2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS) | 979-8-3503-7222-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICETSIS61505.2024.10459414

Analysis and prediction of electric vehicle costs: A machine learning-based approach

Abdelfetah Ouadah [1], Belgacem Said Khaldi [1], Abdelhamid Iratni [2], Ahmed Hafaifa [1]
1 Applied Automation and Industrial Diagnostics Laboratory, Faculty of Science and Technology, University of Djelfa 17000 DZ, Algeria
2 Faculty of Science and Technology, University Mohamed El Bachir El Ibrahimi of Bordj Bou Arreridj, 34000 DZ, Algeria
Emails: abdelfetah.ouadah@univ-djelfa.dz; bs.khaldi@univ-djelfa.dz; iratni@univ-bba.dz; a.hafaifa@univ-djelfa.dz

Abstract. Although electric vehicles (EVs) have many benefits for protecting the environment and lowering emissions, their widespread adoption mainly depends on their price. With machine learning (ML) algorithms, costs can be predicted. This research aims to compare the performance of some of the most well-known ML algorithms to determine which algorithm will best predict the price of electric vehicles. To identify the key characteristics, we examined the literature to research the elements that determine the price of electric vehicles in order to estimate their cost. We theoretically compared these ML algorithms to validate our findings and then compared the output of this comparative study to the outcomes of the simulations.

Keywords: Machine learning, Electric Vehicles (EVs), EVs costs, supervised machine learning algorithm.

I. INTRODUCTION

When it comes to lowering CO2 emissions and the transportation sector's reliance on fossil fuels, electric vehicles (EVs) provide a number of benefits. As a result, several nations have established a variety of measures to help them reach their goals for the growth of electric cars and ease the demand on their energy supplies. Governments should establish additional measures to encourage a wider spectrum of EV usage, notwithstanding the rise in EV use in recent years, such as financial incentives, technological assistance, or recharging infrastructure [1-2].

Any invention or device depends on cost. This is linked to the extent of its diffusion to customers, particularly in electric cars, as they have become essential to human life. As a result, cost is a fundamental criterion for the human response to the electric car, more so than other criteria, such as preserving the environment and reducing emissions [3-4].

The high cost of electric vehicles contributes to their success. This is a result of obstacles including the relatively low manufacturing volume of electric vehicle (EVs) in comparison to conventional internal combustion engine (ICE) vehicles and the growth of the supply chain for EV components like batteries, electric motors, and power electronics.

We have examined and studied research articles on the costs of electric vehicles based on state-of-the-art research. As a result, we could identify the variables that affect the cost of electric vehicles. To estimate this cost, we ran simulations using various supervised machine-learning algorithms. We looked at these findings to determine which algorithm would best predict the cost of an electric vehicle.

The following describes how the paper is structured: Section II describes how research papers deal with electric vehicle costs. The cost prediction based on machine learning algorithms is then described in section III, and simulation results are analyzed. Finally, we conclude this work by highlighting its main findings and prospects.

II. COST ANALYSIS OF ELECTRIC VEHICLES

Several research articles and studies are currently analyzing the economics of EVs with a focus on the cost of electric vehicles, including:

In this case study [6], Using a precise technical model of an electric vehicle and a diesel vehicle to the identical section, a techno-economic model has been developed. The study will connect precise and validated car models utilizing a real driving cycle to comprehensive economic models in order to enhance TCO estimations.

The current investigation [7] looks at the conventional, hybrid, and electric vehicle total cost associated with ownership (TCO) during a five-year period in 14 US cities from 2011 to 2015. This study emphasizes how different state and municipal taxes and levies, as well as the price of gasoline, maintenance, and coverage, affect cost variability in the biggest cities in the United States.

This study [8] provides a framework for multiple-objective optimization of electric vehicle (EV) charging prices and/or emissions with minimal computational cost by combining the features of individual EV batteries into a single EV charging model, accounting for the link between the vehicle and the electrical grid.

By combining literature studies and assessments, the research study [9] attempts to provide a more thorough methodology for estimating the overall cost of ownership of electric vehicles. This research makes a scientific contribution by demonstrating the lack of an elaborate framework for calculating the total cost of ownership of electric vehicles.

