Filomat 39:6 (2025), 1909–1916
https://doi.org/10.2298/FIL2506909A

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Machine Learning in optimizing carbon nanotubes and graphene production

Viktor Andonovikj[a,c,*], Aleksandar T. Dimitrov[b], Mimoza Kovaci Azemi[b], Stefan Kuvendziev[b], Beti Andonovic[b]

[a] *Institute Jožef Stefan, Ljubljana, Slovenia*
[b] *Faculty of Technology and Metallurgy, Skopje, N. Macedonia*
[c] *International Postgraduate School Jožef Stefan, Ljubljana, Slovenia*

Abstract. Machine learning plays a key role in this work, designing new technologies for producing carbon nanotubes (CNTs) and graphene by electrolysis in molten salts. The aim is to achieve non-expensive, high-quality materials, making them economically viable for various applications. For the production of multi-walled carbon nanotubes (MWCNTs), experiments employ both non-stationary and stationary current regimes, while for graphene production, constant and reversing cell voltage as well as constant and reversing overpotential methods are considered. The electrolysis process offers ecological and economical advantages with precise control over parameters such as applied voltage, current density, temperature, electrolyte type, and graphite material. To determine the relationship between these parameters and material quality, explainable tree-based Machine Learning (ML) models are employed, trained using labeled data from domain experts. The extracted rules from the ML model guide optimal production, resulting in high-yield materials that are up to ten times more cost-effective than existing technologies. This contributes to the advance of cost-efficient and high-quality carbon nanomaterials for a wide range of applications.

1. Introduction

Nanomaterials are materials that have at least one dimension (length, width, or thickness) in the nanometer scale, typically between 1 and 100 nm (Fig. 1). Nanotechnologies and nanomaterials in the scientific research over the last couple of decades have increasingly focused on carbon nanostructures due to their exceptional properties and potential applications across various fields [2]. Among such nanostructures, carbon nanotubes (CNTs) and graphene, the one atom thick material, have attracted significant attention due to their remarkable electrical, mechanical, and chemical properties.

2020 *Mathematics Subject Classification*. Primary 62H99; Secondary 68T07.
Keywords. machine learning, optimisation, CNT, graphene.
Received: 10 September 2024; Revised: 22 December 2024; Accepted: 26 December 2024
Communicated by Ljubiša D. R. Kočinac
* Corresponding author: Viktor Andonovikj
Email addresses: viktor.andonovikj@ijs.si (Viktor Andonovikj), aco@tmf.ukim.edu.mk (Aleksandar T. Dimitrov), stefan@tmf.ukim.edu.mk (Stefan Kuvendziev), beti@tmf.ukim.edu.mk (Beti Andonovic)
ORCID iDs: https://orcid.org/0000-0002-8901-7890 (Viktor Andonovikj), https://orcid.org/0000-0002-8845-0641 (Aleksandar T. Dimitrov), https://orcid.org/0000-0002-2933-5408 (Mimoza Kovaci Azemi), https://orcid.org/0000-0003-2347-2038 (Stefan Kuvendziev), https://orcid.org/0000-0002-5386-457X (Beti Andonovic)

