Journal of Nano- and Electronic Physics
Vol. **15** No 3, 03011(4pp) (2023)

Журнал нано- та електронної фізики
Том **15** № 3, 03011(4cc) (2023)

Ensemble Approach for Capacitance Prediction of Heteroatom Doped Carbon Based Electrode Materials

Richa Dubey*, Velmathi Guruviah†, Ravi Prakash Dwivedi‡

Vellore Institute of Technology – VIT Chennai, Kelambakkam – Vandalur Rd, Rajan Nagar, Chennai, 600127 Tamil Nadu, India

(Received 18 May 2023; revised manuscript received 12 June 2023; published online 30 June 2023)

An ensemble approach-based machine learning modeling is used in the current study for unveiling the effect of various electrode parameters on the electrochemical performance of hetero-atom doped nanocarbons. This is achieved using three meta-classifiers in combination with traditional Multi-Layer Perceptron and Random Forest models. The three meta-classifiers used are namely (i) bagging, (ii) classification via regression (CVR) and (iii) multi class classifier (MCC). Amongst these three models, bagging and classification via regression provided greater accuracy in terms of correctly classified instances (%) and area under region of convergence values. The designed models are used to predict class of specific capacitance values. 94.5 % of the considered dataset is classified correctly proving a better accuracy of the designed models. Lowest root mean square value of 0.1787 was obtained for RF model. Compared to the models defined in the literature, the suggested models in this work provide best fit of the experiment and predicted values with highest accuracy and lowest error performance values. The lowest error value for RF and MLP models are 0.18 and 0.19 respectively.

Keywords: Carbon based electrode, Energy storage, Heteroatom-doped, Machine learning, Nitrogen doped, Supercapacitor.

DOI: 10.21272/jnep.15(3).03011 PACS numbers: 42.90. + p, 45.10.Hj, 66.30. – h

1. INTRODUCTION

The current era of electrification and global connectivity has attributed to the energy crisis aroused due to the imbalance between the world's energy supply and demand. This expanding gap urges the implementation of various renewable energy resources for maintaining uninterrupted energy supply. But, the intermittent nature of these renewable energy sources demands energy storage systems which ensure continuity and security in energy supply. The trending lithium-ion battery technology does not satisfy the need for high-power applications [1].

Electrochemical capacitors, preferable known as supercapacitors exhibiting some inimitable properties like high charge capacitance retention rate (i.e. > 89 %), high power density (~ 5 kW/kg), excellent cyclic stability (> half a million cycles) and rapid charge/discharge (in milliseconds) bridges the gap between the conventional electrostatic capacitors and electrochemical batteries [1]. However, the low energy density restricts their current application. Electrode material being an integral and essential part of SC-related research has been extensively explored. Hence, optimum combination of nanocarbons, dopants and electrolytes is essential for broadening the scope of application and materializing the wide application potential of electrochemical capacitors [2].

Data driven modeling has emerged as the method of choice for various applications [3-4]. In contrast to the theoretical modeling, data driven machine learning approach allows establishment of quantitative correlation between the different important parameters of nanocarbons based electrode materials and their performance [5-6]. While the research in the area of heteroatom effects on the properties of nanocarbons based electrode materials has been done intensively however, some aspects require in-depth discussion and exhaustive coverage. In this work, we deploy three meta-classifiers in combination with traditional MLP and RF models for unveiling the effect hetero-atom doped on the electrochemical performance of nanocarbons based on extensive literature driven data.

2. DATA COLLECTION AND MODEL EVALUATION

For this research, a data driven approach was carried out to predict specific capacity, energy density and power density of heteroatom doped nanocarbons. Datasets used in this work were structured by collecting the data's from the literature [6-8]. A total of 16 parameters were considered based on proximate component analysis (PCA). Initially all the datasets were categorized into four classes/ grades namely: Class A: < 120 F/g, Class B: 120 F/g-200 F/g, Class C: 200 F/g-230 F/g and Class D: 230 F/g-290 F/g. Table 1 presents the set of input and output features employed for this work.