This paper [10] assesses the costs of battery electric vehicles over 2020-2030 using the best available data on electric vehicle battery and component costs up to 2018. The expected timeframe for parity is also analyzed by evaluation. Prices for electric sports cars, crossovers, and representative electric cars. In the US light commercial vehicle market. Questions about electric vehicle cost parity are generally critical in helping to inform the types of regulatory policies and incentives that would be most effective in transitioning to a mainstream electric vehicle market.

This article [11] examines the cost-effectiveness of several electric propulsion technologies in the German automotive industry in 2020. To do so, this research intends to present a comprehensive analysis of the expenses related to several electric propulsion alternatives, emphasizing PHEVs. The overall expense of maintaining for several automobiles that run to varying extents by electricity is comprehensively analyzed in this article, including the purchase price, continuous running expenses, maintenance costs, and resale prices for different unit configurations.

The number of miles driven, the length of ownership, the battery's residual value, various tax incentives, and a mileage metering system are some of the factors that have an impact on the total cost of ownership (TCO) of electric light commercial vehicles over time. This article [12] explores this development. It illustrates how lowered kilometer fees and tax benefits for conventional vans may help to reduce the overall cost of ownership of electric light commercial vehicles. This piece illustrates how usage may impact the entire cost of ownership for an electric vehicle.

This study examined the variation in ownership costs according to driving style and geographic location [13]. By examining the neighborhood-level geographic diversity of OCR, this work contributes to the corpus of knowledge. In Los Angeles County, the five-year TCO of four BEV-ICEV automobile pairs—the Nissan Leaf and Versa Note, the Chevrolet Bolt and Trax, the Volkswagen Golf and Golf, the Tesla Model 3, and the Toyota Camry—has been assessed for each area.. They offer sensitivity studies that investigate a variety of presumptions, including the discount rate, depreciation, gasoline prices, government subsidies, and annual vehicle miles traveled (AVM).

According to these articles, we can identify the elements that affect the overall costs of electric vehicles, including their brand, model, acceleration, top speed, range, battery backup, countries, efficiency, rapid charging, powertrain, plug type, body style, segment, and seats.
Using supervised machine-learning algorithms, we aim to forecast the price of electric vehicles based on these variables.

III. ELECTRIC VEHICLE COST PREDICTIONS

It is crucial to provide more results when determining the cost of an electric vehicle, considering requirements and spending limits, and future developments.

According to cutting-edge and electric vehicle certification locations, we are going to predict the cost of an electric vehicle based on these evaluation criteria: brand, model, acceleration, maximum speed, range, battery back, Efficiency, Rapid Charge, Powertrain, Plug Type, Body Style, Segment, and Seats.

Using supervised machine learning (SLM) techniques, this cost prediction is generated. For regression prediction, a variety of methods that utilize supervised learning are utilized. These are a few of the most popular algorithms:

- **Linear regression:** One of the most well-liked and straightforward regression methods is linear regression. A linear link between the target and the independent variables is established. [20].
- **Decision trees:** Regression can also use decision trees. Using a series of conditional tests, they divide the feature space into different parts and assign output values according to the leaves of the tree [19-21].
- **Random forests:** This method based on decision trees uses a variety of decision trees to make predictions. Each tree is built on a random subset of features and data [22].
- **Support vector machines (SVMs):** SVMs may be used for regression but are often employed for classification issues. They look for the ideal distance among the points of data using kernel-based algorithms [23].
- **Artificial neural networks (ANNs):** ANNs are models based on the brain's workings. They comprise layers of interconnected neurons and can be used for regression tasks by modifying the weight and bias of the connections between neurons [24].

When comparing machine learning algorithms, several criteria can be considered to assess their performance and determine which is best suited to a specific task. We have chosen a few critical criteria to consider:

1. **Accuracy:** This measures the algorithm's ability to produce accurate results. In machine learning, precision measures the effectiveness of a model in terms of the proportion of accurate results out of the total number of cases. A model that produces no false positives has an accuracy of 1.0. [14]
2. **Learning speed and execution time:** It is essential to consider how quickly the algorithm learns from training data and how long it takes to predict new data. Some algorithms may be faster to learn but less accurate, while others may take longer but perform better [17].
3. **Complexity and interpretability:** The algorithm's complexity can affect the user's ability to interpret and understand it. [14]
4. **Overfitting Tendency:** The issue of overfitting arises while learning from a noisy database. Overfitting is the result of an overly complex model with too many parameters. An overfitted model is inaccurate because the trend does not reflect the reality of the data. Many techniques can be used to mitigate overfitting, including cross-validation, regularisation, early stopping, pruning, Bayesian priors, dropout, and model comparison [18].
5. **Parameterization:** When setting up an algorithm, data scientists adjust parameters. They are numbers that affect the algorithm's behavior, such as error tolerance, number of iterations, or variants of its behavior. The learning time and accuracy of the algorithm can

sometimes depend on the choice of appropriate parameters. Generally, algorithms with significant parameters require more trials to find the right combination [15].