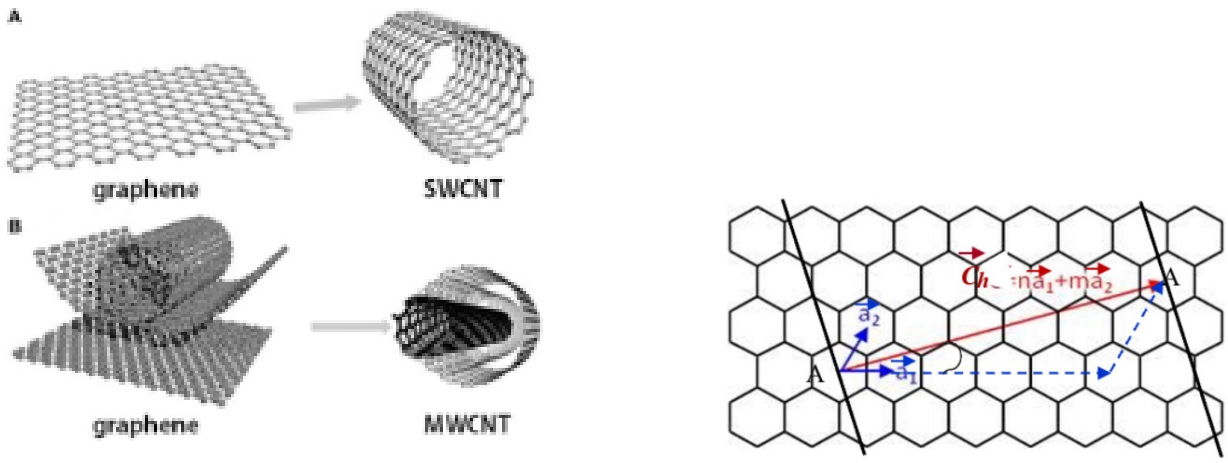

Figure 1: Carbon nanomaterials: graphene and CNTs

Discovered by Lijima [10], CNTs are highly promising in a wide range of applications, including energy storage, materials for electrochemical capacitors, electrocatalytic supports for fuel cells, fillers in polymer composites, etc. Similarly, graphene, first isolated by Geim and Novoselov in 2003, is unique in terms of its extraordinary electrical, thermal, optical, and mechanical properties, earning it the title of "material of the future" [6, 7, 11]. The production of CNTs and graphene has advanced significantly, with various methods developed to optimize their quality and yield. One promising approach to producing CNTs is through the electrolysis of alkali molten salts. This method, pioneered by [9], involves the electrolysis of molten lithium chloride on a graphite cathode, resulting in the formation of multi-walled nanotubes due to the erosion of the cathode during the process. Subsequent research has linked this cathodic erosion to the intercalation of alkali metals into graphite, a process further refined by Fray and colleagues [4, 13]. These advancements have led to the successful incorporation of CNTs in various applications, including as electrodes in lithium-ion batteries, where they have demonstrated improved charge-discharge capacity compared to standard carbon materials [5]. Parallel to the developments in CNT production, the discovery of graphene has motivated extensive research into its applications, ranging from electronics to energy storage. The superior properties of graphene and CNTs have encouraged the development of advanced production techniques, such as chemical vapor deposition (CVD) for graphene and electrolysis in molten salts for CNTs. Moreover, the complex relationship between the quality of CNTs and the parameters of their production process has opened up new approaches for data-driven modeling, including the application of machine learning algorithms. These algorithms offer the potential to optimize production processes by analyzing experimental data and identifying key factors that influence CNT quality [1, 8, 15, 16]. The integration of machine learning in optimizing production processes represents a significant step forward in utilizing the full capabilities of these remarkable materials. This work contributes to production of high quality carbon nanomaterials by an ecological method that is up to ten times cheaper than other existing technologies in the field.

2. Production of Carbon Nanomaterials: Graphene and Carbon Nanotubes (CNTs)

The production of carbon nanomaterials, such as graphene and multi-wall carbon nanotubes (MWCNTs), is evolving towards methods that enhance quality and yield while reducing costs. Graphene and CNTs, typically produced through expensive methods with low yield, can be synthesized more efficiently using electrochemical methods in molten salts, especially with non-stationary current regimes. Our approach allows precise control of parameters like overpotential, current density, and electrolyte temperature,

enabling the production of high-quality graphene at a lower cost. This method enables production of carbon nanomaterisls via ion intercalation during electrolysis, where cations reduce and intercalate at the cathode, causing exfoliation. The process takes place in a graphite crucible with molten chloride salts, such as LiCl, inside a sealed Inconel reactor (Fig. 2).