Once the datasets were extracted, PCC was implemented for deriving the linear relation between the two considered factors and corresponding equation is given as in (1)

$$r = \frac{cov(A,B)}{\sigma_A \sigma_B} \quad (1)$$

where,

* richa.dubey@vit.ac.in
† velmathi.g@vit.ac.in
‡ raviprakash.dwivedi@vit.ac.in

 © 2023 Sumy State University

$cov(A, B)$ – Covariance between A and B
σ_A, σ_B – Standard deviation of A and B respectively
r – Pearson correlation coefficient

Table 1 – Statistics of different considered parameters

Parameter	Mini-mum	Maxi-mum	Mean	SD
ED	1.09	11.33	6.922	2.369
PD	0.06	11.7	1.404	2.081
MISA	160	2316	1093.5	473.82
MESA	16	764	362	262.869
Oxygen	4.94	17.19	9.457	3.51
Pyrollic	0.13	3.66	1.43	0.978
Pyridinic	0.18	3.32	1.3	0.885
Quaternary	0.12	6.58	1.06	1.024
Others	0.16	2.36	0.876	0.626
Capacitance	5.72	286.41	176.88	71.32

* SD – Standard Deviation

The success of the ML models relies on the appropriate selection of parameters. For materializing the full potential of developed supercapacitor, the proper combination of electrolyte, nanocarbons and dopants is essential.

Different ML models were trained and optimized for predicting the specific capacitance and energy density via WEKA (Waikato Environment for Knowledge Analysis) [9]. For implementation of any ML model the entire dataset was randomly distributed into 80:20 ratios for training and test set. Using trial-and-error method, the hyper parameters were tuned.

The performance of the developed models was assessed on the basis of error performance parameter i.e. RMSE and is expressed as:

$$RMSE = \sqrt{\frac{\left(\sum_{n=1}^{N} |x_n - y_n|^2\right)}{N}} \qquad (2)$$

3. RESULTS AND DISCUSSION

3.1 Algorithms Used

For the preparation of prediction model based on heteroatom doped nanocarbons mainly, two different algorithms are used: (i) random forest (RF) and (ii) multi layer perceptron (MLP) [10]. RF is an ensemble learning method which makes predictions based on multiple trees. It improves accuracy and avoids the overfitting problems. It is considered a fast classifier as its implementation in multicore processor enables concurrent use of processor core thereby reducing the time required to build the model.

3.2 Prediction Model

For studying the significance of precise control of oxygen functionality on graphene based network, we have considered tunable C/O ratio reported in literature [7].

In comparison with the pristine graphene, 49.2 % increase in capacitance value under optimized conditions was achieved as shown in Fig. 2 (A-B). As observed, RF provides a good fit of the predicted values with the experimental.