6. **Scalability**: The ability of the algorithm to handle large amounts of data is crucial, especially if you are working with large datasets. When used with large amounts of data, some algorithms can be more efficient regarding execution time and memory consumption.

We have selected the star system to evaluate the algorithm in accordance with the criteria we have specified for the comparison and evaluation of supervised machine learning algorithms (Low [<25%] */Medium [<50%] **/High [<75%] ***/Very High [<100%] ****).

Table.01 compares supervised machine learning algorithms based on the assessment standards.

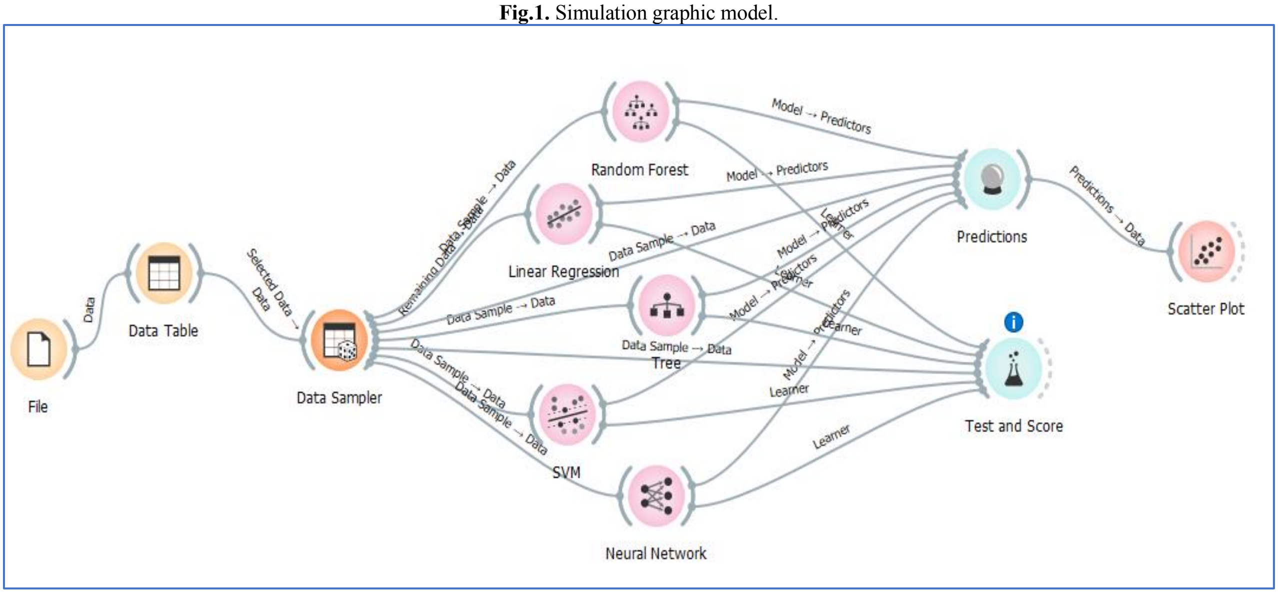
Algo /criteria	Linear regression	Decision trees	Random forests	SVM	ANN
Accuracy	**	**	***	****	***
Learning speed and execution time	***	****	***	**	*
Complexity and interpretability	*	**	***	***	**
Overfitting Tendency	*	***	***	***	**
Parameterization	***	***	***	*	*
Scalability	**	**	***	***	**
Score	12	16	18	16	11

Tab.1. Comparison of algorithms.

To validate this comparison between the algorithms according to the criteria we have chosen from the state of the art, we will predict the price of electric vehicles using these algorithms and compare the results obtained to compare the algorithms in terms of state-of-the-art and terms of simulation.

