 6.2 depicts PCC heatmap between all features, where color and numerical points indicate strongness and direction of linear relationship between features. The positive numeric signify which feature increase together, whereas negative numeric depicts the inverse relationship between each other. The model demonstrates on these patterns such as positive correlation with SC for SSA, PV, and NC have correlation value of 0.35, 0.11 and 0.21 which depict the SC increase when these features increases whereas negative correlation link for PS and PW. The correlations help ML model to accurately predict SC which will enhance its prediction accuracy [191].

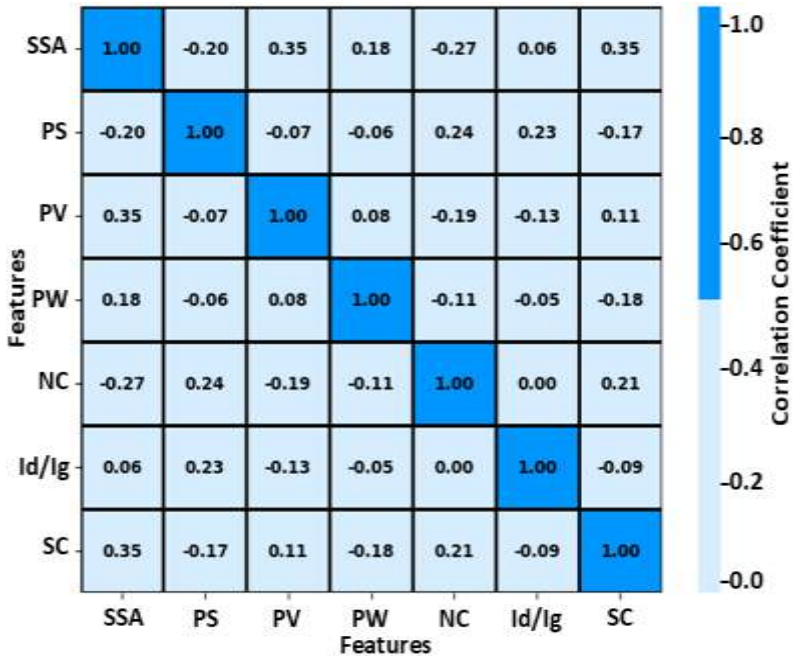

Figure 6.2 Pearson correlation coefficients matrix of all features among each other.

### 6.5.3 ML model performance

The ML models used for this study are already discussed in above section. After data formation and preprocessing, the dataset is ready for training the ML models in which the dataset is divided into 80%: 20% training and testing subset for performance analysis.

Table 6.2 ML model performance indicator

| S. No. | Models | RMSE | MAE | $R^2$ score |
|---|---|---|---|---|
| 1. | XG boost (XGBR) | 36.60 | 28.17 | 0.893 |
| 2. | Gradient boost (GBR) | 36.53 | 29.22 | 0.893 |
| 3. | Ada boost (ADBR) | 56.29 | 48.99 | 0.747 |

Table 6.2 describes the performance metrics values and Figure 6.3 depicts the performance metrics plot of all used ML models. Among all the models evaluated, XGBR yields RMSE value of 36.60, MAE of 28.17 and $R^2$ score of 0.869 which depicts the best performance results to predict specific capacitance. While seeing GBR and ADBR, MAE value of 29.22 and 48.99 which is more than XGBR. The error should be low for prediction the output SC. This model used grid search ensemble learning to train the model. The XGBR achieved best performance due to incorporation of tree pruning, advanced optimization techniques. XGBR is refined implementation of GBR that builds on its foundation by introducing L1 and L2 regularization in its objective function.

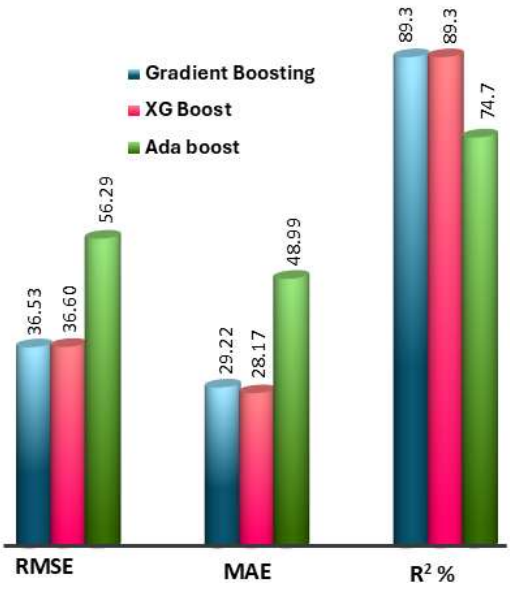


Figure 6.3 ML model performance using performance metrics

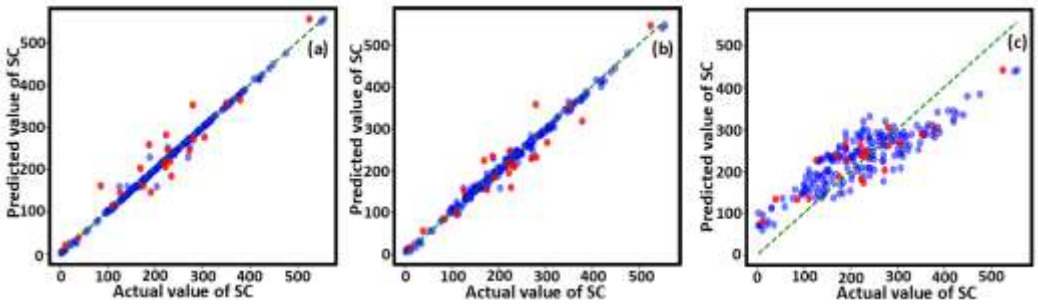


Figure 6.4 Real value and estimated value of SC with (a) XGBR result, (b) GBR results (c) ADBR results

Figure 6.4 depicts the plot of actual value of SC to the predicted value of SC from multiple models used in the study. Model accuracy is determined by the proximity to the best fit line.

The blue data points are train set, and red data points are test set. The test data points in XGBR model lie very close to the best fit line, while for GBR model, test data points are also close to the best fit while for ADBR model it is not very close as compared to other two models. For prediction of SC, XGBR give the best prediction result for new real time test set by analyzing the Figure 6.4. XGBR performed best

by seeing its best fit line which illustrates that the training and test data points lie closer to the fit line.

Table 6.3 ML model comparison table for demonstrating best model.

| Feature | ADBR | GBR | XGBR | Why XGBR |
|---|---|---|---|---|
| Goal | Reduce previous errors by assigning weight again to data | Minimize residual error | Minimize regularization objective function | Robust objective |
| Regularization | No, sensitive to overfitting | Limited | Yes, with introduction of L1 and L2 regularization | Reduce overfitting |
| Loss optimization method | No | $I^{st}$ order gradient function | Use $II^{nd}$ order (gradient and Hessian) | Precise overfitting |
| Outliers | Sensitive to outliers | Prone to overfitting if not tuned properly | Stable due to regularization and shrinkage | Robust, high resilience |
| Speed | Slow | Slow | Optimized and parallel processing | Fast training |

Table 6.3 describes the comparison of various machine learning models used in the study for predicting the performance of supercapacitor with the focus on XGBR model for give the best performance metrics result among other evaluated models. The table compares the key features of ML regression models such as regularization techniques to reduce overfitting, optimization methods to reduce loss, outliers and algorithm speed.

## 6.6 ML model features importance

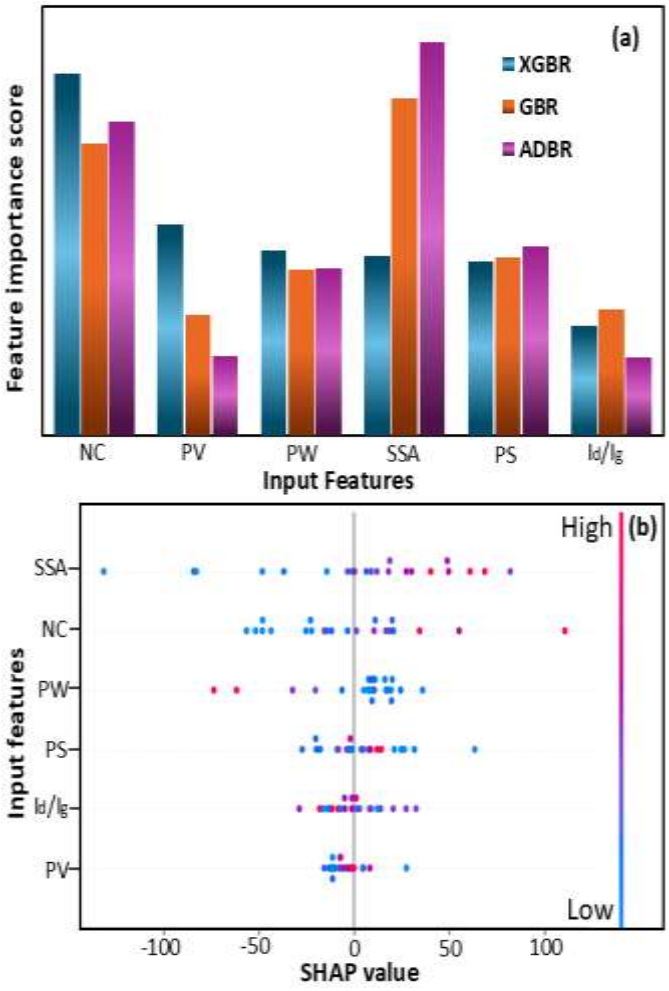


Figure 6.5 Feature analysis of ML models (a) Feature importance (FI) score of all ML models, (b) SHAP analysis of best XGBR model

The feature importance score is important parameter to describe the role of its key feature in model evaluation. The score of its key feature is shown in Table 6.4. Figure 6.5a depicts feature importance plot using bar chart revealing that NC and PV are most critical features for XGBR prediction with the value of 0.295 and 0.172 while PS value of 0.142 and Id/Ig value of 0.09 have least impact on the XGBR performance. Figure 6.5b depicts the sHAP analysis of all important features for XGBR to determine the feature influence on the performance of best XGBR model.

Table 6.4 ML models with its feature importance score

| | XGBR | GBR | ADBR |
|---|---|---|---|
| NC | 0.295 | 0.238 | 0.256 |
| PV | 0.172 | 0.099 | 0.065 |
| PW | 0.151 | 0.136 | 0.137 |
| SSA | 0.147 | 0.275 | 0.321 |
| PS | 0.142 | 0.146 | 0.154 |
| Id/Ig | 0.09 | 0.103 | 0.064 |

## 6.7 Conclusion and future scope

This chapter evaluated various ML model for predicting SC of AC based electrode for supercapacitor using experiment extracted data from previously published research literature. The input features included physicochemical and electrochemical characteristics that include PS, PV, NC, $I_d/I_g$, PW, SSA, and SC. The XGBR model outperforms other models, achieving near close to perfect prediction of SC by getting RMSE of 36.60 and $R^2$ of 0.893. The features of importance plot and sHAP analysis show NC and PV are the most impactful features governing accuracy of the ML model. These results underline the importance of features in optimizing the performance of models for energy storage technologies. This ML model helps in guiding the designing and fabrication process of supercapacitor by determining its specific capacitance. The predictive ML models hold substantial potential for material science community by providing the rational design of high-performance energy storage material. Furthermore, these methods can be extended to large-scale material discovery and optimization of material contributing to next generation energy storage devices for applications in

electrical vehicles, electronics devices and grid energy storage systems. The next chapter uses biomass-based precursor dataset for evaluating the performance of carbon electrode-based supercapacitor

# Chapter 7

# Biomass derived carbon electrode Supercapacitor performance prediction for energy storage devices.

*[Biomass derived carbons from natural sources are emerging as sustainable options provide waste management solutions and cost-effective materials for supercapacitor electrodes. Traditionally, biomass materials are developed for supercapacitors through trial-and-error process, linking electrochemical performance with specific descriptors. This chapter focusses on prediction of electrochemical performance of biomass derived from carbon electrode material for supercapacitors using four machine learning based approaches: Random forest, XG boost, Cat boost and Gradient boost were evaluated with an 80:20 training to test ratio. All models demonstrated strong correlation between experimental results and predicted outcomes, validating their accuracy when compared with their existing state of art models. The results depict that Cat boost model outperforms other ML models, achieving near perfect correlation coefficient ($R^2$ score) of 94.6 % with low root mean square error (RMSE) of 22.06. These findings provide a faster, accurate alternative to traditional experimental methods for supercapacitor optimization and help guide the fabrication and experimentation process]*

## 7.1 Introduction

Biomass is rapidly emerging as a central focus in the pursuit of sustainable energy storage solutions, especially as advanced materials for supercapacitors. Historically, biomass such as agriculture waste, hemp, and bamboo has been valued for their abundance, renewability, and non-toxic nature, making them essential in industry and construction. However, their unique structure and chemical versatility are now being harnessed to create high-performance activated carbon and programmable carbon, both of which are crucial for next-generation supercapacitor electrodes [210]. Despite their promise, biomass have often been overlooked in favour of synthetic alternatives, but recent research demonstrates that their carbonization and activation can yield electrode materials with high surface area, tunable porosity, and excellent capacitance, key factors for efficient energy storage. Supercapacitors stand out in the energy storage landscape due to their ability to charge rapidly and discharge slowly, making them highly sought after for applications ranging from electric vehicles to grid stabilization and portable electronics [136]. Unlike batteries, which rely on chemical reactions, supercapacitors store energy through the physical adsorption of ions at the electrode–electrolyte interface, enabling high power densities and exceptional cycle lifespans [211]. The performance of a supercapacitor is fundamentally determined by the properties of its electrode materials: high surface area, optimal pore structure, and excellent electrical conductivity are essential for maximizing specific capacitance (SC) and energy storage efficiency. Here, biomass-based precursors for developing carbon material for energy storage present a unique and largely untapped opportunity.

The process of converting biomass into activated carbon is both an art and a science. It typically involves carbonization, heating the precursor in an inert atmosphere to form a carbon-rich matrix, followed by activation using physical (steam, $CO_2$) or chemical agents (KOH, $ZnCl_2$). This process creates a network of micro- and mesopores, dramatically increasing the surface area and providing abundant sites for adsorption and charge storage [212]. For example, cotton shell-derived activated carbon has demonstrated SC as high as 247.82 F/g, outperforming many commercial carbons, and highlighting the untapped potential of agricultural byproducts as energy materials [213]. Similarly, cellulose-based supercapacitors have shown superior electric storage capacity and high-power density, making them promising candidates for flexible and renewable paper-based electronic devices [214].

Despite these advantages, the development of biomass-based carbons for supercapacitors has traditionally relied on trial-and-error methods, which are time-consuming and resource-intensive. This inefficiency not only slows innovation but also increases costs, making it difficult to fully realize the economic and environmental benefits of it. However, the integration of machine learning and data-driven research is now transforming this landscape. By leveraging large datasets consist of carbon electrodes excellent material features such as surface area, pore structure, chemical composition, and electrochemical performance. So, the machine learning algorithms can predict the properties of new materials with remarkable accuracy, often reducing experimental workload by more than half. Machine learning based model includes boosting and ensemble techniques are being used to identify optimal synthesis parameters, forecast SC, and guide the design of new carbon architectures before they are even produced in the lab [157].

The synergy between biomass, advanced carbonization techniques, and machine learning-driven optimization holds the promise of revolutionizing supercapacitor manufacturing. Automated synthesis platforms, guided by predictive models, can rapidly generate and characterize new materials, feeding real-time data back into the design loop and accelerating the discovery of high-performance, and low-cost supercapacitors[172]. This approach not only saves money, time, and resources but also minimizes environmental impact by reducing waste and energy consumption during production. The benefits of using biomass for making supercapacitors extend well beyond technical performance: their production valorizes agricultural waste, providing new revenue streams for farmers and supporting rural economies [215]. The use of non-toxic, abundant, and renewable materials aligns with global efforts to reliance on fossil fuels and minimize the ecological footprint of energy storage technologies [213]. Furthermore, the development of programmable and activated carbons from biomass-based precursors creates opportunities for interdisciplinary collaboration between materials scientists, data researchers, and engineers, fostering innovation across the energy sector.

Electrolyte selection is another critical factor in supercapacitor design, as the compatibility between the pore structure of biomass-derived carbons and the chosen electrolyte determines ion accessibility and ultimately device performance [91]. Studies have shown that the ideal micropore/mesopore ratio can be tailored to match specific electrolytes, such as PVA/KOH or organic liquids, further enhancing SC and energy density [216]. This level of customization, enabled by both material synthesis and data-driven optimization, ensures that biomass-based supercapacitors can meet the diverse requirements of modern energy applications. As the field continues to advance, the convergence of biomass, energy storage, and intelligent

design will drive the creation of next-generation supercapacitor devices that charge rapidly, store energy efficiently, and contribute to a more sustainable and equitable future [7].