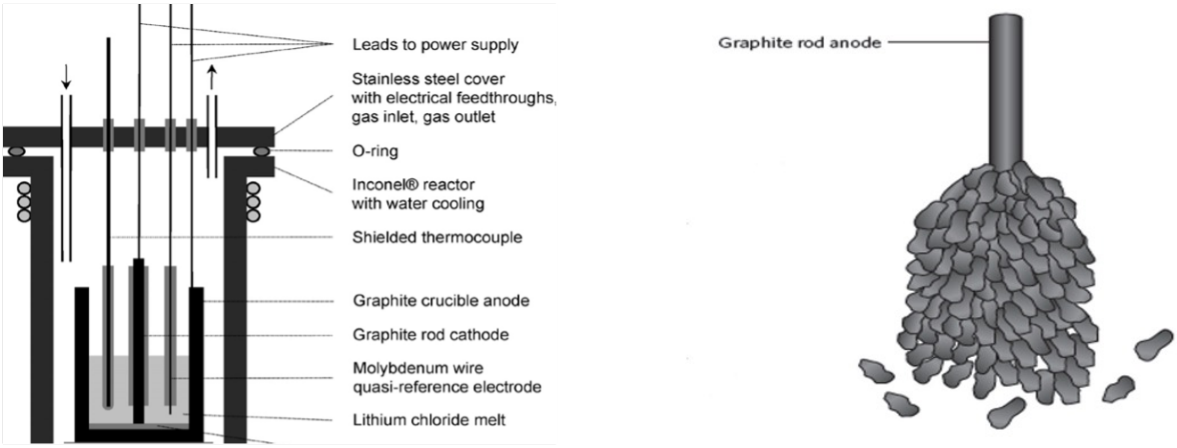

Figure 2: Direct exfoliation of graphene. Schematic of experimental set-up used for conventional electrolysis

Characterization techniques as SEM, TEM, and Raman spectroscopy were used to analyze the obtained structures (Fig. 3).

3. Methodology

Small datasets were subjects of both graphene and CNT research: 50 experiments for producing multi-wall CNTs and 30 experiments for producing graphene. The samples were represented with 9 parameters (features), such as (Table 1):

- Cathode overpotential – measured relevant to a referent electrode
- Type of graphite – EC4, EC17, MSG34
- Time of polarity change – interval in which the polarity of electrodes change
- Experiment type
  - Reversing overpotential
  - Constant overpotential
  - Constant cell potential

A Decision Tree (DT) model was developed to classify the experimentally produced carbon nanomaterials into one of three quality classes (1, 2, or 3). These classes serve as indicators of samples' quality, where a higher class number corresponds to superior quality, characterized by increased yield and fewer defects. The rules extracted from the trained DT model are employed to optimize the large-scale production of cost-effective, high-quality CNTs. Prior to model training, several data preprocessing steps were undertaken to ensure the data was suitable for input into the DT model. The preprocessing and modeling tasks were conducted in Python, utilizing modules such as pandas, NumPy, and scikit-learn. Initially, feature normalization was applied to scale the attributes so that their mean value was zero, while their variance was preserved using their standard deviation. This process ensures that each attribute is given