The selected dataset is downloaded from the Kaggle website [25], containing 104 instances and 14 variables. The target is the price of an electric vehicle. A visualization graph summarizing the simulation and retaining the five machine learning techniques is shown in Figure 01.

Fig.1. Simulation graphic model.

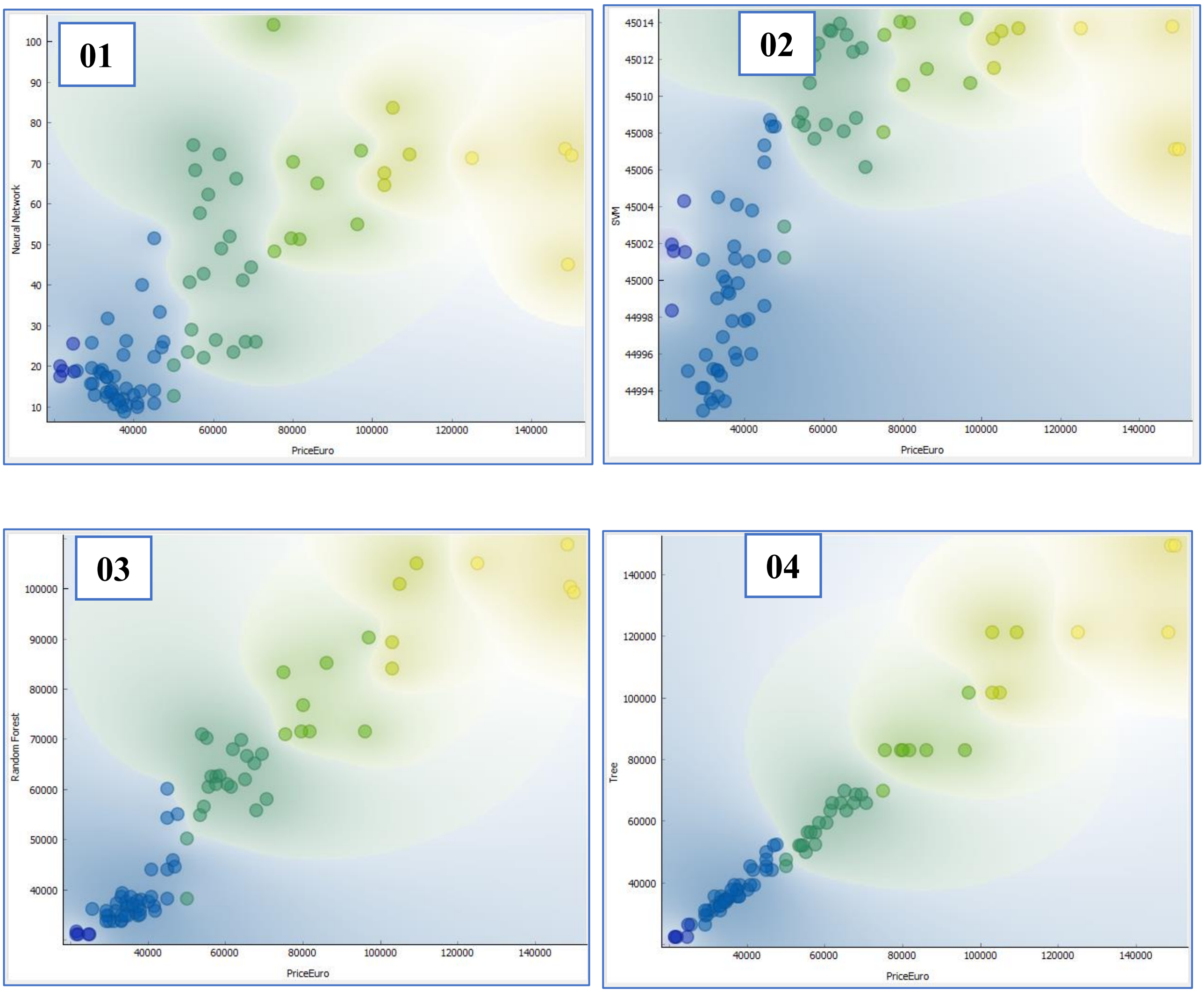


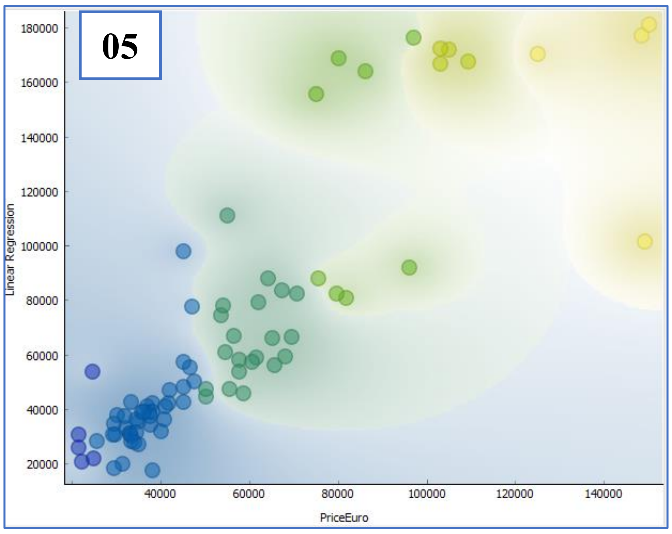
We acquired the following results and displayed the results of machine learning techniques as a function of the cost of electric vehicles according to the data simulated using the algorithms. like that:

- **MSE**: Mean square error
- **RMSE**: Root means square error
- **MAE**: Mean absolute error
- **R2:** Coefficient of determination
- **CVRMSE**: Coefficient of variation of RMSE

Model	Train time [s]	Test time [s]	MSE	RMSE	MAE	R2	CVRMSE
Linear Regression	0.745	0.416	309342633.629	17588.139	9374.677	0.699	30.668
Neural Network	5.928	0.654	4312377717.417	65668.696	57317.244	-3.194	114.505
Random Forest	1.340	0.452	296724999.574	17225.707	10174.97	0.711	30.036
SVM	1.215	0.594	1139126775.997	33750.952	22742.834	-0.108	58.850
Decision Tree	3.435	0.000	387959716.284	19696.693	11852.672	0.622	34.344

Tab.2. Simulation results.





Figures 01, 02, 03, 04, and 05 represent the variation of algorithms (neural network, SVM, random forest, decision tree, and lean regression) as a function of electric vehicle costs.

IV. DISCUSSION AND CONCLUSION

Five supervised machine learning algorithms (Linear Regression, SVM, ANN, Decision tree, and Random forest) were used to simulate the data, and the results are shown in Table 02.

We observe that the Linear Regression algorithm is faster than the others for the train time, the SVM and the Random forest are nearly identical, and the ANN takes longer. The test time yielded the same results, except that the decision tree algorithm's test time is almost zero, but even so, the ANN takes up more time than the others. On the other hand, the RMSE and MAE are altered proportionally for all algorithms, with the Linear Regression algorithm having a lower error than the others, the ANN algorithm having a significant error, and the same for the coefficients of variation (CVRMSE).