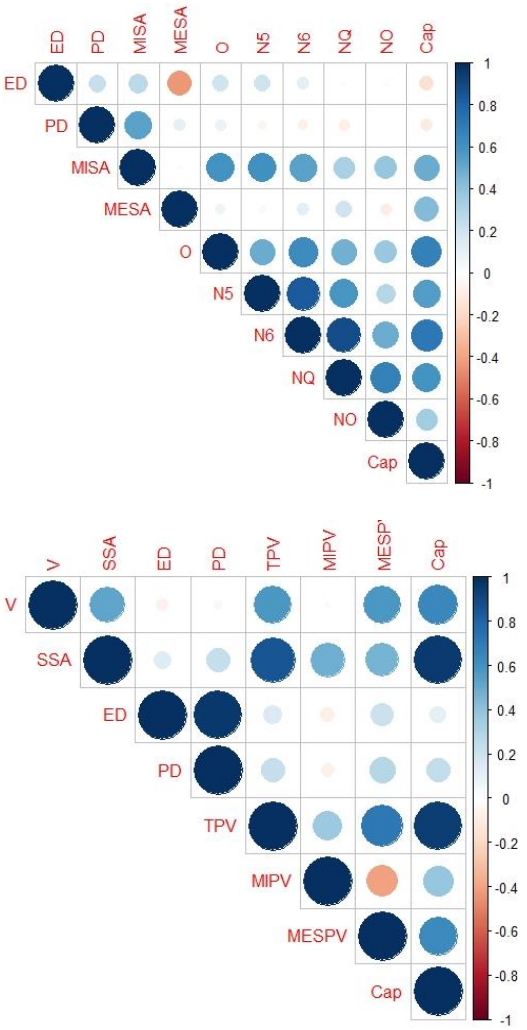
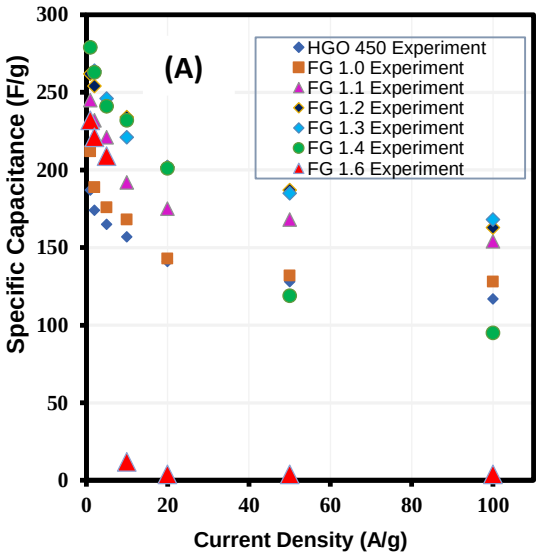


Fig. 1 – Pearson Correlation Coefficient matrix



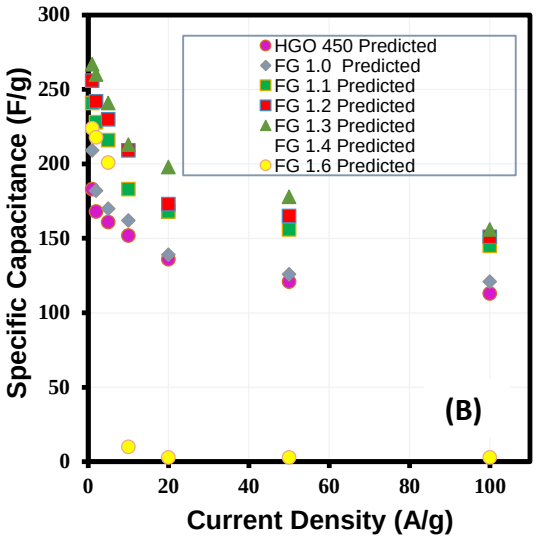


Fig. 2 – (A) Experiment values for Hydrothermally reduced graphene oxide (HGO) and Functionalized Graphene (FG) (B) ANN predicted results for the same. "FG potential" represents the apex potential at which the sample was cycling. Experiment datas were collected from ref. [7]

Table 2 – Sensitivity and specificity parameters of models

Grade	TPR		FPR		Precision		AUROC	
	RF	MLP	RF	MLP	RF	MLP	RF	MLP
A	0.8	0.8	0.1	0.1	0.8	0.7	0.9	0.8
B	0.7	0.6	0.0	0.1	0.6	0.5	0.9	0.8
C	0.3	0.3	0.1	0.1	0.5	0.4	0.7	0.6
D	0.5	0.6	0.0	0.0	0.5	0.7	0.7	0.8
WA	0.6	0.6	0.1	0.1	0.6	0.6	0.8	0.8

Table 3 – Statistical parameters of models

Model Name	Kappa	MAE	RMSE	CC instances (%)
RF (70%)	0.8414	0.1175	0.21	89.1667
MLP (70%)	0.8012	0.0776	0.24	86.6667
RF	0.914	0.0951	0.17	94.5
MLP	0.8518	0.0553	0.19	90.5
Bagging with RF	0.8753	0.098	0.18	92
CVR with RF	0.8828	0.1095	0.19	92.5
MCC with RF	0.832	0.1342	0.22	89.25
Bagging with MLP	0.8638	0.0782	0.18	91.25
CVR with MLP	0.8638	0.0782	0.18	91.25
MCC with MLP	0.8403	0.1143	0.21	89.75

1. CONCLUSION

An ensemble data driven approach in combination with two traditional ML models namely: Random Tree and Multiple layer perceptron were prepared using WEKA 3.9.5 software. The electrochemical performance of hetero atom doped nano-carbon electrode material-based supercapacitor was predicted based on collected data from the literature. A total of 16 parameters were selected. Initially, all the datasets were sub divided into four classes. Of all the models considered, RF model provides the best fit of the experiment and predicted values. 94.5 % of the considered dataset is classified correctly proving a better accuracy of the designed models. Results show the superiority of designed models with higher AUROC values.

REFERENCES

1. Murat Ates, Achref Chebil, Ozan Yoruk, Chérif Dridi, Murat Turkyilmaz, *Ionics* **28**, 27 (2022).
2. Richa Dubey, Velmathi Guruviah, *Ionics* **25**, 1419 (2019).
3. Richa Dubey, Velmathi Guruviah, *Arabian J. Sci. Eng.* **48**, 9329 (2022).
4. Richa Dubey, Velmathi Guruviah, *Energy Source. Part A* **44**, 3381 (2022).
5. R. Dubey, V. Guruviah, *Ionics* **28**, 2169 (2019).
6. M. Rahimi, M.H. Abbaspour-Fard, A. Rohani, *Renew. Energy* **180**, 980 (2021).
7. J. Li, K. Han, D. Wang, Z. Teng, Y. Cao, J. Qi, M. Li, M. Wang, *Carbon* **164**, 42 (2020).
8. D. Zhang, S. Shen, X. Xiao, D. Mao, B. Yan, *RSC Adv.* **10**, 26546 (2020).
9. E. Frank, M.A. Hall, I.H. Witten, *Data Min.* **53**, 553 (2017).
10. D. Yates, M.Z. Islam, *Inf. Sci. (Ny)* **557**, 130 (2021).

Ансамблевий підхід для прогнозування ємності електродних матеріалів на основі вуглецю, легованих гетероатомами

Richa Dubey, Velmathi Guruviah, Ravi Prakash Dwivedi

Vellore Institute of Technology – VIT Chennai, Kelambakkam – Vandalur Rd, Rajan Nagar, Chennai, 600127 Tamil Nadu, India

Моделювання машинного навчання на основі ансамблевого підходу використовується в поточному дослідженні для виявлення впливу різних параметрів електродів на електрохімічні характеристики нановуглеців, легованих гетероатомами. Це досягається за допомогою трьох метакласифікаторів у поєднанні з традиційними моделями багаторівневого персептрона та випадкового лісу. Використані три метакласифікатори, а саме (i) пакетування, (ii) класифікація за допомогою регресії (CVR) і (iii) мультикласовий класифікатор (MCC). Серед цих трьох моделей пакетування та класифікація за допомогою регресії забезпечили більшу точність з точки зору правильно класифікованих екземплярів (%) і площі під значеннями області конвергенції. Розроблені моделі використовуються для прогнозування класів питомих значень ємності. 94,5 % розглянутого набору даних класифіковано правильно, що підтверджує кращу точність розроблених моделей. Найменше середньоквадратичне значення 0,1787 було отримано для радіочастотної моделі. У порівнянні з моделями, визначеними в літературі, запропоновані моделі в цій роботі забезпечують найкращу відповідність експерименту та прогнозованих значень з найвищою точністю та найнижчими значеннями продуктивності похибок. Найменше значення похибки для моделей RF і MLP становить 0,18 і 0,19 відповідно.

Ключові слова: Електрод на основі карбону, Зберігання енергії, Легування гетероатомами, Машинне навчання, Легування азотом, Суперконденсатор