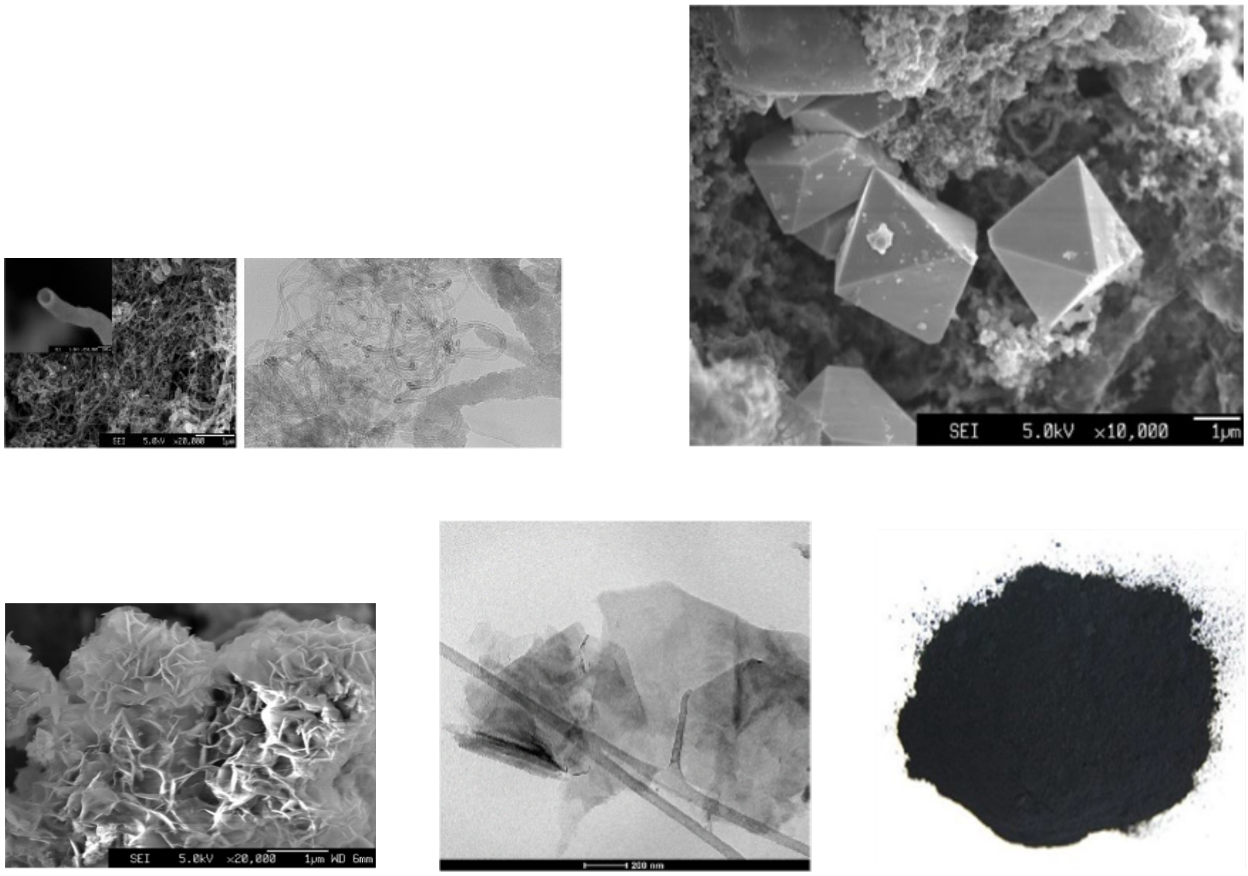

Figure 3: A. TEM and SEM images of MWCNTs obtained experimentally by electrolyses in molten salts at the Faculty of Technology and Metallurgy in Skopje; B. TEM and SEM images of graphene obtained experimentally by electrolyses in molten salts at the Faculty of Technology and Metallurgy in Skopje; C. Nanodiamonds obtained by electrolyses in molten salts; D. Graphene and CNTs in natural dimensions.

equal weight, preventing any single feature from disproportionately influencing the model due to its value range. Among various normalization techniques, we employed the Standard Scaler, defined as:

$$z = \frac{x - \mu}{s} \tag{1}$$

where x represents the actual value of the feature, μ is the mean, and s is the standard deviation.

To model the yield of graphene and carbon nanotubes and using a decision tree, a supervised learning framework was considered, where the dataset $\mathcal{D} = \{(X_i, y_i)\}_{i=1}^{N}$ consists of N instances. Each instance comprises a feature vector $X_i \in \mathbb{R}^d$, representing d features related to the synthesis parameters of CNTs or graphene, and a corresponding target variable $y_i \in \mathbb{R}$, representing the yield.

Data Preprocessing: To prepare the feature matrix $X \in \mathbb{R}^{N \times d}$, for input into the decision tree algorithm, categorical features are encoded using one-hot encoding [12]. This encoding process transforms each categorical variable into a binary vector, expanding the original feature space to $X^* \in \mathbb{R}^{N \times d^*}$, where $d^* \geq d$ accounts for the increased dimensionality due to the encoding.