These findings lead us to the conclusion that the algorithms for cost prediction should be used in the following order: linear regression, random forest, decision tree, SVM, and finally, the ANN.

The Linear Regression algorithm in the bibliographic part is ranked fourth by Conte in the simulation when we compare the performance of the algorithms in the simulation with the comparison we made from the state-of-the-art. Depending on the size of the dataset, this variation of location occurs. The operating principles of the decision tree and the random forest are remarkably similar. The random forest exhibits similar performance regardless of the simulation or biographic data.

The ANN algorithm, the foundation of the bibliographic study, is a subpar predictor of cost.

It is critical to remember that electric car costs are steadily declining as technology advances, production volumes increase, and the supply chain is consolidated. As these elements develop, it is anticipated that the price of electric vehicle (EVs) will eventually be comparable to that of vehicles powered by internal combustion engines.

Because the simulation data are small datasets, this study has some limitations. The five supervised machine learning algorithms should be used to simulate both small and large datasets, and the results and performance of each algorithm should be compared across all dataset sizes.

V. ACKNOWLEDGMENT

We would like to express our gratitude to Applied Science University in Bahrain (https://www.asu.edu.bh/) for offering free registration to the students who authored this academic paper.

VI. REFERENCES

[1] European Environment Agency. (2016). Electric Vehicles and the Energy Sector—Impacts on Europe's Future Emissions.