## 7.2 Machine learning Based proposed framework for specific capacitance predictions

We used several tree-based and ensemble based supervised learning algorithms, including Random forest, Gradient boosting, XG boost, and Cat boost, to predict the specific capacitance of carbon-based supercapacitor electrodes obtained from biomass precursors. These models capture intricate, nonlinear relationships between the electrochemical performance metrics of interest and physicochemical features (e.g., specific surface area, pore size and volume, SC, and carbon content) [217].

### 7.2.1 Proposed algorithm

ML model's algorithm for optimization
Input
*M*: biomass derived carbon material features (specific surface area, pore size and volume, heteroatom doping, morphology etc.)
*E*: Electrochemical test conditions (electrolyte, current density, potential window etc.)
*H*: Historical dataset of material performance (experimental specific capacitance value)
*Λ*: Hyperparameter grid for search
*T*: Decision threshold for prediction error
Output
Predicted specific capacitance *SC* with feature importance score
1: Read carbon material properties *M*, electrochemical condition *E*, and historical specific capacitance data *H*.
2: Normalize features $x_i = \frac{x_{i-\min(x)}}{\max(x)-\min(x)}$ to handle missing values
3: Divide the dataset into training and testing subsets
4: Select regression model (Random forest, Gradient boost, XG boost, Cat boost)
5: Define hyperparameter grid search parameters (learning rate, depth, estimator, regularization)
6: Grid search optimization for each $\lambda \in \Lambda$
7: Train model $f(x,\ \lambda)$
8: Compute loss $L = RMSE\ (y, f(x,\ \lambda))$

9: Model selection $\lambda^* = arg\,min_{\lambda \in \Lambda} L$
10: Specific capacitance prediction $\widehat{SC} = f(x, \lambda^*)$
11. Feature importance score $\varphi_i = E[f(x)|x_i] - \mathrm{E}[f(x)]$
12: $SCI_c = \sum_{i=1}^{k} w_i.\ s_i + \widehat{SC}$
13: $SC = \arg max_{c \in C} SCI_c$ if $SCI_c \geq T$
12: Performance metrics evaluation metrics *(RMSE, $R^2$)*

### 7.2.2 Random forest regression

A Decision tree is a supervised learning model that serves in which every internal node confirms a specific input feature, the branches show the outcome, and the leaf nodes offer the final prediction. Nodes near the top of the tree usually play an important role in deciding the result.

The value prediction is determined by errors such mean square error and mean absolute error. At node, with $N$ number of samples $y_1, y_2 \ldots . y_N$

$$Mean\ square\ error = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2 \qquad \text{Equation 7.1}$$

$$Mean\ absolute\ error = \frac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}| \qquad \text{Equation 7.2}$$

$$\bar{y} = \frac{1}{N}\sum y_i \qquad \text{Equation 7.3}$$

To improve prediction accuracy, Random forest is an ensemble learning technique, combining the benefits of multiple Decision trees. Instead of relying on a single Decision tree, which often overfits training data, Random forest uses a technique known as bootstrapping to generate multiple Decision trees, each trained on a different random sample from the original dataset. Additionally, at each split in the tree, it randomly selects a subset of features rather than using all available features. This randomness introduces diversity among the trees, making the overall model more robust. Each tree gives its own prediction, and the final output is obtained by averaging the results in regression tasks [186].

### 7.2.3 XG boost

XG boost, also known as Extreme Gradient boosting, is a clever machine learning technique that constructs Decision trees sequentially in a particular order. First, the input features are assigned weights, which the first tree uses to generate predictions. When a prediction is incorrect, the model learns from it and moves on to the next tree, which attempts to correct it. Each tree is built upon the one that came before it. Finally, far more accurate results are produced by the combining result of all the trees.

The two main elements of the objective function that powers XG boost are the regularization term and the loss function. The loss function tells the model how well it is doing on the training data, while the regularization term keeps things simple to prevent the model from overfitting or memorizing the data too thoroughly. The mathematical computation for XG boost model is

The training data: $P = \{(x_i, y_i)\}_{i=1}^{n}$, the objective is to find the function $F(x_i)$ that minimizes the loss,

$$L = \sum_{i=1}^{n} l(y_i, F(x_i)) + \sum_{m=1}^{M} \Omega(h_m) \qquad \text{Equation 7.4}$$

$l$ is the loss function to measure difference between actual and predicted output value and $\Omega(h_m)$ is regularization function to penalize the complexity of the model.

$$\Omega(h_m) = cK + \frac{1}{2}\lambda \sum_{j=1}^{K} w_j^2 \qquad \text{Equation 7.5}$$

K is number of leaves in the tree, $w_j$ is the weight of leaf

Initialize the model, $F_0(x)$ is mean of $y$

$$F_0(x) = \arg\min_c \sum_{i=1}^{n} L(y_i, c) \qquad \text{Equation 7.6}$$

Now compute gradient and hessian function, at number of iterations $t$, compute the value for each $i$

$$g_i = \frac{\partial l(y_i, F_{t-1}(x_i))}{\partial F_{t-1}(x_i)}, h_i = \frac{\partial^2 l(y_i, F_{t-1}(x_i))}{\partial F_{t-1}(x_i)^2} \quad \text{Equation 7.7}$$

Now, build tree $h_t(x)$ to minimize the function

$$\sum_{j=1}^{K} \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + cK \quad \text{Equation 7.8}$$

The optimum weight for new leaf j is computed by function

$$w_j^* = \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad \text{Equation 7.9}$$

Update the model,

$$F_t(x) = F_{t-1}(x) + \rho . h_t(x) \quad \text{Equation 7.10}$$

$\rho$ is the learning rate,

The final model is computed by,

$$F_T(x) = F_0(x) + \rho . \sum_{t=1}^{T} h_t(x) \quad \text{Equation 7.11}$$

### 7.2.4 Gradient boost

Gradient boosting is a machine learning method that builds Decision trees one at a time, with each new tree trying to fix the mistakes made by the previous ones. It starts by making a basic guess usually the average of all target values. Next it examines the residual, the difference between that estimate and the actual values of output. To anticipate those mistakes, this tree has been trained. Each tree learns from its predecessor in this ongoing process. In the end, a more precise final prediction is produced by combining all the trees [188]. The loss function evaluates

the model's performance, and the regularization term keeps things straightforward to prevent overfitting.

The training data is defined as: $P = \{(x_i, y_i)\}_{i=1}^{n}$

The objective here is to find a function that helps to minimize the loss

$$Loss = \sum_{i=1}^{n} L(y_i, F(x_i)) \qquad \text{Equation 7.12}$$

Let initialize the model with constant:

$$F_0(x) = \arg\min_c \sum_{i=1}^{n} L(y_i, c) \qquad \text{Equation 7.13}$$

For boosting let assign boosting round as $m = 1, 2, \ldots M$

The gradient equation is:

$$g_i^{(m)} = -\left[\frac{\partial\, L(y_i, F(x_i))}{\partial\, F(x_i)}\right]_{F=F_{m-1}} \qquad \text{Equation 7.14}$$

The base learner $b_m(x)$ to the residuals $g_i^{(m)}$

The multiplier finds by equation:

$$y_m = \arg min_\gamma \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma\, b_m(x) \qquad \text{Equation 7.15}$$

The model is evaluated using functions

$$F_m(x) = F_{m-1}(x) - \delta y_m b_m(x) \qquad \text{Equation 7.16}$$

Where $\delta$ is the learning rate $\delta \in (0,1)$

### 7.2.5 Cat boost

Cat boost is dependable boosting algorithm that gradually constructs Decision trees to improve prediction accuracy. Its ability to handle numerical and categorical data with ease is what sets it apart and is very beneficial for real-world problems. Prior

to making predictions, the first tree gives the input features weights. The following tree notices any errors it makes and attempts to correct them. Each new tree learns from the one before it in this ongoing process, and eventually all the trees combine to produce a far more accurate result.

The objective function used by the model is composed of a regularization term and a loss function. Regularization prevents the model from becoming overly complicated, allowing it to remain accurate without overfitting, while the loss function evaluates the model's performance [218]. The mathematical formulation for Cat boost regression is:

Let the training data be $\{(x_i, y_i)\}_{i=1}^{n}$ to find a function $F(x)$ to predict the final output $y$ and minimize the error.

$$\widehat{y_i} = F(x) \qquad \text{Equation 7.17}$$

Where, $\hat{y}$ is the predicted output for sample i,

First, initialize the model with constant predictions.

$$F_0(x) = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad \text{Equation 7.18}$$

Then, compute the residual function at iteration $p$:

$$r_i^{(p)} = y_i - F_{p-1}(x_i) \qquad \text{Equation 7.19}$$

Train the Decision tree $h_p(x)$ to predict the residual value $r_i^{(p)}$.

Update the model, $\alpha$ is the learning rate

$$F_p(x_i) = F_{p-1}(x) + \alpha h_p(x) \qquad \text{Equation 7.20}$$

The final prediction after $P$ iterations

$$F_P(x) = F_0(x) + \alpha \sum_{p=1}^{P} h_p(x)$$ Equation 7.21

## 7.3 Experimental setting

In this section, we describe dataset as well as cleaning and preprocessing of raw data. The reason for the pre-processing is that unprocessed data is likely to result in poor, or even unreliable performance of learning algorithms.

### 7.3.1 Description of the dataset

To establish prediction models for predicting the performance of supercapacitors, comprehensive datasets were developed through an extensive review and extraction of data from previously published research articles, resulting in more than 1,000 data entries. The dataset comprises 9 important physicochemical properties of activated carbon derived from biomass-based precursors, reported in several studies [99], [212], [219], [220], [221]. The dataset entries contain different units to report key parameters. For example, specific surface area (SSA) was sometimes expressed in $m^2/g$ and other times in $cm^2/g$. To maintain consistency across the dataset, all values were converted to a common set of physical units wherever required. The final selected input features include pore size and pore volume, SSA, nitrogen and oxygen doping, potential window, Current density, D and G band ratio of Raman spectroscopy to detect defects in the material ($I_d/I_g$ ratio). The SC is treated as target output feature. All the independent physical and chemical features that are used to predict SC are explained below. In supercapacitors, there are two mechanisms through which charge storage occurs. The first mechanism is the electrochemical double layer capacitance (EDLC) in which physical ion adsorption occurs at the electrode-electrolyte interface. The second mechanism is pseudo-capacitance in which reversible redo reaction occurs at the surface of electrodes.

Therefore, the electrode material characteristics play an important role in determining SC [91].

The SSA is considered important feature for prediction of SC because it directly influences the amount of charge stored on the electrode surface. The higher SSA can create more infrastructure (pores hierarchy, and conductivity) for adsorption of ion leads to high SC, but only if pores are compatible with electrolyte and have proper structure for ion transport [29], [222]. Pore size of the electrode material decides which ion is accessible and how efficiently it allows ions to enter while pore volume decides numbers of ions can be accommodated on the material to ensure storage capacity [75]. Together, they regulate both SC and charge/discharge rate [216]. The $I_D/I_G$ value is taken to indicate the degree of disorder or defects in the carbon electrode material used for developing supercapacitors. The D-band (approx. 1350 $cm^{-1}$) is due to structural defects and disorder. Graphitic ($sp^2$-hybridized) carbon structure is the G-band (1580 $cm^{-1}$) [223]. A slight increase in defects (higher $I_d/I_g$) can improve the SC of electrodes by introducing additional active sites for the storage of charge. However, if defects are too high, it will disturb the electrical conductivity and structural stability leading to poor performance of the supercapacitor. Heteroatom doping introduces substitutions of non-carbon atoms (like nitrogen, and oxygen) into the carbon matrix framework of a material. Therefore, improving its properties, creating hydrophilic surfaces and ion mobility. This alteration creates active sites for ion interaction and double layer charge storage, significantly enhancing electrochemical performance and charge storage efficiency of supercapacitors [169]. The potential window is the stable operating range without electrolyte decomposition and electrode degradation for maximum energy storage. The current density decides how fast supercapacitor is charged and

discharged. The good electrode material will maintain high SC even when charge/ discharge cycle is very quick. It also influences the efficiency of ions access to pores under working conditions, affecting capacity retention and performance rate of supercapacitor [224]. Therefore, to maximize the supercapacitor's performance and ensure its long-term stability, these parameters should be balanced. After deciding energy storage device parameters that influence the SC performance, next step is to build the ML models to predict the performance of biomass derived supercapacitor.

### 7.3.2 Preprocessing of the dataset

The dataset consists of (1139 × 9) rows and columns. The missing values and outliers from the raw dataset were removed using code prior to final model development. However, some categorical features were dropped from raw dataset to mainly focused on numerical features to develop regression models. These columns were excluded from the analysis because they contained non-numeric or less relevant information. Only the most impactful features were selected for this study to simplify the model and make it more effective. This process of selecting relevant input variables has come under feature selection. It helps to avoid overfitting, as well as reduce noise and redundancy in the data.

### 7.3.3 Implementation detail for training and evaluation

After assembling dataset, and feature selection, the final dataset is divided into training dataset and testing dataset in the ratio of 80:20 to train the ML model and predict its results using both testing and training set. 80% data used for model development and remaining 20% used for independent performance evaluation. The various features of the biomass-based carbon as described in previous sections were used to train the ML models. Each model was optimized using the same input

conditions. To improve the robustness of models, cross validation was implemented using hyperparameter optimization technique. Subsequently, the ML models were further assessed using ML performance metrics using root mean square error (RMSE), and correlation coefficients ($R^2$) to ensure better models performance evaluation. The test dataset was finally used to validate the prediction applicability for practical energy storage devices.

### 7.3.4 Model evaluation

The ML model performance is evaluated on two performance metrics. These metrics are root mean square error (RMSE) and correlation coefficient $R^2$ score result. These metrics denote the model performance by analyzing the difference between true and predicted value of specific capacitance attained from the ML model.

In the data, actual value denoted as $x_i$ and predicted value as $y_i$. RMSE denotes error due to deviation square, therefore, it is important metric for checking accuracy of the model. It is represented by the equation.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \qquad \text{Equation 7.22}$$

Here, n is total number of observations in data. For good prediction, RMSE value should be low, showing minimum deviation between the actual and obtained value of specific capacitance.

The mean absolute error (MAE) is performance evaluation metrics to find the average value between predicted and actual value of specific capacitance. The smaller value indicates good performance of the model. It tells about how close the predicted value is from the actual value.

$$MAE = \sqrt{\frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n}} \quad \text{Equation 7.23}$$

The third performance metric is correlation coefficient ($R^2$ score). It represents the variance in predicted value by seeing the correlation with independent input features. The value near to 1 depicts good model performance. The metric helps in determining how model fits with the unknown dataset. The equation for $R^2$ score is

$$R^2 = 1 - \frac{\sum_{1=1}^{n}(x_i - y_i)^2}{\sum_{1=1}^{n}(x_i - \bar{x_i})^2} \quad \text{Equation 7.24}$$

Where $\bar{x_i}$ is the mean of actual value of specific capacitance.

## 7.4 Results and discussions

### 7.4.1 Statistical analysis

In this work, the box plot was made for statistical analysis of the data to describe all features used to predict specific capacitance of biomass-based carbon electrode is shown in Figure 7.1. “Figure 7.1” shows the comprehensive description of features, which defines the value range of features. From figure it can be seen that the SSA are concentrated between 0 to 3977.3 $m^2/g$ with the mean of 1558.73 $m^2/g$, The distribution of pore size mainly concentrated between 0 to 13 nm with the mean of 2.28 nm. The distribution of pore volume mainly concentrated between 0.003 $cm^3/g$ to 2.38 $cm^3/g$ with the mean of 0.90 $cm^3/g$. The nitrogen content, oxygen content, $I_d/I_g$ ratio, potential window and current density in the elemental analysis of feature are concentrated in 0- 13.4%, 0 – 31.2%, 0.46 - 2.74, 0 – 1.8 V, 0.05 – 100 A/g. These statistical variations across datasets likely originate from the inherent diversity and structural complexity of different biomass precursors. The specific capacitance is in the range of 1-550 F/g. From Fig 1, data span of specific

surface area is relatively high, this may be due to different types of biomass precursor and also due to magnitude [175].

In this chapter, RMSE metrics is employed as an important parameter for evaluating model performance during hyperparameter optimization. RMSE is robust in the presence of outliers, as it is less sensitive to deviation compared to other evaluation metrics like MAE and mean square error (MSE). The characteristics allow for stable and reliable assessment of model accuracy without removing outliers. This approach facilitates the retention of data information and enables more realistic assessment of model performance under real world practical conditions.

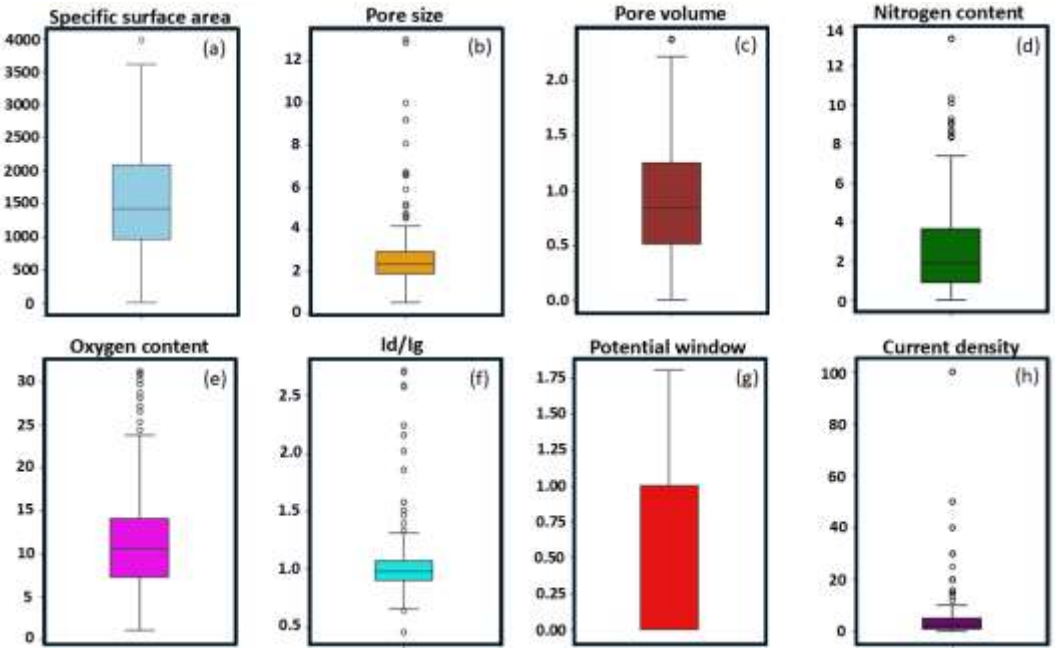


Figure 7.1 The statistical box plot of all features used in the study to predict specific capacitance.

### 7.4.2 Pearson correlation coefficients (PCC) analysis

The Pearson correlation coefficients (PCC) is a statistical measure used to determine the linear relationship between features. In this work, Figure 7.2 describes the PCC analysis among various input features and their correlation with the specific capacitance and among each other of supercapacitor. It highlights

strong positive correlation such as between SSA and pore volume (PV) (PCC = 0.85), and between SSA and specific capacitance (SC) have PCC value of 0.43 indicating positive correlation among them. The pore volume (PV) also has positive correlation with specific capacitance with PCC value of 0.45. This analysis helps to identify which features influence the supercapacitor's performance. This comprehensive approach increases the ML model ability to predict specific capacitance more accurately by taking into account both linear and non-linear effects, it also considered as supercapacitor performance evaluation.

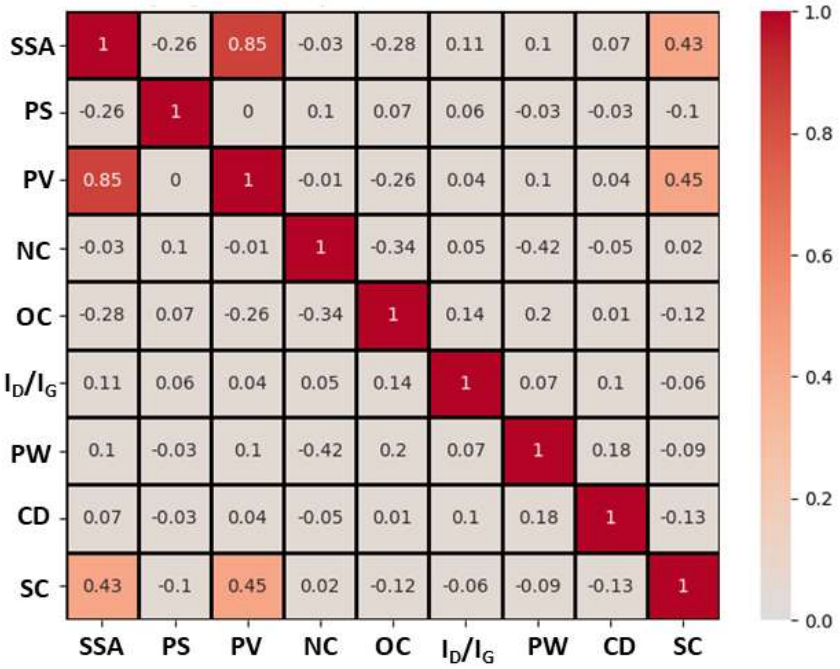


Figure 7.2 The Pearson correlation coefficient (PCC) plot of all features used in the study to predict specific capacitance.

(Notes: SSA – specific surface area, PS – Pore size, NC – Nitrogen content (wt %), OC – oxygen content (wt %), $I_D/I_G$, PW – potential window, CD – current density, SC – specific capacitance.)

### 7.4.3 Machine learning Model results analysis

The machine learning models were described in the method section above. The are mainly four ML models that are used to predict the specific capacitance of biomass-

derived carbon electrode for supercapacitor. The models are Random forest, XG boost, Cat boost, and Gradient boost. The three models used boosting techniques to build algorithms to predict the results in this case specific capacitance. The RMSE, MAE, and $R^2$ score evaluation metrics were used to evaluate and visualize the result as shown in Table 7.1 and Figure 7.3. The Cat boost model works best among all four ML models with lowest RMSE value of 22.06, lowest MAE value of 15.06, and highest $R^2$ score % value of 94.6. The Cat boost works best due to its ordered boosting techniques that help to build models to simulate future test data easily which will reduce overfitting and ensure better accuracy. It uses symmetric tree structure which is balanced that leads to fast prediction time and efficient regularization. It also handles missing data properly without any external care. The Cat boost, XG boost, Random forest, and Gradient boost work in descending order. The low RMSE value shows that the error is less between the actual specific capacitance and predicted specific capacitance. The $R^2$ score value used in regression to evaluate how well ML model fits the dataset, it explains the variance in the specific capacitance prediction. According to the evaluation matrix, it can clearly prove that the predicted performance of the Gradient boost model is not satisfactory compared to Random forest, XG boost and Cat boost. The errors of the Gradient boost model are relatively large and $R^2$ score % is much less than 100 %, which indicates that ML model cannot work accurately to predict the specific capacitance of supercapacitor.

In order to compare each ML model's performance, Figure 7.4 depicts the comparison of predicted and actual value of specific capacitance value through data visualization. In conclusion, the various ML models were applied but among all, the four ML models work best, and results analysis were done in detail to prove that

it predicts the specific capacitance of biomass-based carbon accurately. It can be seen in Figure 7.4 predicted and actual curve of high-performance ML model (Cat boost and XG boost) almost equal, while the predicted curve of Gradient boost fluctuates more. In conclusion, all four ML models like Random forest, XG boost, Cat boost, and Gradient boost successfully predict the specific capacitance with high accuracy. The results depicted that these models can effectively be used for prediction of specific capacitance without laboratory experiments. This can reduce laboratory time and cost, help in guide the fabrication process accordingly.

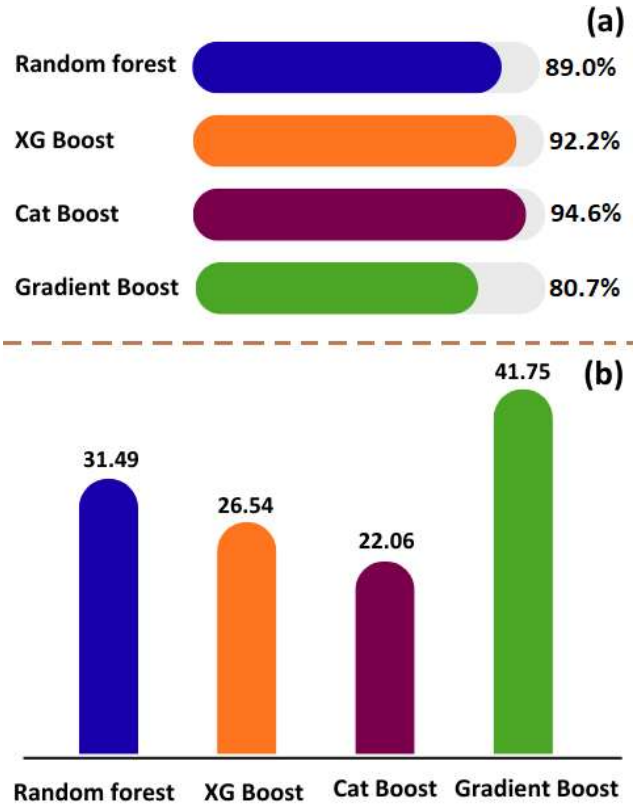


Figure 7.3 The Machine learning results of various models considered in the study in the form of bar chart (a) R2 score result with percentage values, (b) RMSE result with error values

- **Why do Cat boost models work best among other models?**

The Cat boost performs best due to its ability to identify and capture non-linear and non-uniform relationships and higher order feature interactions among various

physicochemical properties which affect supercapacitor performance. Furthermore, its ordered boosting technique and inbuilt regularization reduce overfitting and enhance model capability, especially for heterogeneous experimental dataset in this case biomass derived carbon electrode-based supercapacitor dataset.

Electrochemical performance in biomass derived carbon material depends on coupled physicochemical properties rather than single and isolated material parameters. For instance, measured specific capacitance is influenced by various parameters such as SSA, pore size, pore volume, degree of graphitization, and testing conditions also like potential window and current density. Among other models, Cat boost effectively and fast in learning these higher order feature interactions. In contrast, Random forest tends to smooth extreme specific capacitance values, while XG boost and Gradient boost models are relatively more sensitive to heterogeneity in dataset and hyperparameter tunning. The result of Gradient boost and Random forest suggests limited learning for predicting the relationship between material structure and specific capacitance.

Table 7.1 ML model's performance metrics results

| ML Model | RMSE | MAE | $R^2$ score % |
|---|---|---|---|
| Cat boost | 22.06 | 15.06 | 94.6 |
| XG boost | 26.54 | 17.86 | 92.2 |
| Gradient boost | 41.75 | 32.74 | 80.7 |
| Random forest | 31.49 | 21.93 | 89.0 |

### 7.4.4 The effect of input features on the model's performance

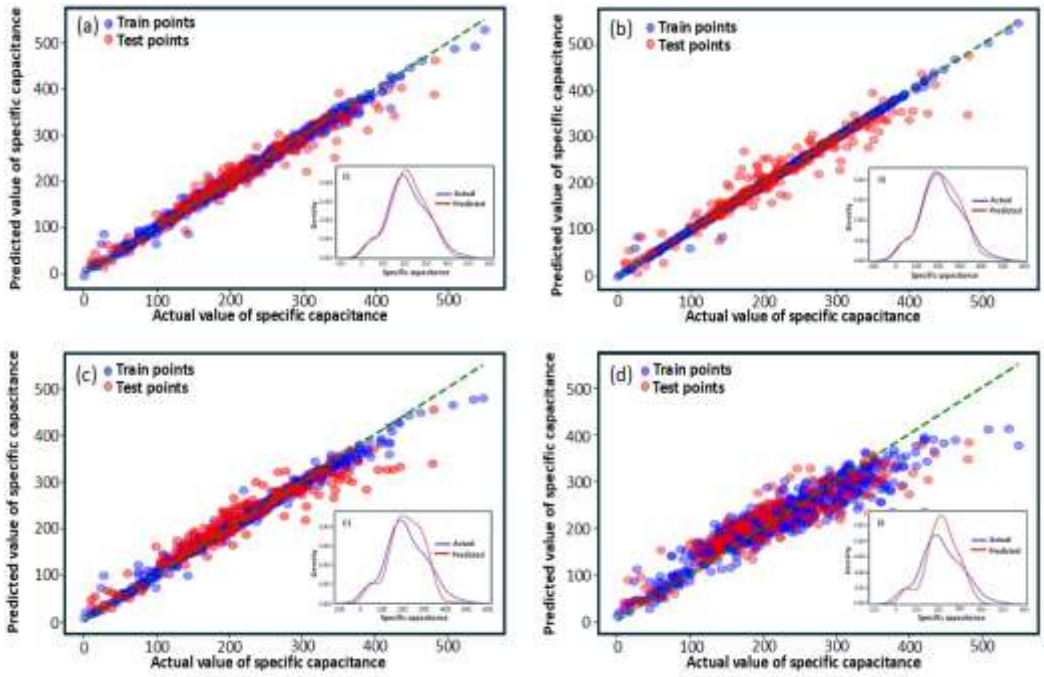


Figure 7.4 The Machine learning train and test set results of Actual value vs predicted value of specific capacitance with KDE plot of various ML models (a) Cat boost model, (b) XG boost (c) Random forest model result (d) Gradient boost model result

The effect of the biomass precursor physicochemical and electrochemical characteristics on the specific capacitance was ranked in Figure 7.5. The input features used to predict the specific capacitance of biomass-based carbon include specific surface area, pore size and volume, the ratio of Id/Ig, heteroatom doping, potential window, and current density. These factors influence the model's performance to predict specific capacitance of supercapacitor. The result showed that pore volume was the most influential factor affecting specific capacitance of Cat boost model accounting for feature importance score of 18.85, followed by the specific surface area and nitrogen content with the feature importance score of 17.29 and 13.98. The lowest influential feature of Cat boost model is pore size with the feature importance score of 7.5.

Table 7.2 Feature importance score of the important ML models

| Features | Cat boost | XG boost | Random forest | Gradient boost |
|---|---|---|---|---|
| Specific surface area | 17.29 | 0.08 | 0.22 | 0.23 |
| Pore size | 7.5 | 0.056 | 0.06 | 0.05 |
| Pore volume | 18.85 | 0.41 | 0.27 | 0.33 |
| Nitrogen content | 13.98 | 0.09 | 0.1 | 0.09 |
| Oxygen content | 11.52 | 0.08 | 0.08 | 0.05 |
| $I_d/I_g$ ratio | 13.66 | 0.11 | 0.11 | 0.09 |
| Potential window | 8.28 | 0.12 | 0.02 | 0.03 |
| Current density | 8.87 | 0.02 | 0.11 | 0.08 |

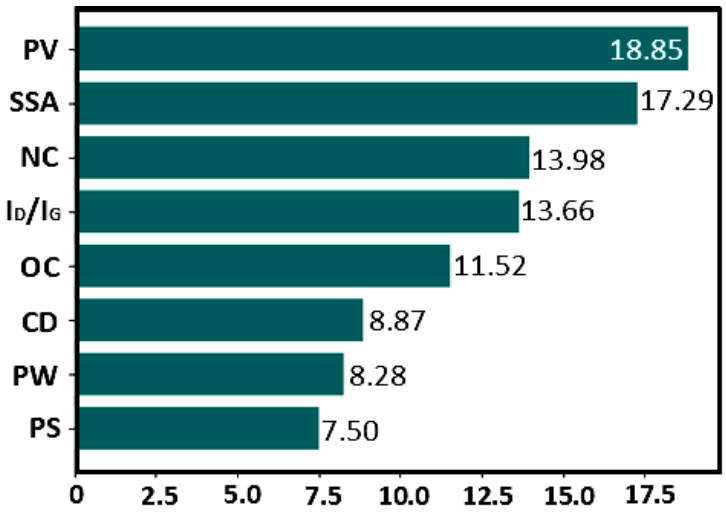


Figure 7.5 Machine learning feature importance score of Cat boost Model

Similarly, for the XG boost model, the pore volume and potential window are the most influential features with the importance score of 0.41 and 0.12, and the least influential features are pore size and current density with the importance score of 0.05 and 0.02. For the Random forest model, pore volume has the highest importance score of 0.27 and potential window with the lowest importance score of 0.02. For the Gradient boost model, the highest influential feature is pore volume

and lowest influential feature is potential window with the score value of 0.33 and 0.03. As illustrated in the Pearson correlation coefficient matrix in Figure 7.3 and Table 7.2, both pore volume and specific surface area exhibit positive correlation with the specific capacitance of biomass-based carbon electrode. Hence, it is proved here also that these features demonstrate high feature importance score, highlighting their strong influence on performance of supercapacitor.

### 7.4.5 Synergistics effect between electrochemical performance and ML results

The Cat boost model performance in predicting the specific capacitance of biomass derived carbon electrode material can be driven by both the factors which are intrinsic characteristics of the dataset and non-linear complex electrochemical behavior which governs supercapacitor performance. Among all the discussed models, Cat boost achieved the highest predictive accuracy with best performance metrics. The electrochemical performance of supercapacitors is highly affected by physicochemical properties and testing conditions rather than one factor. In the considered dataset, the input features influence charge storage through interconnected phenomenon involving electric double layer capacitance, ion diffusion, electrolyte behavior, conductivity, and pseudocapacitive surface reactions. The statistical result depicts the SSA is most influential feature which ranges from 6 to 3977 $m^2/g$ with standard deviation of 861.22. Such a high distribution depicts variability in pore development and activation conditions among the biomass derived carbon. However, high SSA does not always give the high specific capacitance because other factors also play important role such as excess microporous structure also isolate ion transport and electrolyte accessibility.

Therefore, specific capacitance depends on combined interaction between SSA and pore architecture

The feature importance further supports the obtained results, where pore volume of importance score 18.85, SSA has importance score of 17.29 of the best models among all, these have emerged as dominant features, depicting the important role of hierarchical porosity and accessibility of electrolyte. The nitrogen content, oxygen content and Id/Ig ration reflect the contribution of heteroatom doping in improving conductivity, surface wettability and graphitization degree improve transport of electrons. The contribution of all features suggests that specific capacitance is governed by couple interactions of the features rather than single feature. Cat boost model effectively captures these complex feature patterns through order boosting mechanism and symmetrical Decision trees, allowing more efficient learning of electrochemical performance.

### 7.4.6 Comparison with the related work

Table 7.3 presents the novelty of this study by comparing previously published work on predicting supercapacitor performance using different types of electrode materials. The results highlight that Cat boost model achieves the best performance for biomass derived carbon electrode, a sustainable and eco-friendly material.

Table 7.3 Comparison table with previously published studies

| S. No. | Material | Features | Models | Error metrics | |
|---|---|---|---|---|---|
| 1. | Carbon | 6 | Random forest | RMSE-38.13, $R^2$ score– 0.68 | [208] |
| 2. | Carbon | 8 | Random tree | RMSE - 67.62, R score– 0.76 MAE – 52.03 | [128] |
| 3. | Activated Carbon | 8 | Linear regression | RMSE - 97.10, R score– 0.53, MAE – 76.89 | [128] |

| | | | | | |
|---|---|---|---|---|---|
| 4. | Biomass derived carbon | 7 | DNN | RMSE –29.198, $R^2$ score–0.822, MAE – 19.922 | [175] |
| 5. | Biomass derived carbon | 7 | Light GBM | RMSE –17.756, $R^2$ score– 0.95, MAE – 11.09 | [175] |
| 6. | Carbon nanotube | 9 | Random forest | RMSE – 39.01, $R^2$ score– 0.79, | [195] |
| 7. | Porous Carbon | 11 | XG boost | RMSE – 25.50, $R^2$ score–0.892, MAE – 19.56 | [153] |
| 8. | Porous Carbon | 11 | Random forest | RMSE –35.715, $R^2$ score–0.856, MAE – 29.268 | [153] |
| 9. | Various material composition | 8 | Random forest | RMSE – 27.13, $R^2$ score–0.972, MAE – 19.95 | [132] |
| 10. | Biomass derived carbon | 9 | Cat boost | RMSE – 22.06, $R^2$ score–0.946, MAE – 15.06 | This work |

### 7.4.7 Challenges with problem characteristics

The development of biomass derived carbon electrode for supercapacitors faces both material and methodological challenges. To address these challenges, novel predictive model is designed to overcome trial and error method to data driven optimization using material characteristics and features. The main challenges are material complexity, inadequate material features and experimental inefficiency to predict the performance of the carbon electrodes.

The proposed method overcomes the issues by applying Machine learning techniques. The first method was to utilize source features which means measuring properties before material was prepared. This method uses basic composition information of biomass material such as input condition, structural composition and elemental analysis which were easier to find for experimental verification. The second challenge was the selection of optimized and efficient ML algorithms. The method employed Cat boost outperforms over other ML algorithms for the

considered dataset. The Cat boost performs better because it can efficiently process large datasets, reduce memory consumption and build accurate trees faster than traditional models. The Cat boost model achieved R2 score % of 94.6 and RMSE of 22.06 provide accurate prediction than other models. The gap between basic raw material and electrochemical performance was filled with prediction modelling, reduced experimental costs and provided theoretical analysis for designing electrodes for high performance supercapacitors.

## 7.5 Conclusion and Future outlook

This chapter describes a novel methodology that utilizes the physicochemical properties of biomass-based carbon electrode for supercapacitor. By applying ML models, including Cat boost, XG boost, Gradient boost and Random forest, the proposed methodology effectively predicted the specific capacitance of supercapacitor. Among the tested ML models, Cat boost and XG boost outperformed the other two ML models. With the Cat boost achieving the lowest RMSE value of 22.06 and highest $R^2$ score % of 94.6. The key input features that are most influential features are pore volume and specific surface area. The results aligned with previous studies win the ML approach for predicting supercapacitor performance, highlighting Cat boost suitability for specific capacitance predictions. This method not only offers insight for optimizing biomass derived supercapacitor performance but also serves as a predictive framework for other key input features, offering broad application in energy storage research. The future research would validate findings by fabricating biomass derived supercapacitor electrode, demonstrate that it reduces cost and time in experiments. Expand datasets, refine ML models, and integrate advances analysis to increase robustness, and broaden research impact in this field.

# Chapter 8

# Conclusion

*[This chapter concluded different frameworks to optimize the development of supercapacitors and other energy storage devices. These data-driven models optimize synthesis parameters and simulate operational behavior, reducing experimental reliance. The models can be used to guide the performance predictions of different properties of energy storage devices such as energy density, conductivity, and life cycle and optimizing material synthesis parameters, thereby reducing reliance on costly laboratory trial-and-error.]*

## 8.1 Summary of the research

The thesis systematically describes the development, synthesis, characterization, and performance prediction of activated carbon electrode derived from various precursors like natural fiber and biomass for supercapacitor applications, with main emphasis on using machine learning (ML) technique as a tool for performance predictions. By mixing different knowledge domain derived from material science, electrochemistry and machine learning, the present research relies on both theoretical and practical implementation of sustainable energy storage technologies.

The research flows with focusing on reviewing and analyzing carbon nanomaterials and activated carbon obtained from various renewable resources as promising electrode materials for supercapacitors. It was observed that the physical and chemical properties of carbon such as electrical conductivity, pore size structure, cyclic stability, and large specific surface area make it highly compatible for energy storage applications. The review describes that electrochemical performance is mainly dependent on the structural and morphological characteristics of developed carbon nanomaterial formed, which mostly depends on choice of precursors, synthesis methods, and activation process. The different properties such as pore size distribution, specific surface area, surface chemistry and material doping were known as critical factors influencing charge storage mechanisms, specific capacitance, energy density, and power density.

The major outcome obtained from the analysis is that the small modification in carbon structure like nitrogen doping, introduction of defects, and controlled pore development can lead to significant increases in performance of supercapacitors. These findings underline the importance of electrode material for supercapacitors

design, particularly those that require sustainable and environmentally friendly energy storage solutions.

After understanding precursors and electrode materials, the main contribution of the thesis is to develop and evaluate ML models to predict the specific capacitance of carbon electrode originating from various sources for supercapacitors. The experimental datasets were extracted from previously published research articles and formed to include important physicochemical and electrochemical properties of electrode material such as specific surface area, pore size, nitrogen doping, potential window, and current density. These descriptors were used as input features for classification and regression ML models with specific capacitance as output features.

After formation of dataset, an exploratory data analysis was performed to analyze and understand the statistical distribution of features and their correlations with each other. This analysis identifies correlations between parameters and their electrochemical performance as well as potential elimination among various features. After following data exploration, multiple machine learning classification and regression models were used to predict specific capacitance of activated carbon material electrode. These models were trained and use validation techniques like cross-validation strategies to ensure equality of models to predict results, and to reduce overfitting

The comprehensive comparison of multiple ML models demonstrated the Random forest and boosting algorithms consistently outperformed linear and single tree-based models. The best predictive accuracy was obtained with low root mean square error (RMSE), mean absolute error (MAE), strong correlation coefficient ($R^2$),

verifying the suitability of machine learning for complexity of model, non-linear relationships between material properties and their performance.

The ML model results show that physicochemical properties can successfully predict the specific capacitance from various precursors in different test environments and conditions. These models perform better than the traditional empirical approaches. This proves that the data-driven methodology is effective in solving complex problems that control charge storage in carbon nanomaterials.

For the deep analysis of ML model predictions, shapley additive explanation (SHAP) analysis and feature prioritization score analysis was employed to assess the contribution of individual features to predict the output specific capacitance. The interpretability analysis depicts specific surface area, and pore size distribution plays dominant role in determining specific capacitance, which aligns consistent with the charge storage mechanism. Furthermore, surface chemistry related features including heteroatom doping were recognized as influencing feature to affect specific capacitance, through the enhancement of wettability and pseudocapacitive effect. These results confirmed that the ML models capture physically meaningful relationships rather than coincidental correlations.

The integration of classification and regression methods has offered an additional advantage by first identifying the probable range in which specific capacitance lies and thereafter predicting numerical values. These two step approaches improve model reliability for high performance electrode predictions while simultaneously illustrate the potential of ML assisted screening techniques in material selection and fabrication process.

Overall, the thesis establishes that machine learning offers a powerful and effective alternative to empirical trial and error experimental methods. These methodological

enhancements are not only enabling the identification of novel materials suitable for carbon electrodes but also optimize the procedure, making it efficient and less resource consuming. This progress holds significant potential for future of electrode material performance in energy storage technologies.

From sustainable development perspective, research supports the use of natural fiber and biomass-based carbon material as eco-friendly, easy availability, and low-cost alternatives with data-driven optimization techniques combine with effort move towards green energy technologies and material economic cost as also depicted in schematic diagram in fig. 1.

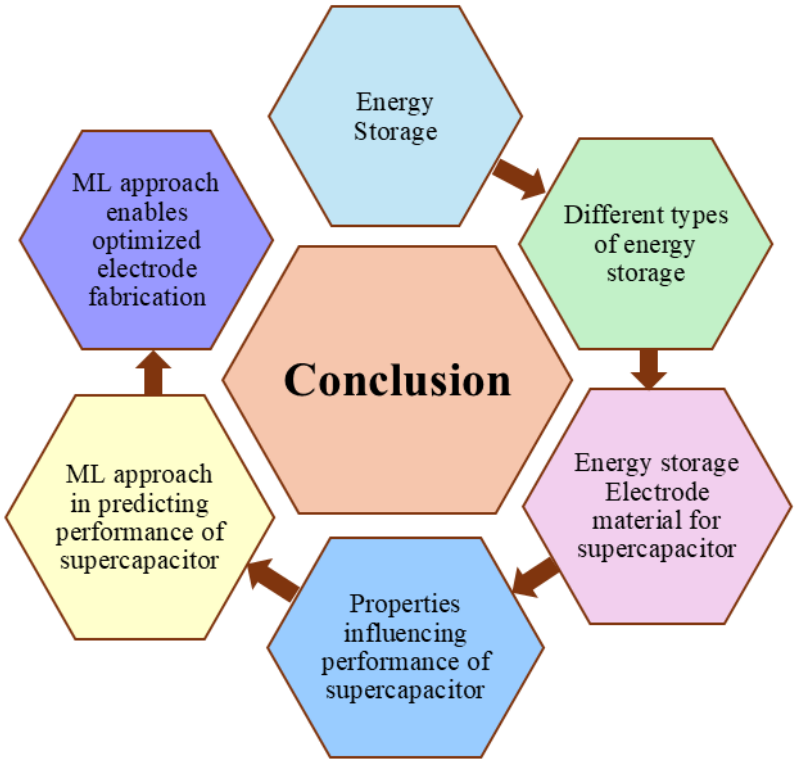


Figure 8.1 Schematic representation of data-driven machine learning based performance prediction of supercapacitor.

## 8.2 Future research goals

Despite the promising results, several other opportunities persist for future research. The dataset with newly experimentally validated natural fiber-derived carbon nanomaterials would be explored and further enhance the model robustness. Incorporating other physicochemical and electrochemical features related to its synthesis methods, electrolyte composition and type, device configuration could improve prediction accuracy and practical relevance, to also find cycle stability and energy density of supercapacitor. The future research would also use other explainable artificial intelligence methodology that can further deepen the understanding of structure, property and performance relationships.

In future work, the trained ML model can be used to predict the performance of newly designed graphene and carbon composite electrodes obtained from natural resources prior to fabricating the device, therefore assisting the experimental process more efficiently.

In conclusion, the thesis demonstrates the integration of material engineering and data-driven machine learning methods to provide scalable, robust and cost-effective pathways for supercapacitor technology. The findings provide valuable insight for researchers working towards the innovative development of high-performance, sustainable and next generation energy storage systems.

# REFERENCES


[1] J. Chang *et al.*, "Activated porous carbon prepared from paulownia flower for high performance supercapacitor electrodes," *Electrochim. Acta*, vol. 157, 2015, doi: 10.1016/j.electacta.2014.12.169.

[2] J. Singh, S. Srivastava, S. Shrivas, R. Pandey, and A. Dubey, "Comparative study of Structures of Electrochromic Device for Flexible Electrochromic Display," in *2023 International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT)*, 2023, pp. 290–295. doi: 10.1109/DICCT56244.2023.10110119.

[3] H. Chen *et al.*, "Nitrogen-doped porous carbon using ZnCl2 as activating agent for high-performance supercapacitor electrode materials," *J. Mater. Sci.*, vol. 53, no. 4, 2018, doi: 10.1007/s10853-017-1453-3.

[4] A. Dubey, X. Tao, C. Cochrane, and V. Koncar, "Textile based three-layer robust flexible and stable electrochromic display," *IEEE Access*, vol. 8, pp. 182918–182929, 2020, doi: 10.1109/ACCESS.2020.3028655.

[5] J. Wang *et al.*, "Biomass derived carbon for energy storage devices," *Royal Society of Chemistry*, 2017, doi: 10.1039/c6ta08742f.

[6] L. Zhang and X. S. Zhao, "Carbon-based materials as supercapacitor electrodes," *Chem. Soc. Rev.*, vol. 38, no. 9, 2009, doi: 10.1039/b813846j.

[7] A. Dubey *et al.*, "Nano iron pyrite (FeS2) exhibits bi-functional electrode character," *RSC Adv.*, vol. 6, no. 20, pp. 16859–16867, 2016, doi: 10.1039/c6ra01973k.

[8] G. B. Barin, I. De Fátima Gimenez, L. P. Da Costa, A. G. S. Filho, and L. S. Barreto, "Influence of hydrothermal carbonization on formation of curved

graphite structures obtained from a lignocellulosic precursor," *Carbon N. Y.*, vol. 78, pp. 609–612, Nov. 2014, doi: 10.1016/J.CARBON.2014.07.017.

[9] E. Taer, M. Deraman, I. A. Talib, A. A. Umar, M. Oyama, and R. M. Yunus, "Physical, electrochemical and supercapacitive properties of activated carbon pellets from pre-carbonized rubber wood sawdust by CO2 activation," *Current Applied Physics*, vol. 10, no. 4, pp. 1071–1075, 2010, doi: https://doi.org/10.1016/j.cap.2009.12.044.

[10] B. Tulachan *et al.*, "Electricity from the silk cocoon membrane," *Sci. Rep.*, vol. 4, Jun. 2014, doi: 10.1038/srep05434.

[11] A. Dubey, D. Philip, and M. Das, "A sustainable, eco-friendly charge storage device from bio-charred jute: An innovative strategy to empower the jute farmers of India," in *Smart Innovation, Systems and Technologies*, Springer Science and Business Media Deutschland GmbHpp. 397–407, , 2017, doi: 10.1007/978-981-10-3521-0_34.

[12] Z. Wang, D. Shen, C. Wu, and S. Gu, "State-of-the-art on the production and application of carbon nanomaterials from biomass," *Green Chemistry*, vol. 20, no. 22, pp. 5031–5057, 2018, doi: 10.1039/c8gc01748d.

[13] A. Dubey *et al.*, "Biocharring of natural fibers of insect and plant origin: a green route for the production of 'carbon-based charge storage nanomaterials,'" *Mater. Renew. Sustain. Energy*, vol. 7, no. 4, Nov. 2018, doi: 10.1007/s40243-018-0127-7.

[14] B. Liu, Y. Liu, H. Chen, M. Yang, and H. Li, "Oxygen and nitrogen co-doped porous carbon nanosheets derived from Perilla frutescens for high volumetric performance supercapacitors," *J. Power Sources*, vol. 341, 2017, doi: 10.1016/j.jpowsour.2016.12.022.

[15] W. J. Liu, K. Tian, Y. R. He, H. Jiang, and H. Q. Yu, “High-yield harvest of nanofibers/mesoporous carbon composite by pyrolysis of waste biomass and its application for high durability electrochemical energy storage,” *Environ. Sci. Technol.*, vol. 48, no. 23, 2014, doi: 10.1021/es504184c.

[16] Y. Cao *et al.*, “Hierarchical porous activated carbon for supercapacitor derived from corn stalk core by potassium hydroxide activation,” *Electrochim. Acta*, vol. 212, pp. 839–847, Sep. 2016, doi: 10.1016/J.ELECTACTA.2016.07.069.

[17] M. Farrokhnia, A. Safekordi, M. Rashidzadeh, G. Khanbabaei, R. Akbari Anari, and M. Rahimpour, “Development of porous nanocomposite membranes for gas separation by identifying the effective fabrication parameters with Plackett–Burman experimental design,” *Journal of Porous Materials*, vol. 23, no. 5, 2016, doi: 10.1007/s10934-016-0187-y.

[18] A. Tripathy, S. Mohanty, S. K. Nayak, and A. Ramadoss, “Renewable banana-peel-derived activated carbon as an inexpensive and efficient electrode material showing fascinating supercapacitive performance,” *J. Environ. Chem. Eng.*, vol. 9, no. 6, Dec. 2021, doi: 10.1016/j.jece.2021.106398.

[19] L. Ge, Y. Wu, F. Wang, and Y. Huang, “Algae-Derived Nitrogen Self-Doped Porous Carbon Materials with High Supercapacitor Performances,” *Energy and Fuels*, vol. 35, no. 18, 2021, doi: 10.1021/acs.energyfuels.1c01827.

[20] K. Liang *et al.*, “Peanut shell waste derived porous carbon for high-performance supercapacitors,” *J. Energy Storage*, vol. 70, p. 107947, Oct. 2023, doi: 10.1016/J.EST.2023.107947.

[21] G. Qiu *et al.*, “Bamboo-based hierarchical porous carbon for high-performance supercapacitors: the role of different components,” *Colloids Surf. A Physicochem. Eng. Asp.*, vol. 650, p. 129575, Oct. 2022, doi: 10.1016/J.COLSURFA.2022.129575.

[22] D. Khalafallah, X. Quan, C. Ouyang, M. Zhi, and Z. Hong, “Heteroatoms doped porous carbon derived from waste potato peel for supercapacitors,” *Renew. Energy*, vol. 170, pp. 60–71, Jun. 2021, doi: 10.1016/J.RENENE.2021.01.077.

[23] A. Volperts, G. Dobele, A. Zhurinsh, D. Vervikishko, E. Shkolnikov, and J. Ozolinsh, “Wood-based activated carbons for supercapacitor electrodes with a sulfuric acid electrolyte,” *New Carbon Materials*, vol. 32, no. 4, pp. 319–326, Aug. 2017, doi: 10.1016/S1872-5805(17)60125-2.

[24] X. Song, X. Ma, Y. Li, L. Ding, and R. Jiang, “Tea waste derived microporous active carbon with enhanced double-layer supercapacitor behaviors,” *Appl. Surf. Sci.*, vol. 487, pp. 189–197, Sep. 2019, doi: 10.1016/j.apsusc.2019.04.277.

[25] B. Liu, Y. Liu, H. Chen, M. Yang, and H. Li, “Oxygen and nitrogen co-doped porous carbon nanosheets derived from Perilla frutescens for high volumetric performance supercapacitors,” *J. Power Sources*, vol. 341, 2017, doi: 10.1016/j.jpowsour.2016.12.022.

[26] J. C. Acomb, C. Wu, and P. T. Williams, “The use of different metal catalysts for the simultaneous production of carbon nanotubes and hydrogen from pyrolysis of plastic feedstocks,” *Appl. Catal. B*, vol. 180, pp. 497–510, Jan. 2016, doi: 10.1016/j.apcatb.2015.06.054.

[27] P. Mahanandia and K. K. Nanda, “A one-step technique to prepare aligned arrays of carbon nanotubes,” *Nanotechnology*, vol. 19, no. 15, Apr. 2008, doi: 10.1088/0957-4484/19/15/155602.

[28] F. Chen *et al.*, “Preparation of High-Performance Porous Carbon Materials by Citric Acid-Assisted Hydrothermal Carbonization of Bamboo and Their Application in Electrode Materials,” *Energy and Fuels*, vol. 36, no. 16, pp. 9303–9312, Aug. 2022, doi: 10.1021/acs.energyfuels.2c01828.

[29] Z. Hu and M. P. Srinivasan, “Preparation of high-surface-area activated carbons from coconut shell,” *Microporous and Mesoporous Materials*, vol. 27, no. 1, pp. 11–18, Jan. 1999, doi: 10.1016/S1387-1811(98)00183-8.

[30] L. Sun *et al.*, “Isolated boron and nitrogen sites on porous graphitic carbon synthesized from nitrogen-containing chitosan for supercapacitors,” *ChemSusChem*, vol. 7, no. 6, pp. 1637–1646, 2014, doi: 10.1002/cssc.201400048.

[31] R. Pandey, A. Dubey, and M. K. Sinha, “5 - Lotus fibre drawing and characterization,” in *Sustainable Fibres for Fashion and Textile Manufacturing*, R. Nayak, Ed., in The Textile Institute Book Series, Woodhead Publishing, pp. 95–108, 2023, doi: https://doi.org/10.1016/B978-0-12-824052-6.00001-9.

[32] R. Pandey, M. K. Sinha, and A. Dubey, “6 - Macrophyte and wetland plant fibres,” in *Sustainable Fibres for Fashion and Textile Manufacturing*, R. Nayak, Ed., in The Textile Institute Book Series., Woodhead Publishing, pp. 109–127, 2023, doi: https://doi.org/10.1016/B978-0-12-824052-6.00006-8.

[33] M. Li *et al.*, “Recent advancements of plant-based natural fiber–reinforced composites and their applications,” 2020. doi: 10.1016/j.compositesb.2020.108254.

[34] R. Pandey *et al.*, “A Comparative Study of Phenotypic Variability and Physico-Mechanical Properties of Dual-Purpose Flax Fiber Varieties in India,” *Journal of Natural Fibers*, vol. 19, no. 17, pp. 15680–15689, 2022, doi: 10.1080/15440478.2022.2133048.

[35] D. and D. M. Dubey Amarish and Philip, “A Sustainable, Eco-friendly Charge Storage Device from Bio-charred Jute: An Innovative Strategy to Empower the Jute Farmers of India,” in *Research into Design for Communities, Volume 2*, D. Chakrabarti Amaresh and Chakrabarti, Ed., Singapore: Springer Singapore, 2017, pp. 397–407.

[36] M. Roy *et al.*, “Soft magnetic memory of silk cocoon membrane,” *Sci. Rep.*, vol. 6, Jul. 2016, doi: 10.1038/srep29214.

[37] S. Parida and D. P. Dutta, “Nanostructured Materials from Biobased Precursors for Renewable Energy Storage Applications,” in *Biorenewable Nanocomposite Materials, Vol. 1: Electrocatalysts and Energy Storage*, vol. 1410, in ACS Symposium Series, vol. 1410, American Chemical Society, ch. 13, pp. 307–366, 2022, doi: doi:10.1021/bk-2022-1410.ch013.

[38] H. Jirimali, J. Singh, R. Boddula, J.-K. Lee, and V. Singh, “Nano-Structured Carbon: Its Synthesis from Renewable Agricultural Sources and Important Applications,” *Materials*, vol. 15, no. 11, 2022, doi: 10.3390/ma15113969.

[39] S. Ravi and S. Vadukumpully, “Sustainable carbon nanomaterials: Recent advances and its applications in energy and environmental remediation,” *J.*

*Environ. Chem. Eng.*, vol. 4, no. 1, pp. 835–856, 2016, doi: https://doi.org/10.1016/j.jece.2015.11.026.

[40] B. Yan *et al.*, "High performance supercapacitors based on wood-derived thick carbon electrodes synthesized via green activation process," *Inorg. Chem. Front.*, vol. 9, no. 23, 2022, doi: 10.1039/d2qi01914k.

[41] A. Gholampour and T. Ozbakkaloglu, "A review of natural fiber composites: properties, modification and processing techniques, characterization, applications," 2020. doi: 10.1007/s10853-019-03990-y.

[42] J. Tian, T. Zhang, D. Talifu, A. Abulizi, and Y. Ji, "Porous carbon materials derived from waste cotton stalk with ultra-high surface area for high performance supercapacitors," *Mater. Res. Bull.*, vol. 143, p. 111457, Nov. 2021, doi: 10.1016/j.materresbull.2021.111457.

[43] H. Liu *et al.*, "Lignin-based electrodes for energy storage application," *Ind. Crops Prod.*, vol. 165, 2021, doi: 10.1016/j.indcrop.2021.113425.

[44] M. Jawaid and H. P. S. Abdul Khalil, "Cellulosic/synthetic fibre reinforced polymer hybrid composites: A review," Aug. 01, 2011. doi: 10.1016/j.carbpol.2011.04.043.

[45] L. Jino *et al.*, "Review on natural fibre composites reinforced with nanoparticles," *Mater. Today Proc.*, 2023, doi: 10.1016/j.matpr.2023.01.126.

[46] S. Darmawan, N. J. Wistara, G. Pari, A. Maddu, and W. Syafii, "Characterization of Lignocellulosic Biomass as Raw Material for the Production of Porous Carbon-based Materials." *BioResources*, *11*(2), 2016, doi: https://doi.org/10.15376/biores.11.2.3561-3574

[47] A. Vilén, P. Laurell, and R. Vahala, “Comparative life cycle assessment of activated carbon production from various raw materials,” *J. Environ. Manage.*, vol. 324, 2022, doi: 10.1016/j.jenvman.2022.116356.

[48] R. Pandey *et al.*, “Physico-Chemical Characterization of Lignocellulosic Seed Microfibers,” *Journal of Natural Fibers*, vol. 21, no. 1, 2024, doi: 10.1080/15440478.2024.2360493.

[49] F. Wang, D. Ouyang, Z. Zhou, S. J. Page, D. Liu, and X. Zhao, “Lignocellulosic biomass as sustainable feedstock and materials for power generation and energy storage,” *Journal of Energy Chemistry*, vol. 57, pp. 247–280, Jun. 2021, doi: 10.1016/j.jechem.2020.08.060.

[50] C. E. Wyman, “Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals.” Hoboken, NJ, USA: *John Wiley & Sons,* 2013. doi: 10.1002/9780470975831.

[51] Z. Wang, P. Tammela, M. Strømme, and L. Nyholm, “Cellulose-based Supercapacitors: Material and Performance Considerations,” *Adv. Energy Mater.*, vol. 7, no. 18, Sep. 2017, doi: 10.1002/aenm.201700130.

[52] M. M. Pérez-Madrigal, M. G. Edo, and C. Alemán, “Powering the future: Application of cellulose-based materials for supercapacitors,” *Royal Society of Chemistry*, 2016, doi: 10.1039/c6gc02086k.

[53] Z. Sun *et al.*, “Overview of cellulose-based flexible materials for supercapacitors,” *Royal Society of Chemistry*, Mar. 28, 2021, doi: 10.1039/d0ta10504j.

[54] H. Lin, Y. Liu, Z. Chang, S. Yan, S. Liu, and S. Han, “A new method of synthesizing hemicellulose-derived porous activated carbon for high-

performance supercapacitors," *Microporous and Mesoporous Materials*, vol. 292, Jan. 2020, doi: 10.1016/j.micromeso.2019.109707.

[55] S. Mehta, S. Jha, and H. Liang, "Lignocellulose materials for supercapacitor and battery electrodes: A review," *Elsevier Ltd*., Dec. 01, 2020, doi: 10.1016/j.rser.2020.110345.

[56] A. M. Abioye and F. N. Ani, "Recent development in the production of activated carbon electrodes from agricultural waste biomass for supercapacitors: A review," *Elsevier Ltd*., Dec. 01, 2015, doi: 10.1016/j.rser.2015.07.129.

[57] L. Manjakkal *et al.*, "Sustainable electrochemical energy storage devices using natural bast fibres," *Elsevier B.V.*, Jun. 01, 2023, doi: 10.1016/j.cej.2023.142845.

[58] L. Lyu *et al.*, "Recent development of biomass-derived carbons and composites as electrode materials for supercapacitors," *Royal Society of Chemistry*., Dec. 01, 2019, doi: 10.1039/c9qm00348g.

[59] Y. Zhai, Y. Dou, D. Zhao, P. F. Fulvio, R. T. Mayes, and S. Dai, "Carbon Materials for Chemical Capacitive Energy Storage," *Advanced Materials*, vol. 23, no. 42, pp. 4828–4850, Nov. 2011, doi: 10.1002/adma.201100984.

[60] H. Wei *et al.*, "Advanced porous hierarchical activated carbon derived from agricultural wastes toward high performance supercapacitors," *J. Alloys Compd.*, vol. 820, 2020, doi: 10.1016/j.jallcom.2019.153111.

[61] P. De *et al.*, "Role of porosity and diffusion coefficient in porous electrode used in supercapacitors – Correlating theoretical and experimental studies," *Electrochemical Science Advances*, vol. 3, no. 1, 2023, doi: 10.1002/elsa.202100159.

[62] Mohit, N. Yadav, and S. A. Hashmi, “High energy density solid-state supercapacitors based on porous carbon electrodes derived from pre-treated bio-waste precursor sugarcane bagasse,” *J. Energy Storage*, vol. 55, Nov. 2022, doi: 10.1016/j.est.2022.105421.

[63] Y. Liu, Z. Shi, Y. Gao, W. An, Z. Cao, and J. Liu, “Biomass-Swelling Assisted Synthesis of Hierarchical Porous Carbon Fibers for Supercapacitor Electrodes,” *ACS Appl. Mater. Interfaces*, vol. 8, no. 42, pp. 28283–28290, Oct. 2016, doi: 10.1021/acsami.5b11558.

[64] A. Dubey *et al.*, “An eco-friendly, low-power charge storage device from bio-tolerable nano cerium oxide electrodes for bioelectrical and biomedical applications,” *Biomed. Phys. Eng. Express*, vol. 4, no. 2, 2018, doi: 10.1088/2057-1976/aaa282.

[65] S. Unknown, P. Chand, and A. Joshi, “Biomass derived carbon for supercapacitor applications: Review,” 2021. doi: 10.1016/j.est.2021.102646.

[66] D. Puthusseri, V. Aravindan, S. Madhavi, and S. Ogale, “3D micro-porous conducting carbon beehive by single step polymer carbonization for high performance supercapacitors: the magic of in situ porogen formation,” *Energy Environ. Sci.*, vol. 7, no. 2, pp. 728–735, 2014, doi: 10.1039/C3EE42551G.

[67] A. Jain, C. Xu, S. Jayaraman, R. Balasubramanian, J. Y. Lee, and M. P. Srinivasan, “Mesoporous activated carbons with enhanced porosity by optimal hydrothermal pre-treatment of biomass for supercapacitor applications,” *Microporous and Mesoporous Materials*, vol. 218, pp. 55–61, Dec. 2015, doi: 10.1016/J.MICROMESO.2015.06.041.

[68] S. Guo *et al.*, “Effects of CO2 activation on porous structures of coconut shell-based activated carbons,” *Appl. Surf. Sci.*, vol. 255, no. 20, pp. 8443–8449, Jul. 2009, doi: 10.1016/J.APSUSC.2009.05.150.

[69] C.-F. Chang, C.-Y. Chang, and W.-T. Tsai, “Effects of Burn-off and Activation Temperature on Preparation of Activated Carbon from Corn Cob Agrowaste by CO2 and Steam,” *J. Colloid Interface Sci.*, vol. 232, no. 1, pp. 45–49, 2000, doi: https://doi.org/10.1006/jcis.2000.7171.

[70] B. Li *et al.*, “Nitrogen-doped activated carbon for a high energy hybrid supercapacitor,” *Energy Environ. Sci.*, vol. 9, no. 1, pp. 102–106, 2016, doi: 10.1039/C5EE03149D.

[71] Y. Fu, Y. Shen, Z. Zhang, X. Ge, and M. Chen, “Activated bio-chars derived from rice husk via one- and two-step KOH-catalyzed pyrolysis for phenol adsorption,” *Science of The Total Environment*, vol. 646, pp. 1567–1577, Jan. 2019, doi: 10.1016/J.SCITOTENV.2018.07.423.

[72] Q. Sun, T. Jiang, G. Zhao, and J. Shi, “Porous carbon material based on biomass prepared by MgO template method and ZnCl2 activation method as electrode for high performance supercapacitor,” *Int. J. Electrochem. Sci.*, vol. 14, no. 1, 2019, doi: 10.20964/2019.01.50.

[73] J. Hou, C. Cao, F. Idrees, and X. Ma, “Hierarchical porous nitrogen-doped carbon nanosheets derived from silk for ultrahigh-capacity battery anodes and supercapacitors,” *ACS Nano*, vol. 9, no. 3, pp. 2556–2564, Mar. 2015, doi: 10.1021/nn506394r.

[74] P. Lin, Y. Xia, and Z. Liu, “Influence of Different Activators on the Structure and Properties of Activated Carbon Based on Bamboo Fiber,” *Polymers (Basel).*, vol. 14, no. 24, Dec. 2022, doi: 10.3390/polym14245500.

[75] P. T. Williams and A. R. Reed, “Development of activated carbon pore structure via physical and chemical activation of biomass fibre waste,” *Biomass Bioenergy*, vol. 30, no. 2, pp. 144–152, Feb. 2006, doi: 10.1016/j.biombioe.2005.11.006.

[76] H. Yamada, H. Nakamura, F. Nakahara, I. Moriguchi, and T. Kudo, “Electrochemical Study of High Electrochemical Double Layer Capacitance of Ordered Porous Carbons with Both Meso/Macropores and Micropores,” *The Journal of Physical Chemistry C*, vol. 111, no. 1, pp. 227–233, 2007, doi: 10.1021/jp063902g.

[77] P. Yu *et al.*, “Rational Synthesis of Highly Porous Carbon from Waste Bagasse for Advanced Supercapacitor Application,” *ACS Sustain. Chem. Eng.*, vol. 6, no. 11, pp. 15325–15332, Nov. 2018, doi: 10.1021/acssuschemeng.8b03763.

[78] A. Gopalakrishnan and S. Badhulika, “Effect of self-doped heteroatoms on the performance of biomass-derived carbon for supercapacitor applications,” *Elsevier B.V.*, Dec. 31, 2020, doi: 10.1016/j.jpowsour.2020.228830.

[79] Z. Mao, S. Zhao, J. Wang, Y. Zeng, X. Lu, and Y. Tong, “Facile synthesis of nitrogen-doped porous carbon as robust electrode for supercapacitors,” *Mater. Res. Bull.*, vol. 101, pp. 140–145, May 2018, doi: 10.1016/J.MATERRESBULL.2018.01.028.

[80] M. J. Mostazo-López *et al.*, “Ultraporous nitrogen-doped zeolite-templated carbon for high power density aqueous-based supercapacitors,” *Carbon N. Y.*, vol. 129, pp. 510–519, Apr. 2018, doi: 10.1016/J.CARBON.2017.12.050.

[81] Y. Huang *et al.*, “Sulfurized activated carbon for high energy density supercapacitors,” *J. Power Sources*, vol. 252, pp. 90–97, Apr. 2014, doi: 10.1016/j.jpowsour.2013.12.004.

[82] H. Chen, F. Yu, G. Wang, L. Chen, B. Dai, and S. Peng, “Nitrogen and Sulfur Self-Doped Activated Carbon Directly Derived from Elm Flower for High-Performance Supercapacitors,” *ACS Omega*, vol. 3, no. 4, pp. 4724–4732, Apr. 2018, doi: 10.1021/acsomega.8b00210.

[83] X. Du *et al.*, “Effect of ZnCl2 impregnation concentration on the microstructure and electrical performance of ramie-based activated carbon hollow fiber,” *Ionics (Kiel).*, vol. 22, no. 4, pp. 545–553, Apr. 2016, doi: 10.1007/s11581-015-1571-3.

[84] H. D. Tran, D. Li, and R. B. Kaner, “One-Dimensional Conducting Polymer Nanostructures: Bulk Synthesis and Applications,” *Advanced Materials*, vol. 21, no. 14–15, pp. 1487–1499, 2009, doi: https://doi.org/10.1002/adma.200802289.

[85] W. Chen, H. Yu, S.-Y. Lee, T. Wei, J. Li, and Z. Fan, “Nanocellulose: a promising nanomaterial for advanced electrochemical energy storage,” *Chem. Soc. Rev.*, vol. 47, no. 8, pp. 2837–2872, 2018, doi: 10.1039/C7CS00790F.

[86] R. J. Moon, A. Martini, J. Nairn, J. Simonsen, and J. Youngblood, “Cellulose nanomaterials review: Structure, properties and nanocomposites,” *Chem. Soc. Rev.*, vol. 40, no. 7, 2011, doi: 10.1039/c0cs00108b.

[87] H. Wang, H. Yi, C. Zhu, X. Wang, and H. Jin Fan, “Functionalized highly porous graphitic carbon fibers for high-rate supercapacitive electrodes,” *Nano Energy*, vol. 13, 2015, doi: 10.1016/j.nanoen.2015.03.033.

[88] J. Yan, Q. Wang, T. Wei, and Z. Fan, “Recent advances in design and fabrication of electrochemical supercapacitors with high energy densities,” *Adv. Energy Mater.*, vol. 4, no. 4, 2014, doi: 10.1002/aenm.201300816.

[89] L. Peng, Z. Fang, Y. Zhu, C. Yan, and G. Yu, “Holey 2D Nanomaterials for Electrochemical Energy Storage,” 2018. doi: 10.1002/aenm.201702179.

[90] C. Chen *et al.*, “Facile synthesis of crumpled nitrogen-doped porous carbon nanosheets with ultrahigh surface area for high-performance supercapacitors,” *Nanoscale Adv.*, vol. 5, no. 7, 2023, doi: 10.1039/d2na00949h.

[91] Z. Bi *et al.*, “Biomass-derived porous carbon materials with different dimensions for supercapacitor electrodes: A review,” 2019, *Royal Society of Chemistry*. doi: 10.1039/c9ta04436a.

[92] S. Song, F. Ma, G. Wu, D. Ma, W. Geng, and J. Wan, “Facile self-templating large scale preparation of biomass-derived 3D hierarchical porous carbon for advanced supercapacitors,” *J. Mater. Chem. A Mater.*, vol. 3, no. 35, pp. 18154–18162, Jul. 2015, doi: 10.1039/c5ta04721h.

[93] B. Pant, G. P. Ojha, J. Acharya, and M. Park, “Preparation, characterization, and electrochemical performances of activated carbon derived from the flower of Bauhinia variegata L for supercapacitor applications,” *Diam. Relat. Mater.*, vol. 136, Jun. 2023, doi: 10.1016/j.diamond.2023.110040.

[94] E. A. Arkhipova, R. Y. Novotortsev, A. S. Ivanov, K. I. Maslakov, and S. V. Savilov, “Rice husk-derived activated carbon electrode in redox-active electrolyte – New approach for enhancing supercapacitor performance,” *J. Energy Storage*, vol. 55, Nov. 2022, doi: 10.1016/j.est.2022.105699.

[95] M. D. Mehare, A. D. Deshmukh, and S. J. Dhoble, “Preparation of porous agro-waste-derived carbon from onion peel for supercapacitor application,” *J. Mater. Sci.*, vol. 55, no. 10, 2020, doi: 10.1007/s10853-019-04236-7.

[96] S. S. Shah, E. Cevik, M. A. Aziz, T. F. Qahtan, A. Bozkurt, and Z. H. Yamani, “Jute Sticks Derived and Commercially Available Activated Carbons for Symmetric Supercapacitors with Bio-electrolyte: A Comparative Study,” *Synth. Met.*, vol. 277, 2021, doi: 10.1016/j.synthmet.2021.116765.

[97] C. Zequine *et al.*, “High-Performance Flexible Supercapacitors obtained via Recycled Jute: Bio-Waste to Energy Storage Approach,” *Sci. Rep.*, vol. 7, no. 1, 2017, doi: 10.1038/s41598-017-01319-w.

[98] P. Jakubec, S. Bartusek, J. J. Dvořáček, V. Šedajová, V. Kupka, and M. Otyepka, “Flax-derived carbon: A highly durable electrode material for electrochemical double-layer supercapacitors,” *Nanomaterials*, vol. 11, no. 9, 2021, doi: 10.3390/nano11092229.

[99] Y. Li *et al.*, “Biomass-derived microporous carbon with large micropore size for high-performance supercapacitors,” *J. Power Sources*, vol. 448, 2020, doi: 10.1016/j.jpowsour.2019.227396.

[100] L. K. Shrestha *et al.*, “Phyllanthus emblica Seed-Derived Hierarchically Porous Carbon Materials for High-Performance Supercapacitor Applications,” *Materials*, vol. 15, no. 23, 2022, doi: 10.3390/ma15238335.

[101] L. Huang, S. Wang, Y. Zhang, X. H. Huang, J. J. Peng, and F. Yang, “Preparation of a N-P co-doped waste cotton fabric-based activated carbon for supercapacitor electrodes,” *Xinxing Tan Cailiao/New Carbon Materials*, vol. 36, no. 6, 2021, doi: 10.1016/S1872-5805(21)60054-9.

[102] X. Liu, N. Ostrovsky-Snider, M. Lo Presti, T. Kim, G. Guidetti, and F. G. Omenetto, “Use of Silk Fibroin Material Composites for Green, Flexible Supercapacitors,” *ACS Biomater. Sci. Eng.*, vol. 10, no. 8, pp. 5390–5398, Aug. 2024, doi: 10.1021/acsbiomaterials.4c00716.

[103] H. Y. Park, M. Huang, T.-H. Yoon, and K. H. Song, “Electrochemical properties of kenaf-based activated carbon monolith for supercapacitor electrode applications,” *RSC Adv.*, vol. 11, no. 61, pp. 38515–38522, 2021, doi: 10.1039/D1RA07815A.

[104] J. Sun *et al.*, “Fabrication of three-dimensional microtubular kapok fiber carbon aerogel/RuO2 composites for supercapacitors,” *Electrochim. Acta*, vol. 300, pp. 225–234, Mar. 2019, doi: 10.1016/J.ELECTACTA.2019.01.095.

[105] L. Wang, X. Ma, Z. Ma, P. Li, and L. Zhang, “KHCO3 Chemical-Activated Hydrothermal Porous Carbon Derived from Sugarcane bagasse for Supercapacitor Applications,” *Chem. Asian J.*, Sep. 2024, doi: 10.1002/asia.202400530.

[106] K. K. R. Reddygunta, R. Beresford, L. Šiller, L. Berlouis, and A. Ivaturi, “Activated Carbon Utilization from Corn Derivatives for High-Energy-Density Flexible Supercapacitors,” *Energy and Fuels*, vol. 37, no. 23, pp. 19248–19265, Dec. 2023, doi: 10.1021/acs.energyfuels.3c01925.

[107] Q. Wang, Z. Chen, Q. Luo, H. Li, J. Li, and W. Yang, “Capillary Evaporation on High-Dense Conductive Ramie Carbon for Assisting Highly Volumetric-Performance Supercapacitors,” *Small*, vol. 19, no. 42, Oct. 2023, doi: 10.1002/smll.202303349.

[108] S. Sawadsitang, T. Duangchuen, A. Karaphun, T. Putjuso, P. Kumnorkaew, and E. Swatsitang, “Synthesis, characterization and electrochemical properties of activated coconut fiber carbon (ACFC) and CuO/ACFC nanocomposites for applying as electrodes of supercapacitor devices,” *Surfaces and Interfaces*, vol. 25, Aug. 2021, doi: 10.1016/j.surfin.2021.101174.

[109] S. S. Gunasekaran and S. Badhulika, “High-performance solid-state supercapacitor based on sustainable synthesis of meso-macro porous carbon derived from hemp fibres via CO2 activation,” *J. Energy Storage*, vol. 41, p. 102997, Sep. 2021, doi: 10.1016/J.EST.2021.102997.

[110] S. Iqbal, H. Khatoon, A. Hussain Pandit, and S. Ahmad, “Recent development of carbon-based materials for energy storage devices,” *Mater. Sci. Energy Technol.*, vol. 2, no. 3, pp. 417–428, Dec. 2019, doi: 10.1016/j.mset.2019.04.006.

[111] Q. Abbas, R. Raza, I. Shabbir, and A. G. Olabi, “Heteroatom doped high porosity carbon nanomaterials as electrodes for energy storage in electrochemical capacitors: A review,” 2019. doi: 10.1016/j.jsamd.2019.07.007.

[112] L. Zhang *et al.*, “Facile synthesis of bio-based nitrogen- and oxygen-doped porous carbon derived from cotton for supercapacitors,” *RSC Adv.*, vol. 8, no. 7, pp. 3869–3877, 2018, doi: 10.1039/c7ra11475c.

[113] V. V. N. Obreja, “On the performance of supercapacitors with electrodes based on carbon nanotubes and carbon activated material—A review,” *Physica E Low. Dimens. Syst. Nanostruct.*, vol. 40, no. 7, pp. 2596–2605, May 2008, doi: 10.1016/J.PHYSE.2007.09.044.

[114] T. Yumak, D. Bragg, and E. M. Sabolsky, “Effect of synthesis methods on the surface and electrochemical characteristics of metal oxide/activated carbon composites for supercapacitor applications,” *Appl. Surf. Sci.*, vol. 469, pp. 983–993, Mar. 2019, doi: 10.1016/J.APSUSC.2018.09.079.

[115] A. Awitdrus *et al.*, “Biomass conversion into activated carbon as a sustainable energy material for the development of supercapacitor devices,” *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 44, no. 2, pp. 3349–3359, Jun. 2022, doi: 10.1080/15567036.2022.2064941.

[116] A. Dubey *et al.*, “Biocharring of natural fibers of insect and plant origin: a green route for the production of ‘carbon-based charge storage nanomaterials,’” *Mater. Renew. Sustain. Energy*, vol. 7, no. 4, p. 20, 2018, doi: 10.1007/s40243-018-0127-7.

[117] M. Sevilla and R. Mokaya, “Energy storage applications of activated carbons: supercapacitors and hydrogen storage,” *Energy Environ. Sci.*, vol. 7, no. 4, pp. 1250–1280, 2014, doi: 10.1039/C3EE43525C.

[118] Y. L. Zhang and Z. S. Tang, “Porous carbon derived from herbal plant waste for supercapacitor electrodes with ultrahigh specific capacitance and excellent energy density,” *Waste Management*, vol. 106, pp. 250–260, Apr. 2020, doi: 10.1016/J.WASMAN.2020.03.032.

[119] O. Boujibar, F. Ghamouss, A. Ghosh, O. Achak, and T. Chafik, “Activated carbon with exceptionally high surface area and tailored nanoporosity obtained from natural anthracite and its use in supercapacitors,” *J. Power Sources*, vol. 436, p. 226882, Oct. 2019, doi: 10.1016/J.JPOWSOUR.2019.226882.

[120] H. Jangir *et al.*, “Sequential entrapping of Li and S in a conductivity cage of N-doped reduced graphene oxide supercapacitor derived from silk cocoon: a hybrid Li–S-silk supercapacitor,” *Appl. Nanosci.*, vol. 8, no. 3, pp. 379–393, 2018, doi: 10.1007/s13204-018-0641-z.

[121] S. Zhang and N. Pan, “Supercapacitors Performance Evaluation,” *Adv. Energy Mater.*, vol. 5, no. 6, p. 1401401, 2015, doi: https://doi.org/10.1002/aenm.201401401.

[122] Y. Zhou, “Advances of machine learning in multi-energy district communities– mechanisms, applications and perspectives,” *Energy and AI*, vol. 10, p. 100187, Nov. 2022, doi: 10.1016/J.EGYAI.2022.100187.

[123] K. Pan *et al.*, “Integrated data mining for prediction of specific capacitance of porous carbon materials for flexible energy storage devices,” *J. Energy Storage*, vol. 73, p. 109072, Dec. 2023, doi: 10.1016/J.EST.2023.109072.

[124] M. Pardakhti, E. Moharreri, D. Wanik, S. L. Suib, and R. Srivastava, “Machine Learning Using Combined Structural and Chemical Descriptors for Prediction of Methane Adsorption Performance of Metal Organic Frameworks (MOFs),” *ACS Comb. Sci.*, vol. 19, no. 10, pp. 640–645, Oct. 2017, doi: 10.1021/acscombsci.7b00056.

[125] S. Zhu *et al.*, “Artificial neural network enabled capacitance prediction for carbon-based supercapacitors,” *Mater. Lett.*, vol. 233, pp. 294–297, Dec. 2018, doi: 10.1016/j.matlet.2018.09.028.

[126] A. G. Saad, A. Emad-Eldeen, W. Z. Tawfik, and A. G. El-Deen, “Data-driven machine learning approach for predicting the capacitance of graphene-based supercapacitor electrodes,” *J. Energy Storage*, vol. 55, Nov. 2022, doi: 10.1016/j.est.2022.105411.

[127] X. Yang, C. Yuan, S. He, D. Jiang, B. Cao, and S. Wang, “Machine learning prediction of specific capacitance in biomass derived carbon materials: Effects of activation and biochar characteristics,” *Fuel*, vol. 331, Jan. 2023, doi: 10.1016/j.fuel.2022.125718.

[128] H. Su, S. Lin, S. Deng, C. Lian, Y. Shang, and H. Liu, “Predicting the capacitance of carbon-based electric double layer capacitors by machine learning,” *Nanoscale Adv.*, vol. 1, no. 6, pp. 2162–2166, 2019, doi: 10.1039/C9NA00105K.

[129] S. Al-Dahidi, O. Ayadi, M. Alrbai, and J. Adeeb, “Ensemble Approach of Optimized Artificial Neural Networks for Solar Photovoltaic Power Prediction,” *IEEE Access*, vol. 7, pp. 81741–81758, 2019, doi: 10.1109/ACCESS.2019.2923905.

[130] Y. Zhou, S. Li, and K. Wang, “Life Prediction of Hybrid Supercapacitor Based on Improved Model-Extreme Learning Machine,” in *2019 IEEE 10th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, 2019, pp. 420–424. doi: 10.1109/PEDG.2019.8807678.

[131] J. R. Zhao, J. Hu, J. F. Li, and P. Chen, “N-doped carbon nanotubes derived from waste biomass and its electrochemical performance,” *Mater. Lett.*, vol. 261, Feb. 2020, doi: 10.1016/j.matlet.2019.127146.

[132] A. Emad-Eldeen, M. A. Azim, M. Abdelsattar, and A. AbdelMoety, “Utilizing machine learning and deep learning for enhanced supercapacitor performance prediction,” *J. Energy Storage*, vol. 100, Oct. 2024, doi: 10.1016/j.est.2024.113556.

[133] T. Gao and W. Lu, “Machine learning toward advanced energy storage devices and systems,” *iScience*, vol. 24, no. 1, p. 101936, Jan. 2021, doi: 10.1016/J.ISCI.2020.101936.

[134] S. Oyucu, F. Doğan, A. Aksöz, and E. Biçer, “Comparative Analysis of Commonly Used Machine Learning Approaches for Li-Ion Battery Performance Prediction and Management in Electric Vehicles,” *Applied Sciences*, vol. 14, no. 6, 2024, doi: 10.3390/app14062306.

[135] H. Liu, Z. Cui, Z. Qiao, X. An, and Y. Wang, “Machine learning-assisted prediction, screen, and interpretation of porous carbon materials for high-performance supercapacitors,” *Journal of Materials Informatics*, vol. 4, no. 4, 2024, doi: 10.20517/jmi.2024.29.

[136] P. Simon and Y. Gogotsi, “Materials for electrochemical capacitors,” *Nat. Mater.*, vol. 7, no. 11, pp. 845–854, 2008, doi: 10.1038/nmat2297.

[137] J. Chmiola, G. Yushin, Y. Gogotsi, C. Portet, P. Simon, and P. L. Taberna, “Anomalous Increase in Carbon Capacitance at Pore Sizes Less Than 1 Nanometer,” *Science (1979).*, vol. 313, no. 5794, pp. 1760–1763, Sep. 2006, doi: 10.1126/science.1132195.

[138] S. Zhu *et al.*, “Soluble salt self-assembly-assisted synthesis of three-dimensional hierarchical porous carbon networks for supercapacitors,” *J. Mater. Chem. A*, vol. 3, no. 44, pp. 22266–22273, 2015, doi: 10.1039/C5TA04646G.

[139] X. J. Li *et al.*, “Excellent Capacitive Performance of a Three-Dimensional Hierarchical Porous Graphene/Carbon Composite with a Superhigh Surface Area,” *Chemistry – A European Journal*, vol. 20, no. 41, pp. 13314–13320, 2014, doi: https://doi.org/10.1002/chem.201402897.

[140] M. Karnan, A. G. K. Raj, K. Subramani, S. Santhoshkumar, and M. Sathish, "The fascinating supercapacitive performance of activated carbon electrodes with enhanced energy density in multifarious electrolytes," *Sustainable Energy Fuels*, vol. 4, no. 6, pp. 3029–3041, 2020, doi: 10.1039/C9SE01298B.

[141] S. Chen, W. Wang, X. Zhang, and X. Wang, "High-Performance Supercapacitors Based on Graphene/Activated Carbon Hybrid Electrodes Prepared via Dry Processing," *Batteries*, vol. 10, no. 6, 2024, doi: 10.3390/batteries10060195.

[142] A. Nordenström *et al.*, "High-surface-area activated carbon from pine cones for semi-industrial spray deposition of supercapacitor electrodes," *Nanoscale Adv.*, vol. 4, no. 21, pp. 4689–4700, 2022, doi: 10.1039/D2NA00362G.

[143] R. Heimböckel, F. Hoffmann, and M. Fröba, "Insights into the influence of the pore size and surface area of activated carbons on the energy storage of electric double layer capacitors with a new potentially universally applicable capacitor model," *Phys. Chem. Chem. Phys.*, vol. 21, no. 6, pp. 3122–3133, 2019, doi: 10.1039/C8CP06443A.

[144] S. Liu *et al.*, "Nitrogen-Doped Porous Carbons Derived from Peanut Shells as Efficient Electrodes for High-Performance Supercapacitors," *Int. J. Mol. Sci.*, vol. 25, no. 14, 2024, doi: 10.3390/ijms25147583.

[145] A. Chen, Y. Yu, T. Xing, R. Wang, Y. Li, and Y. Li, "Synthesis of nitrogen-doped hierarchical porous carbon for supercapacitors," *Mater. Lett.*, vol. 157, pp. 30–33, Oct. 2015, doi: 10.1016/J.MATLET.2015.05.082.

[146] B. Blaiszik *et al.*, “A data ecosystem to support machine learning in materials science,” *MRS Commun.*, vol. 9, no. 4, pp. 1125–1133, 2019, doi: 10.1557/mrc.2019.118.

[147] A. Ibrahem Ahmed Osman, A. Najah Ahmed, M. F. Chow, Y. Feng Huang, and A. El-Shafie, “Extreme Gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia,” *Ain Shams Engineering Journal*, vol. 12, no. 2, pp. 1545–1556, Jun. 2021, doi: 10.1016/J.ASEJ.2020.11.011.

[148] A. Ibrahem Ahmed Osman, A. Najah Ahmed, M. F. Chow, Y. Feng Huang, and A. El-Shafie, “Extreme Gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia,” *Ain Shams Engineering Journal*, vol. 12, no. 2, pp. 1545–1556, Jun. 2021, doi: 10.1016/J.ASEJ.2020.11.011.

[149] S. Rameshpant Kalmegh Associate Professor, “International Journal of Emerging Technology and Advanced Engineering Comparative Analysis of WEKA Data Mining Algorithm RandomForest, RandomTree and LADTree for Classification of Indigenous News Data,” 2008. [Online]. Available: www.ijetae.com

[150] G. Trezza, L. Bergamasco, M. Fasano, and E. Chiavazzo, “Minimal crystallographic descriptors of sorption properties in hypothetical MOFs and role in sequential learning optimization,” *NPJ Comput. Mater.*, vol. 8, no. 1, p. 123, 2022, doi: 10.1038/s41524-022-00806-7.

[151] M. Zhou, A. Gallegos, K. Liu, S. Dai, and J. Wu, “Insights from machine learning of carbon electrodes for electric double layer capacitors,” *Carbon N. Y.*, vol. 157, pp. 147–152, Feb. 2020, doi: 10.1016/j.carbon.2019.08.090.

[152] S. Wang, K. Huang, C. Lian, and H. Liu, “Machine learning and data-driven material exploration for supercapacitors,” *Supercapacitors: Materials, Design, and Commercialization*, pp. 353–370, Jan. 2024, doi: 10.1016/B978-0-443-15478-2.00015-2.

[153] P. Liu *et al.*, “An emerging machine learning strategy for the assisted design of high-performance supercapacitor materials by mining the relationship between capacitance and structural features of porous carbon,” *Journal of Electroanalytical Chemistry*, vol. 899, p. 115684, Oct. 2021, doi: 10.1016/J.JELECHEM.2021.115684.

[154] W. A. M. K. P. Wickramaarachchi, M. Minakshi, X. Gao, R. Dabare, and K. W. Wong, “Hierarchical porous carbon from mango seed husk for electro-chemical energy storage,” *Chemical Engineering Journal Advances*, vol. 8, p. 100158, Nov. 2021, doi: 10.1016/J.CEJA.2021.100158.

[155] Z.-H. Shen, H.-X. Liu, Y. Shen, J.-M. Hu, L.-Q. Chen, and C.-W. Nan, “Machine learning in energy storage materials,” *Interdisciplinary Materials*, vol. 1, no. 2, pp. 175–195, 2022, doi: https://doi.org/10.1002/idm2.12020.

[156] S. Ghosh, G. R. Rao, and T. Thomas, “Machine learning-based prediction of supercapacitor performance for a novel electrode material: Cerium oxynitride,” *Energy Storage Mater.*, vol. 40, pp. 426–438, Sep. 2021, doi: 10.1016/j.ensm.2021.05.024.

[157] S. Mishra *et al.*, “The impact of physicochemical features of carbon electrodes on the capacitive performance of supercapacitors: a machine learning approach,” *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-33524-1.

[158] W. Wang, Y. Shen, Z. Ma, X. Wei, H. Fan, and Q. Bai, “High-performance supercapacitors based on self-supporting thick carbon electrodes from renewable biomass wood,” *Sustainable Materials and Technologies*, vol. 40, p. e00824, Jul. 2024, doi: 10.1016/J.SUSMAT.2024.E00824.

[159] S. Yang, S. Wang, X. Liu, and L. Li, “Biomass derived interconnected hierarchical micro-meso-macro- porous carbon with ultrahigh capacitance for supercapacitors,” *Carbon N. Y.*, vol. 147, pp. 540–549, Jun. 2019, doi: 10.1016/J.CARBON.2019.03.023.

[160] M. D. Stoller and R. S. Ruoff, “Best practice methods for determining an electrode material’s performance for ultracapacitors,” *Energy Environ. Sci.*, vol. 3, no. 9, pp. 1294–1301, 2010, doi: 10.1039/C0EE00074D.

[161] S. Chen, L. Qiu, and H. M. Cheng, “Carbon-Based Fibers for Advanced Electrochemical Energy Storage Devices,” *Chem. Rev.*, vol. 120, no. 5, pp. 2811–2878, Mar. 2020, doi: 10.1021/acs.chemrev.9b00466.

[162] H. Wang, W. Cai, L. He, M. Zhu, and Y. Wang, “Anchoring ternary NiCoMn-S ultrathin nanosheets on porous ZnCo2O4 nanowires to form core-shell composites for high-performance asymmetric supercapacitor,” *J. Alloys Compd.*, vol. 870, p. 159347, Jul. 2021, doi: 10.1016/J.JALLCOM.2021.159347.

[163] Y. Liu *et al.*, “Advanced Supercapacitors Based on Porous Hollow Carbon Nanofiber Electrodes with High Specific Capacitance and Large Energy Density,” *ACS Appl. Mater. Interfaces*, vol. 12, no. 4, pp. 4777–4786, Jan. 2020, doi: 10.1021/acsami.9b19977.

[164] S. and K. K. K. Sinha Prerna and Banerjee, “Activated Carbon as Electrode Materials for Supercapacitors,” in *Handbook of Nanocomposite*

*Supercapacitor Materials II: Performance*, K. K. Kar, Ed., Cham: Springer International Publishing, pp. 113–144, 2020, doi: 10.1007/978-3-030-52359-6_5.

[165] X. Wang *et al.*, "Selective Charging Behavior in an Ionic Mixture Electrolyte-Supercapacitor System for Higher Energy and Power," *J. Am. Chem. Soc.*, vol. 139, no. 51, pp. 18681–18687, Dec. 2017, doi: 10.1021/jacs.7b10693.

[166] P. E. Lokhande, U. S. Chavan, and A. Pandey, "Materials and Fabrication Methods for Electrochemical Supercapacitors: Overview," *Electrochemical Energy Reviews*, vol. 3, no. 1, pp. 155–186, 2020, doi: 10.1007/s41918-019-00057-z.

[167] P. Simon and Y. Gogotsi, "Perspectives for electrochemical capacitors and related devices," *Nat. Mater.*, vol. 19, no. 11, pp. 1151–1163, 2020, doi: 10.1038/s41563-020-0747-z.

[168] A. Dubey *et al.*, "Biocharring of natural fibers of insect and plant origin: a green route for the production of 'carbon-based charge storage nanomaterials,'" *Mater. Renew. Sustain. Energy*, vol. 7, no. 4, Nov. 2018, doi: 10.1007/s40243-018-0127-7.

[169] S. Ghosh, S. Barg, S. M. Jeong, and K. (Ken) Ostrikov, "Heteroatom-Doped and Oxygen-Functionalized Nanocarbons for High-Performance Supercapacitors," *Adv. Energy Mater.*, vol. 10, no. 32, p. 2001239, 2020, doi: https://doi.org/10.1002/aenm.202001239.

[170] N. Kumar, S.-B. Kim, S.-Y. Lee, and S.-J. Park, "Recent Advanced Supercapacitor: A Review of Storage Mechanisms, Electrode Materials,

Modification, and Perspectives," *Nanomaterials*, vol. 12, no. 20, 2022, doi: 10.3390/nano12203708.

[171] E. T. Mombeshora and V. O. Nyamori, "A review on the use of carbon nanostructured materials in electrochemical capacitors," *John Wiley and Sons Ltd*., Dec. 01, 2015, doi: 10.1002/er.3423.

[172] S. Shrivas and A. Dubey, "Prediction of specific capacitance of activated carbon electrode for energy storage device by machine learning based approach," *J. Energy Storage*, vol. 121, p. 116632, Jun. 2025, doi: 10.1016/J.EST.2025.116632.

[173] A. B. Mendhe, B. Panigrahy, and H. S. Panda, "Ensemble machine learning for predictive accuracy and experimental corroboration on transition metal-based electrodes for supercapacitor," *Electrochim. Acta*, vol. 512, Feb. 2025, doi: 10.1016/j.electacta.2024.145482.

[174] M. Tang, Y. Ding, T. Hu, X. Zhu, G. Zheng, and Y. Tian, " Machine Learning-Based Prediction of Supercapacitor Capacitance for MgCo 2 O 4 Electrodes ," *ChemPhysChem*, Nov. 2024, doi: 10.1002/cphc.202400629.

[175] X. Lu, C. Zhao, H. Tu, S. Wang, A. Chen, and H. Zhang, "Research on prediction of energy density and power density of biomass carbon-based supercapacitors based on machine learning," *Sustainable Materials and Technologies*, vol. 44, p. e01309, Jul. 2025, doi: 10.1016/j.susmat.2025.e01309.

[176] K. A. Severson *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nat. Energy*, vol. 4, no. 5, pp. 383–391, May 2019, doi: 10.1038/s41560-019-0356-8.

[177] B. S. Reddy *et al.*, “Modeling capacitance of carbon-based supercapacitors by artificial neural networks,” *J. Energy Storage*, vol. 72, Nov. 2023, doi: 10.1016/j.est.2023.108537.

[178] Y. Zhang, Z. Peng, Y. Guan, and L. Wu, “Prognostics of battery cycle life in the early-cycle stage based on hybrid model,” *Energy*, vol. 221, p. 119901, Apr. 2021, doi: 10.1016/J.ENERGY.2021.119901.

[179] M. Rahimi, M. H. Abbaspour-Fard, and A. Rohani, “Synergetic effect of N/O functional groups and microstructures of activated carbon on supercapacitor performance by machine learning,” *J. Power Sources*, vol. 521, p. 230968, Feb. 2022, doi: 10.1016/J.JPOWSOUR.2021.230968.

[180] Y. Chen *et al.*, “Machine learning-guided prediction of energy storage performance of carbon cathode materials for zinc-ion hybrid capacitors,” *J. Colloid Interface Sci.*, vol. 699, p. 138139, 2025, doi: https://doi.org/10.1016/j.jcis.2025.138139.

[181] X. Lu, C. Zhao, H. Tu, S. Wang, A. Chen, and H. Zhang, “Research on prediction of energy density and power density of biomass carbon-based supercapacitors based on machine learning,” *Sustainable Materials and Technologies*, vol. 44, Jul. 2025, doi: 10.1016/j.susmat.2025.e01309.

[182] A. Emad-Eldeen, M. A. Azim, M. Abdelsattar, and A. AbdelMoety, “Utilizing machine learning and deep learning for enhanced supercapacitor performance prediction,” *J. Energy Storage*, vol. 100, Oct. 2024, doi: 10.1016/j.est.2024.113556.

[183] M. Tang, Y. Ding, T. Hu, X. Zhu, G. Zheng, and Y. Tian, “Machine Learning-Based Prediction of Supercapacitor Capacitance for MgCo2O4 Electrodes,”

*ChemPhysChem*, vol. 25, no. 21, p. e202400629, Nov. 2024, doi: https://doi.org/10.1002/cphc.202400629.

[184] W. Z. Tawfik, M. S. Abdel-Wahab, J. K. Lee, A. M. Al-Enizi, and R. Y. Youssef, "Modeling specific capacitance of carbon nanotube-based supercapacitor electrodes by machine learning algorithms," *Phys. Scr.*, vol. 99, no. 6, Jun. 2024, doi: 10.1088/1402-4896/ad4df3.

[185] A. Almuamri, F. Soysal, and S. Kocaoglu, "Capacitance Classification for Supercapacitors using Machine Learning," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3553187.

[186] Y. Ao, H. Li, L. Zhu, S. Ali, and Z. Yang, "The linear Random forest algorithm and its advantages in machine learning assisted logging regression modeling," *J. Pet. Sci. Eng.*, vol. 174, pp. 776–789, Mar. 2019, doi: 10.1016/J.PETROL.2018.11.067.

[187] W. LI *et al.*, "Implementation of AdaBoost and genetic algorithm machine learning models in prediction of adsorption capacity of nanocomposite materials," *J. Mol. Liq.*, vol. 350, p. 118527, Mar. 2022, doi: 10.1016/J.MOLLIQ.2022.118527.

[188] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient boosting Feature Selection With Machine Learning Classifiers for Intrusion Detection on Power Grids," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1104–1116, 2021, doi: 10.1109/TNSM.2020.3032618.

[189] M. Shariq *et al.*, "Machine learning models for prediction of electrochemical properties in supercapacitor electrodes using MXene and graphene nanoplatelets," *Chemical Engineering Journal*, vol. 484, Mar. 2024, doi: 10.1016/j.cej.2024.149502.

[190] A. Rácz, D. Bajusz, and K. Héberger, “Multi-Level Comparison of Machine Learning Classifiers and Their Performance Metrics,” *Molecules*, vol. 24, no. 15, 2019, doi: 10.3390/molecules24152811.

[191] I. Jebli, F. Z. Belouadha, M. I. Kabbaj, and A. Tilioua, “Prediction of solar energy guided by pearson correlation using machine learning,” *Energy*, vol. 224, p. 120109, Jun. 2021, doi: 10.1016/J.ENERGY.2021.120109.

[192] S. Nanda, S. Ghosh, and T. Thomas, “Machine learning aided cyclic stability prediction for supercapacitors,” *J. Power Sources*, vol. 546, Oct. 2022, doi: 10.1016/j.jpowsour.2022.231975.

[193] C. Zhao *et al.*, “Research on specific capacitance prediction of biomass carbon-based supercapacitors based on machine learning,” *J. Energy Storage*, vol. 97, Sep. 2024, doi: 10.1016/j.est.2024.112974.

[194] E. Y. L. Teo *et al.*, “High surface area activated carbon from rice husk as a high performance supercapacitor electrode,” *Electrochim. Acta*, vol. 192, pp. 110–119, Feb. 2016, doi: 10.1016/J.ELECTACTA.2016.01.140.

[195] W. Z. Tawfik, M. Shaban, A. Raveendran, J. K. Lee, and A. M. Al-Enizi, “Insights into the specific capacitance of CNT-based supercapacitor electrodes using artificial intelligence,” *RSC Adv.*, vol. 15, no. 5, pp. 3155–3167, 2025, doi: 10.1039/D4RA05546B.

[196] T. Wang *et al.*, “Machine-learning-assisted material discovery of oxygen-rich highly porous carbon active materials for aqueous supercapacitors,” *Nat. Commun.*, vol. 14, no. 1, p. 4607, 2023, doi: 10.1038/s41467-023-40282-1.

[197] R. Dubey and V. Guruviah, “Machine Learning Enabled Performance Prediction of Biomass-Derived Electrodes for Asymmetric Supercapacitor,” in *Futuristic Communication and Network Technologies*, N. Subhashini,

Morris. A. G. Ezra, and S.-K. Liaw, Eds., Singapore: Springer Nature Singapore, 2023, pp. 453–460.

[198] R. Mahindar, S. Mukherjee, S. Dutta, T. Dutta, and R. Ghosh, “Particle Swarm-Based Approach for Storage Efficiency Optimization of Supercapacitors,” in *Computational Intelligence in Pattern Recognition*, A. K. Das, J. Nayak, B. Naik, S. K. Pati, and D. Pelusi, Eds., Singapore: Springer Singapore, 2020, pp. 445–461.

[199] M. A. Ebied, M. A. Azim, and A. Emad-Eldeen, “Artificial Intelligence-based Cyclic Voltammetry Behavior Model for Supercapacitance of Zinc oxide (ZnO) Nanocomposite,” in *2022 4th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*, 2022, pp. 1–5. doi: 10.1109/REEPE53907.2022.9731390.

[200] V. Sawant, R. Deshmukh, and C. Awati, “Machine learning techniques for prediction of capacitance and remaining useful life of supercapacitors: A comprehensive review,” *Journal of Energy Chemistry*, vol. 77, pp. 438–451, 2023, doi: https://doi.org/10.1016/j.jechem.2022.11.012.

[201] M. Zhou, A. Gallegos, K. Liu, S. Dai, and J. Wu, “Insights from Machine Learning of Carbon Electrodes for Electric Double Layer Capacitors,” 2019.

[202] P. N. K. Anantha, A. K. Pratihasta, N. Tanwar, R. K. Singh, R. K. Misra, and S. K. Singh, “Interpretable Hybrid CDTCN-GRU Model for Supercapacitor Life Prediction,” *IEEE Sens. J.*, vol. 25, no. 14, pp. 26970–26978, 2025, doi: 10.1109/JSEN.2025.3573331.

[203] M. Tang, Y. Ding, T. Hu, X. Zhu, G. Zheng, and Y. Tian, “Machine Learning-Based Prediction of Supercapacitor Capacitance for MgCo2O4 Electrodes,” *ChemPhysChem*, vol. 25, no. 21, Nov. 2024, doi: 10.1002/cphc.202400629.

[204] E. Frackowiak and F. Béguin, “Carbon materials for the electrochemical storage of energy in capacitors,” 2001. doi: 10.1016/S0008-6223(00)00183-4.

[205] S. Shrivas, A. Dubey, R. Pandey, A. S. M. Tripathi, and S. Shrivastava, “Activated carbon derived from natural fiber precursor for next generation electrode materials for supercapacitors,” Oct. 01, 2025, *Springer*. doi: 10.1007/s10854-025-15878-6.

[206] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, “Machine learning for molecular and materials science,” Jul. 26, 2018, *Nature Publishing Group*. doi: 10.1038/s41586-018-0337-2.

[207] M. Hajihosseinlou, A. Maghsoudi, and R. Ghezelbash, “A Novel Scheme for Mapping of MVT-Type Pb–Zn Prospectivity: LightGBM, a Highly Efficient Gradient boosting Decision tree Machine Learning Algorithm,” *Natural Resources Research*, vol. 32, no. 6, pp. 2417–2438, 2023, doi: 10.1007/s11053-023-10249-6.

[208] S. I. Ahmed, S. Radhakrishnan, B. B. Nair, and R. Thiruvengadathan, “Efficient hyperparameter-tuned machine learning approach for estimation of supercapacitor performance attributes,” *J. Phys. Commun.*, vol. 5, no. 11, 2021, doi: 10.1088/2399-6528/ac3574.

[209] G. Henri and N. Lu, “A Supervised Machine Learning Approach to Control Energy Storage Devices,” *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 5910–5919, Nov. 2019, doi: 10.1109/TSG.2019.2892586.

[210] L. L. Zhang and X. S. Zhao, “Carbon-based materials as supercapacitor electrodes,” *Chem. Soc. Rev.*, vol. 38, no. 9, pp. 2520–2531, 2009, doi: 10.1039/B813846J.

[211] A. G. Pandolfo and A. F. Hollenkamp, “Carbon properties and their role in supercapacitors,” *J. Power Sources*, vol. 157, no. 1, pp. 11–27, Jun. 2006, doi: 10.1016/J.JPOWSOUR.2006.02.065.

[212] J. Wang, X. Zhang, Z. Li, Y. Ma, and L. Ma, “Recent progress of biomass-derived carbon materials for supercapacitors,” 2020. doi: 10.1016/j.jpowsour.2020.227794.

[213] S. Singh, Y. Zhang, S. A. Hashmi, and F. Yang, “Supercapacitors with cotton shell-derived activated carbons and porous polymer electrolyte films,” *RSC Adv.*, vol. 15, no. 13, pp. 9787–9800, 2025, doi: 10.1039/D5RA00696A.

[214] X. Wang, C. Yao, F. Wang, and Z. Li, “Cellulose-Based Nanomaterials for Energy Applications,” *Small*, vol. 13, no. 42, p. 1702240, 2017, doi: https://doi.org/10.1002/smll.201702240.

[215] T. Subramaniam, M. A. M. Nainar, and N. A. Nordin, “A Review on Synthesis and Characterization of Activated Carbon from Natural Fibers for Supercapacitor Application,” Jan. 01, 2022, *Universiti Putra Malaysia Press*. doi: 10.47836/pjst.30.1.20.

[216] F. Zhang, X. Xiao, D. Gandla, Z. Liu, D. Q. Tan, and Y. Ein-Eli, “Bio-Derived Carbon with Tailored Hierarchical Pore Structures and Ultra-High Specific Surface Area for Superior and Advanced Supercapacitors,” *Nanomaterials*, vol. 12, no. 1, 2022, doi: 10.3390/nano12010027.

[217] G. K. Yogesh *et al.*, “A machine learning approach for estimating supercapacitor performance of graphene oxide nano-ring based electrode materials,” *Energy Adv.*, vol. 4, no. 1, pp. 119–139, 2025, doi: 10.1039/D4YA00577E.

[218] D. K. Vishwakarma *et al.*, “Evaluation of CatBoost Method for Predicting Weekly Pan Evaporation in Subtropical and Sub-Humid Regions,” *Pure Appl. Geophys.*, vol. 181, no. 2, pp. 719–747, 2024, doi: 10.1007/s00024-023-03426-4.

[219] “Introduction to Biomass-Derived Carbon Materials,” 2023.

[220] N. Dhiman, V. Sharma, and S. Ghosh, “Perspective on Biomass-Based Cotton-Derived Nanocarbon for Multifunctional Energy Storage and Harvesting Applications,” Apr. 25, 2023, *American Chemical Society*. doi: 10.1021/acsaelm.3c00022.

[221] B. Karamanova, M. Shipochka, M. Georgiev, T. Stankulov, A. Stoyanova, and R. Stoyanova, “Biomass-Derived Carbonaceous Materials to Achieve High-Energy-Density Supercapacitors,” *Front. Mater.*, vol. 8, Apr. 2021, doi: 10.3389/fmats.2021.654841.

[222] H. Xu, Y. Zhang, L. Wang, Y. Chen, and S. Gao, “Hierarchical porous biomass-derived carbon framework with ultrahigh surface area for outstanding capacitance supercapacitor,” *Renew. Energy*, vol. 179, pp. 1826–1835, Dec. 2021, doi: 10.1016/j.renene.2021.08.008.

[223] M. S. Javed *et al.*, “Synthesis of mesoporous defective graphene-nanosheets in a space-confined self-assembled nanoreactor: Highly efficient capacitive energy storage,” *Electrochim. Acta*, vol. 305, pp. 517–527, May 2019, doi: 10.1016/J.ELECTACTA.2019.03.072.

[224] P. Thomas, C. W. Lai, and M. R. Bin Johan, “Recent developments in biomass-derived carbon as a potential sustainable material for super-capacitor-based energy storage and environmental applications,” Jun. 01, 2019, *Elsevier B.V.* doi: 10.1016/j.jaap.2019.03.021.

# LIST OF PUBLICATIONS

## List of Journals Publications:

1. Stuti Shrivas & Amarish Dubey, "Prediction of specific capacitance of activated carbon electrode for energy storage device by machine learning based approach", *Journal of Energy Storage*, Volume 121,116632 (2025), https://doi.org/10.1016/j.est.2025.116632. (I.F. 10.7) (Published)

2. Stuti Shrivas, Amarish Dubey, Ritu Pandey, Atul SM Tripathi & Shivanshu Shrivastava, "Activated carbon derived from natural fiber precursor for next generation electrode materials for supercapacitors", *J Mater Sci: Mater Electron* 36, 1830 (2025). https://doi.org/10.1007/s10854-025-15878-6. (I.F. 3.2) (Published)

3. Ritu Pandey, Amarish Dubey, Ganesh P Mishra, Aakriti Ahuja, Neelu Kambo, Shubham Joshi, Mukesh Kumar Sinha, Ragini Dubey & Stuti Shrivas, "Cellulosic crystallinity and antimicrobial artistry of cactus cilia", *Physica Scripta*, Volume 99, No. 6, 065005 (2024), https://doi.org/10.1088/1402-4896/ad4016. (I.F. 2.6) (published)

4. Stuti Shrivas, Mouli, Gauri Maurya, Chaitanya Yadav & Amarish Dubey, "Quantitative Prediction of Supercapacitor Performance in Biomass-Derived Carbon Electrodes Using Machine Learning Regression Models" in *Journal of Energy Storage.* (I.F. 10.7) (under review).

5. Stuti Shrivas & Amarish Dubey, "Toward Sustainable Data-Driven Energy Devices: Machine Learning Regression Based Prediction of Supercapacitor Performance Using Activated Carbon Electrode" in *IEEE Transactions on Artificial Intelligence*. (under review).

6. Stuti Shrivas & Amarish Dubey, "Utilization of Machine Learning Techniques to Estimate the Specific Capacitance of Activated Carbon Electrode for Supercapacitors" in *Physica Scripta.* (I.F. 2.6) (under review)

## List of Conferences:

1. Stuti Shrivas, Amarish Dubey & Ritu Pandey, "Renewable Energy Storage from Biomass Derived Carbon: Comparative Study," *2024 IEEE 11th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, Lucknow, India, 2024, pp. 1-5, doi: 10.1109/UPCON62832.2024.10983730. (Published)
2. Jovanpreet Singh, Shivangi Srivastava, Stuti Shrivas, Ritu Pandey & Amarish Dubey, "Comparative study of Structures of Electrochromic

Device for Flexible Electrochromic Display," *2023 in IEEE International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT)*, Dehradun, India, 2023, pp. 290-295, doi: 10.1109/DICCT56244.2023.10110119. (Published)

3. Stuti Shrivas and Amarish Dubey, "Insight into the Prediction of Specific Capacitance of Activated Carbon-Based Electrode for Supercapacitor Using Machine Learning," *2025 IEEE 12th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, Varanasi, India, 2025, pp. 1-4, doi: 10.1109/UPCON66414.2025.11453725. (Published)

4. Stuti Shrivas, Amarish Dubey & Ritu Pandey, "Energy from Abundant Sustainable Natural Fibers", *International Conference on Sustainable Design Practices, November 25 - 27, 2023* at National Institute of Fashion Technology, Kangra, India.

5. Stuti Shrivas, Ritu Pandey & Amarish Dubey, "Carbon Nanomaterials from Natural Fiber for Energy Storage" *International e-conference on Textiles for Advanced Technologies Textile Tech November 2022*.

6. Jovanpreet Singh, Stuti Shrivas & Amarish Dubey; "Natural fibers: The Energy Sources for Electrical Vehicle", *Functional Textiles and Clothing Conference November 2022* at IIT Delhi, New Delhi, India.