Model Formulation: The DT model aims to approximate a function $f : \mathbb{R}^d \to \mathbb{R}$ that maps the encoded feature matrix X^* to the target variable y. The tree is constructed by recursively partitioning the feature

Table 1: Data sample for MWCNTs

Electrolyte Temperature	Type of Graphite - Anode	Type of Graphite - Cathode	Cell Potential	Cathode Overpotential	Time of Polarity Change - Cell	Time of Polarity Change - Cathode	Time of Polarity Change - Anode	Experiment Type
780	EC4	EC4	5.5	3.0	600	60	60	Reversing overpotential
780	MSG34	MSG34	4.0	4.0	600	600	600	Constant overpotential
780	EC17	EC17	3.5	2.0	600	60	60	Reversing overpotential
640	EC4	EC4	4.5	2.5	600	60	60	Reversing overpotential
780	EC4	EC4	4.5	2.5	600	60	120	Constant cell potential

space into axis-aligned regions. At each node t of the tree, a split is defined by selecting a feature X_j^* and a threshold θ_j, such that the feature space is divided into two subsets:

$$R_{\text{left}} = \{X^* \in \mathbb{R}^{d^*} | X_j^* \leq \theta_j\}, \quad R_{\text{right}} = \{X^* \in \mathbb{R}^{d^*} | X_j^* > \theta_j\}. \tag{2}$$

The goal is to find the optimal pair (j, θ_j) that minimizes the impurity of the resulting subsets. The impurity at a node t is quantified by the Gini Index (*G.I.*) [3] which for a set $R(t)$ is defined as:

$$G.I.(R(t)) = 1 - \sum_{k=1}^{K} p_k^2(t), \tag{3}$$

where $p_k(t)$ is the proportion of samples in $R(t)$ that belong to class k, and K is the number of classes.

Optimization Problem: At each node t, the optimal split (j^*, θ_j^*) is found by solving the following optimization problem:

$$(j^*, \theta_j^*) = \arg\min_{(j,\theta_j)} \left[\frac{|R_{\text{left}}(t)|}{|R(t)|} G.I.\left(R_{\text{left}}(t)\right) + \frac{|R_{\text{right}}(t)|}{|R(t)|} G.I.\left(R_{\text{right}}(t)\right) \right]. \tag{4}$$

where $|R(t)|$ is the number of samples at node t. The process continues recursively until a stopping criterion is met, such as reaching a maximum depth $D_{\max}$ or reaching a minimum number of samples per leaf $N_{\min}$.

4. Results and Prediction

The performance of the model was tested on a fraction of the original dataset, called a test set. The test set contains randomly chosen 30% of the data instances of the original dataset, while the other part, called a training set, was used for training the model. Classification accuracy was used as a performance metric.

Once the tree is fully grown, the prediction for a new instance $X_{\text{new}}^* \in \mathbb{R}^{d^*}$ is obtained by traversing the tree from the root to a leaf node. The predicted yield $\hat{y}$ is the mean (for regression) or the majority class (for classification) of the target values y_i in the leaf node reached by X_{new}^*.

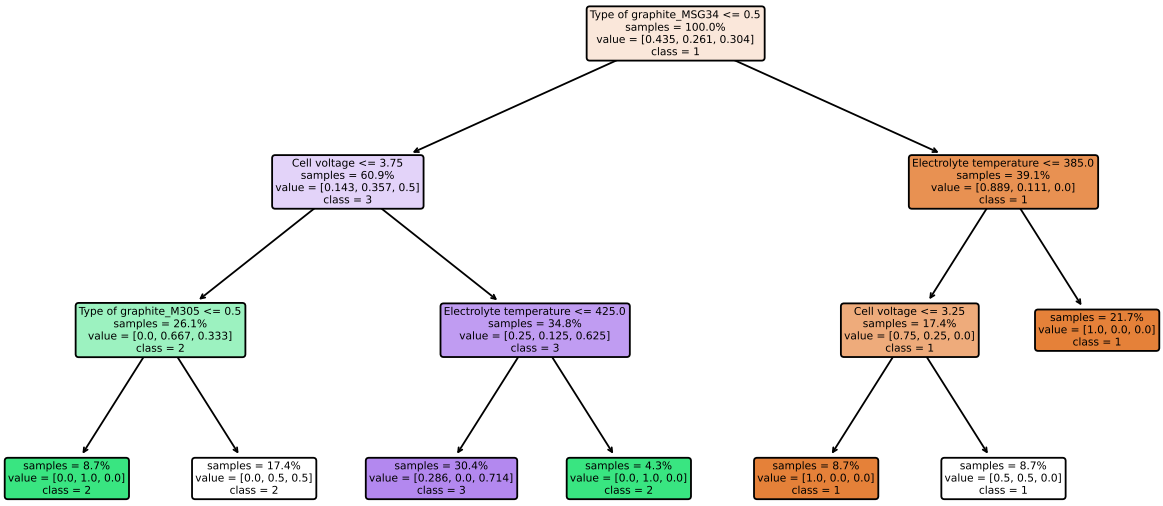

Figure 4: DT model to classify graphene samples

The following rules are extracted for graphene samples:

Class 2: The tree predicts class 2 when the graphite is not of type M305 and the electrolyte temperature is high (above 425°C), or when the graphite is of type M305, the cell voltage is moderate (below 3.75V), and the electrolyte temperature is high (above 385°C).

Class 3: The tree predicts class 3 when the graphite is not of type M305 and the electrolyte temperature is moderate (425°C or lower), or when the graphite is of type M305 and the cell voltage is high (above 3.75V), or when the graphite is of type MSG34 and the electrolyte temperature is moderate (385°C or lower).

Class 1: The tree predicts class 1 when the graphite is of type M305 and the cell voltage is low (3.25V or lower), or when the graphite is of type MSG34, the cell voltage is moderate (above 3.25V), and the electrolyte temperature is high (above 385°C).

The following rules are extracted for CNT production: Class 2: The model predicts class 2 when the electrolyte temperature is moderate to high (above 705°C) or when the time of polarity change at the anode is short (45 minutes or less), and the cathode overpotential is high (above 2.25V). Additionally, if the type of graphite used in the anode is not MSG34 and the time of polarity change at the cathode is short (90 minutes or less), class 2 is predicted.

Class 3: The model predicts class 3 when the cell voltage is low (5.25V or less) and the cathode overpotential is low (2.25V or less).

Class 1: The model predicts class 1 when the electrolyte temperature is low (705°C or less) and the time of polarity change at the anode is long (greater than 45 minutes). It also predicts class 1 when the type of graphite used in the anode is MSG34 or when the time of polarity change at the cathode is long (greater than 90 minutes).

The experimental results underscore the DT model as a powerful tool for optimizing the production of carbon nanomaterials, including both graphene and carbon nanotubes (CNTs). This model is particularly effective in identifying the key parameters that influence the quality of these materials, a task that typically requires specialized expertise in nanomaterials. By integrating the insights gained from the DT model with domain expertise, we established general guidelines for producing high-quality graphene and CNTs. These guidelines enable the production of these materials across various applications without the need for expert knowledge or high consulting costs, thus streamlining the production process, reducing expenses, and

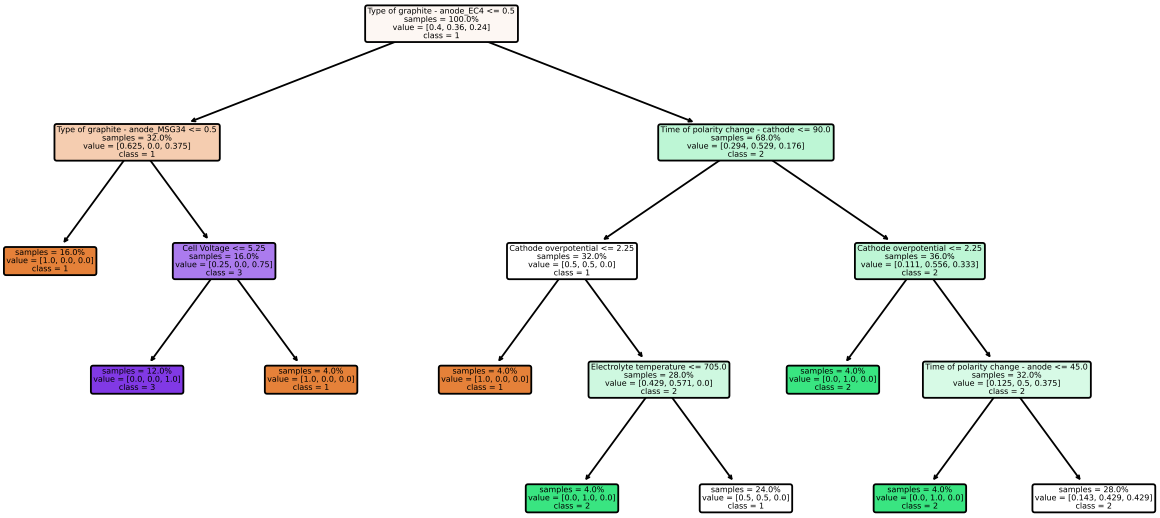

Figure 5: DT model to classify CNT samples

expanding access to superior nanomaterials. In addition to rule extraction, calculating feature importance within the DT model is critical for further optimizing production processes (Fig. 6). Feature importance ranks the parameters based on their influence on the final material quality, providing a clear hierarchy of the most impactful factors. Feature importance measures the decrease in node impurity weighted by the probability of reaching that node. A higher value indicates greater significance. This ranking allows producers to focus on the parameters that matter most, leading to more efficient and effective production. By analyzing the rules derived from the model and validating them with domain expertise, we can ensure the guidelines are robust, allowing for the production of high-quality CNTs and graphene with consistent results. The ability to identify the most critical thresholds at decision nodes provides valuable insight into the parameter ranges that lead to high or low-quality materials, further refining the production process.

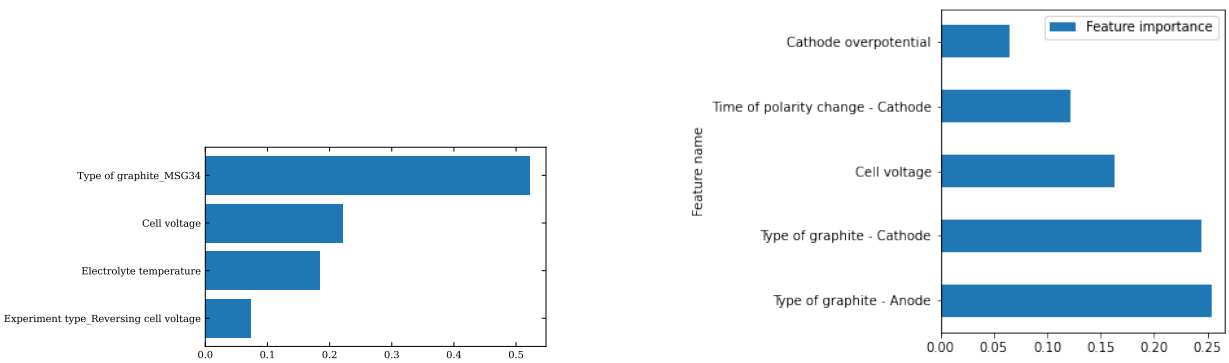

Figure 6: Carbon nanomaterials: graphene and CNTs

5. Conclusion

This research explores a machine learning approach to optimize graphene and multi-walled carbon nanotube (MWCNT) production using electrolysis in molten salts. Electrolysis in molten salts is a simple, ecological, and cost-effective method for producing carbon nanomaterials, and is economical but challenging to control without expensive equipment. Various parameters, such as electrolyte type, voltage, current density, temperature, and graphite type, can be adjusted to enhance the process. Using datasets from experimental graphene and CNT production, a machine learning model was developed to optimize the process. By employing an explainable tree-based machine learning model, rules for optimizing graphene and CNT production were derived. The model was trained with production parameters as inputs and graphene and CNT samples quality as the target, with quality labels provided by experts. The decision tree model, which achieves 80% accuracy, characterizes the obtained samples and describes the relationship between quality and input parameters. The model aligns well with theoretical expertise, providing practical rules for improving graphene and CNT quality. By following these rules, one can produce higher-quality carbon nanomaterials and expand the dataset for better model performance. This method is significantly cheaper, producing high-quality graphene and CNTs at a cost up to ten times lower than other techniques. It simplifies the production process, reducing the need for expert intervention.

References

[1] V. Andonovikj, P. Boškoski, S. Džeroski, B. M. Boshkoska, *Survival analysis as semi-supervised multi-target regression for time-to-employment prediction using oblique predictive clustering trees*, Expert Systems with Applications, **235** (2024), 121246.

[2] B. Andonovic, V. Andonovic, A. T. Dimitrov, M. K. Azemi, *Optimal Selection of Parameters for Production of Multiwall Carbon Nanotubes (MWCNTs) by Electrolysis in Molten Salts using Machine Learning*, ENTRENOVA - ENTerprise REsearch InNOVAtion, **8**(1) (2022), November, DOI:10.54820/entrenova-2022-0003.

[3] T. Daniya, M. Geetha, K. S. Kumar, *Classification and Regression Trees with Gini Index*, Advances in Mathematics: Scientific Journal, **10** (2020).

[4] A. T. Dimitrov, *Study of molten Li_2Co_3 electrolysis as a method for production of carbon nanotubes*, Macedonian Journal of Chemistry and Chemical Engineering, **28**(1) (2009), 111.

[5] A. T. Dimitrov, G. Z. Chen, I. A. Kinloch, D. J. Fray, *A feasibility study of scaling-up the electrolytic production of carbon nanotubes in molten salts*, Electrochimica Acta, **48**(1) (2002), 91–102.

[6] A. K. Geim, K. S. Novoselov, *The rise of graphene*, Nature Materials, **6**(3) (2007), 183–191.

[7] A. K. Geim, *Nobel Lecture: Random walk to graphene*, Reviews of Modern Physics, **83**(3) (2011), 851–862.

[8] R. B. Ghannam, S. M. Techtmann, *Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring*, Computational and Structural Biotechnology Journal, **19** (2021), 1092–1107.

[9] W. K. Hsu, J. P. Hare, M. Terrones, H. W. Kroto, D. R. M. Walton, P. J. F. Harris, *Condensed-phase nanotubes*, Nature, **377**(6551) (1995), 687–687.

[10] S. Iijima, *Helical microtubules of graphitic carbon*, Nature, **354**(6348) (1991), 56–58.

[11] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, A. A. Firsov, *Electric Field Effect in Atomically Thin Carbon Films*, Science, **306**(5696) (2004), 666–669.

[12] K. Potdar, S. Taher, D. Chinmay, *A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers*, International Journal of Computer Applications, **175**(4) (2017), 7–9.

[13] C. Schwandt, A. T. Dimitrov, D. J. Fray, *The preparation of nano-structured carbon materials by electrolysis of molten lithium chloride at graphite electrodes*, Journal of Electroanalytical Chemistry, **647**(2) (2010), 150–158.

[14] C. Schwandt, A. T. Dimitrov, D. J. Fray, *High-yield synthesis of multi-walled carbon nanotubes from graphite by molten salt electrolysis*, Carbon, **50**(3) (2012), 1311–1315.

[15] S. R. Vadyala, S. N. Betgeri, J. C. Matthews, E. Matthews, *A review of physics-based machine learning in civil engineering*, Results in Engineering, **13** (2022), 100316.

[16] C. Xu, S. A. Jackson, *Machine learning and complex biological data*, Genome Biology, **20**(1) (2019).