[2] Thangavel, S., Deepak, M., Girijaprasanna, T., Raju, S., Dhanamjayulu, C., & Muyeen, S. M. (2023). A Comprehensive Review on Electric Vehicle: Battery Management System, Charging Station, Traction Motors. IEEE Access.

[3] Qiu, D., Wang, Y., Hua, W., & Strbac, G. (2023). Reinforcement learning for electric vehicle applications in power systems: A critical review. Renewable and Sustainable Energy Reviews, 173, 113052.

[4] DOUILLARD, C., & AUDETTE, S. Comparaison des coûts totaux de possession de véhicules électriques et conventionnels au Québec.

[5] Solanke, T. U., Ramachandaramurthy, V. K., Yong, J. Y., Pasupuleti, J., Kasinathan, P., & Rajagopalan, A. (2020). A review of strategic charging–discharging control of grid-connected electric vehicles. Journal of Energy Storage, 28, 101193.

[6] Desreveaux, A., Hittinger, E., Bouscayrol, A., Castex, E., & Sirbu, G. M. (2020). Techno-economic comparison of the total cost of ownership of electric and diesel vehicles. IEEE Access, 8, 195752-195762.

[7] Breetz, H. L., & Salon, D. (2018). Do electric vehicles need subsidies? Ownership costs for conventional, hybrid, and electric vehicles in 14 US cities. Energy Policy, 120, 238-249.

[8] Brinkel, N. B. G., Schram, W. L., AlSkaif, T. A., Lampropoulos, I., & Van Sark, W. G. J. H. M. (2020). Should we reinforce the grid? Cost and emission optimization of electric vehicle charging under different transformer limits. Applied Energy, 276, 115285.

[9] Van Velzen, A., Annema, J. A., van de Kaa, G., & van Wee, B. (2019). Proposing a more comprehensive future total cost of ownership estimation framework for electric vehicles. Energy Policy, 129, 1034-1046.

[10] Lutsey, N., & Nicholas, M. (2019). Update on electric vehicle costs in the United States through 2030. Int. Counc. Clean Transp, 12.

[11] Propfe, B., Redelbach, M., Santini, D. J., & Friedrich, H. (2012). Cost analysis of plug-in hybrid electric vehicles, including maintenance & repair costs and resale values. World Electric Vehicle Journal, 5(4), 886-895.

[12] Lebeau, P., Macharis, C., & Van Mierlo, J. (2019). How to improve the total cost of ownership of electric vehicles: An analysis of the light commercial vehicle segment. World Electric Vehicle Journal, 10(4), 90.

[13] Parker, N., Breetz, H. L., Salon, D., Conway, M. W., Williams, J., & Patterson, M. (2021). Who saves money buying electric vehicles? Heterogeneity in the total cost of ownership. Transportation Research Part D: Transport and Environment, 96, 102893.

[14] Ouadah, A., Zemmouchi-Ghomari, L., & Salhi, N. (2022). Selecting an appropriate supervised machine learning algorithm

for predictive maintenance. The International Journal of Advanced Manufacturing Technology, 119(7), 4277-4301.

[15] Microsoft Build https://docs.microsoft.com/fr-fr/azure/machine-learning/how-to-select-algorithms Consulter le 22/06/2023

[16] Vansh Jatana, SRM Institute of Science and Technology (Juin;2019) Machine Learning Algorithms.

[17] Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine learning algorithms: classification and comparison. International Journal of Computer Trends and Technology (IJCTT), 48(3), 128-138.

[18] Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), 2(4), 2231-2307.

[19] Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9, 381-386.

[20]] Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, *4*, 51-62.

[21] Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine learning algorithms: classification and comparison. International Journal of Computer Trends and Technology (IJCTT), 48(3), 128-138.

[22] Alsharif, M. H., Kelechi, A. H., Yahya, K., & Chaudhry, S. A. (2020). Machine learning algorithms for smart data analysis in internet of things environment: taxonomies and research trends. Symmetry, 12(1), 88.

[23] Ray, S. (2019, February). A quick review of machine learning algorithms. In 2019 International Conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 35-39). IEEE.

[24] Méndez, M., Merayo, M. G., & Núñez, M. (2023). Machine learning algorithms to forecast air quality: a survey. Artificial Intelligence Review, 1-36.

[25] https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset