Journal of Basic Physical Research ISSN: 2141-8403 Prints, 2141-8411 Online
Vol.13, August, 2024, pp. 12 - 20. Available online at www.jbasicphyres-unizik.org

OPTIMIZING MACHINE LEARNING-DRIVEN MOBILE CHARGING STATIONS IN POWER-CONSTRAINED ENVIRONMENTS

Ezuruka E. O.[1], Anigbogu S. O.[2], Ekwealor O. U.[3]
[1,2,3]Department of Computer Science, Nnamdi Azikiwe University Awka.
Email Address: eo.ezuruka@unizik.edu.ng, so.anigbogu@unizik.edu.ng ou.ekwealor@unizik.edu.ng
Corresponding author eo.ezuruka@unizik.edu.ng

Abstract
Mobile charging stations serve as essential sources of power for a range of devices, such as electric cars, medical equipment, and communication devices in order to maintain the functionality and security of vital activities. A distinct set of optimization problems, such as resource allocation, forecasting model development, power management, cost optimization, and so on, arise when mobile charging stations are deployed and managed in energy-constrained environments. This work addresses the optimization difficulties in the deployment and management of mobile charging stations by utilizing state-of-the-art strategies to improve the operational efficiency and resilience. The system emphasized on employing supervised learning as a kind of machine learning (ML), in which the model was tested and trained on a labeled dataset derived from a range of energy sources. These sources include solar energy systems, generators, and portable battery packs to get around these optimization challenges. The support vector machine (SVM) and multilayer neural networks (MLNP) along with Naïve Bayesian(NB) optimization strategies were developed to boost the performance of mobile charging stations. Artificial neural network (ANN) outperformed the support vector machine (SVM) by a significant margin when selecting the target variable. The black-box feature incorporated into the framework reduced error probability and promoted high standards, all while increasing the efficiency and reliability of the model learning process. This study provides an approach for places with limited energy resources and promotes the shift to more environmentally friendly means of transportation while simultaneously enhancing the accessibility, efficiency, and dependability of electrical vehicle charging services.
Keywords: *Electric vehicle, Feature engineering, preprocessing, MLPN, Optimization, SVM*

Introduction

The availability of power has become a significant challenge, particularly in regions without established networks of charging stations. Public charging stations, powered by various energy sources such as solar systems, generators, or portable battery packs, have emerged as a solution to this issue, providing a vital resource for mobile devices and electric vehicles (Gruoss *et al*., 2020). The adoption of electric vehicles (EVs) promises to reduce noise and local pollution, contributing to sustainable transportation (Chen *et al*., 2020). However, the widespread use of EVs is hindered by the inadequate network of charging stations, which are often poorly planned and limited in number (Majidpour *et al*., 2019). In response to these challenges, researchers have explored various energy optimization strategies, including the application of machine learning (ML) techniques, to improve the efficiency of charging infrastructure. ML-driven optimization offers a data-centric approach to solving complex problems by leveraging computational power to enhance system performance. This approach is particularly valuable in areas where energy scarcity not only limits economic development and educational opportunities but also exacerbates social inequalities by restricting access to essential services and communication technologies (Lee *et al*., 2019). The importance of optimizing mobile charging stations becomes evident in energy-constrained regions, where consistent and reliable energy supply is crucial for socio-economic development. Numerous studies have examined the challenges and potential solutions for EV charging infrastructure. For instance, the Electric Power Research Institute identified three charging levels—Level 1, Level 2, and Level 3—each offering varying charging speeds and suitability for different environments (Mies *et al*., 2019). Additionally, the dynamic nature of non-residential charging stations, which can be influenced by factors like location and time, underscores the need for intelligent resource allocation to meet varying demand patterns (Shahriar *et al*., 2021).

Machine learning techniques, such as Decision Trees, K-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and random forests, have been employed in various studies to predict EV charging and energy consumption rates, leading to more efficient operations (Liew *et al*., 2019; Zamir *et al*., 2020). The integration of ML in optimizing mobile charging stations not only enhances their reliability and accessibility but also provides cost savings by minimizing downtime and improving scheduling, which is especially important in regions with limited energy availability.

Literature Review

The Electric Power Research Institute identified three levels of charging for electric vehicles: Level 1, Level 2, and Level 3. Level 1 operates at 120V/15A and provides the slowest charging rate, typically through a standard electrical outlet, making it suitable for home use. Level 2 uses 240V AC, offering faster charging for both public and private settings, though it requires specialized charging equipment. Level 3, known as DC fast charging, uses 480V AC and is found in public and commercial locations, offering the quickest charging times, capable of fully charging a vehicle in under thirty minutes. Charging for electric vehicles can be categorized into domestic and non-residential types. Domestic charging, typically involving Level 1 or Level 2, benefits from predictable pricing and straightforward scheduling, with owners often charging their vehicles overnight. In contrast, non-residential charging is more dynamic, with usage patterns influenced by factors such as location, time of week, and weather conditions. For example, the demand at a public charging station near a shopping mall can vary significantly depending on these factors. Understanding these patterns is essential for optimizing resource allocation and enhancing the efficiency of charging infrastructure.

Materials and Methods

(a). Datasets: The dataset was sourced from publicly made available kaggle site at "https://data.cityofpaloalto.org/dataviews/257812/ELECT-VEHIC-CHARG-STATI-83602/".
A few of its features are station name, MAC address, start date, start time, end date, total duration, charging time, power split transmission (PST), pulse discharge test (PDT) and etc, With over 2000 testing and 8000 training sets making up the 10,000 dataset, the suggested collection provides a sizable dataset for building models.

Table 1: Dataset for EV **(source:** https://data.cityofpaloalto.org/dataviews/257812/ELECT-VEHIC-CHARG-STATI -83602/**")**

	Station Name	MAC Address	Org Name	-	Post User_id
0	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	3284
1	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	4169
2	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	4169
3	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	2545
4	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	3765
-	- - -	- - -	- - -	-	- - -
9995	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	55033
9996	PALO ALTO CA / HAMILTON #2	000D:6F00:009E:D39E	City of Palo Alto	-	126779
9997	PALO ALTO CA / HIGH #4	000D6F0000A20F47	City of Palo Alto	-	139203
9998	PALO ALTO CA / HAMILTON #1	000D:6F00:015A:9D76	City of Palo Alto	-	2670
9999	PALO ALTO CA / BRYANT #2	000D6F0000A2108E	City of Palo Alto	-	134699

(b). **Data pre-processing:** The preprocessing is encapsulated in set of routines capable of filtering instances or attributes. The data preparation phase is utilized in order to identify and deal with erroneous values and eliminate missing data values from the current system dataset. The date and the vehicle charging parameters are preprocessed into an appropriate format prior to training. The missing value replace function was used during the preprocessing stage to fill the values for the training dataset before creating the model.

(c). Feature engineering: The feature engineering technique is employed to transform data to have meaningful representation using human knowledge. This is very important and intensive which acts as a weakness to learning models. This phase relies mainly on human ingenuity and prior knowledge to compensate for the inability of algorithms to extract and organize the discriminative information from dataset (Bengio *et al*. 2020).

(d). Support Vector Machines (SVM) Classifier: The SVM is one of the simplest and more preferred machine learning techniques used by data professionals because it tendency of producing better and high accuracy with less computational error (Zhang and Jiang, 2012). The SVM uses two main concepts namely; hypothesis space and the loss functioning finding an "optimal" hyper-plane as a solution to any learning problem (Bryant and Allen, 201**3**). The SVM is memory efficient and uses subsets of training data points in the support vectors called decision function. The simplest formulation of SVM is the linear one, where the hyper-plane lies on the space of the input data (Byun and Lee, 2003). The SVM estimators are defined on the training dataset and tested to effectively predict the target variable. We created an optimizer and invoked a black box function directly instead of specifying the optimization directly.

The SVM can be experienced as:

$$\min \frac{1}{2}|w|^2 + c\sum_{i=1}^{n} \xi_1 \quad 1$$

$$y_i(wx_1 - b) \geq 1 - \xi \quad \xi \quad \geq 0 \quad 2$$

Where x is the input vector and w is the vector that corresponds to the hyper-parameter plane, and i is the feature's dimensionality, which can range from 1, 2, 3,...,n Whereas the ξ indicates no separable data input, C is the capacity constant. A SVM classifier was invoked from the sklearn.svm library in python and SVM model created. The gamma variable set to be scalable, c=1,0 and random states set to 101 with the Python script: svc=SVC(gamma='scale', C=1.0, rando_state=101). The model was trained with training dataset with svm.fit (X_train,y_test) and predicted using the testing dataset. The visualization was done using mat_plot_lib library in python. A SVM classifier was created with the pre-processed training data to make predictions about employees exit. The researcher employed SVM because it can work well with unknown data distribution with high speed and better accuracy compared to other algorithms. The kappa hyperparameter was set to 1.96 and xi = 0.01 to strike a balance between the two processes. We are simultaneously optimizing C and degree in order to have the optimizer recommend new parameter values for us to explore with the designated acquisition function.

(e). Artificial neural network (MLPN): An approach known as artificial neural networks (ANNs) mimics the way the human brain works. In these networks, learning occurs via a collection of basic processing known to be neurological systems (Darmawan *et al*.2018) The ANN processes information similarly to how the organic nervous system of humans does[8]. The input, hidden, and output layers are its three distinct layers. The first input layer receives input from the outside world, which is then forwarded to the hidden and output layers. One way to depict the ANN is as:

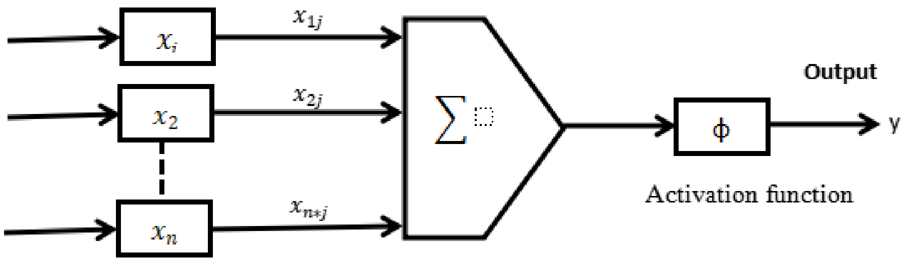

Figure 1: Basic components of ANN(**Source:** Shi et al.2022)

Where w_{1j}, w_{2j} , . . . w_{nj} are the neurons of the input layer and x_1, x_2, x_3,... x_n are the vectors of inputs with corresponding heights. The summation or additive junction is represented by the letter sigma while the activation and output variables are denoted φ and y respectively.

(e).. Model training: The libraries given below are used in Python to train the suggested model.
(i). Scikit-learn: is a popular and freely available Python framework for machine learning predictive data analysis (Ferreira, 2018). The multi-layer perceptron neural network and the support vector machine. We imported the Sci-learn package in order to train two distinct machine learning models.
(ii). TensorFlow is an open-source deep learning system Known as the "big daddy" of deep learning frameworks. The dataset module is being used to build a unique dataset that will be fed into the training model.

(f). Metrics of Evaluation: Comparing the various outcomes requires the use of a consistent model diagnostic tool. Model anticipated results for scenarios involving multi-classification tasks, like the one proposed here, can be visualized in a variety of ways. Standard model evaluation metrics like accuracy, confusion matrices, and ROC learning curves are used to evaluate the efficacy of MLPNN and SVM algorithms. Classification accuracy is defined as the proportion of the dataset's data points that were correctly classified. The error rate can be measured with the general equation given by:

$$\text{Accuracy}=\frac{\text{Total number of correct classification}}{\text{Overall number of classes}}=\frac{TP+TN}{TP+TN+FP+FN} \quad 3$$

Where TN represents true negative, FP is false positive, TP is true positive and FN is false negative cases.

Results and Discussion
The results of the proposed model are shown and discussed with reference to the relevant SVM and MLPN classification approaches. We demonstrated and interpreted experimental results using well-known AI/ML techniques. Better and more accurate findings have been obtained by improving both the theory and its application.

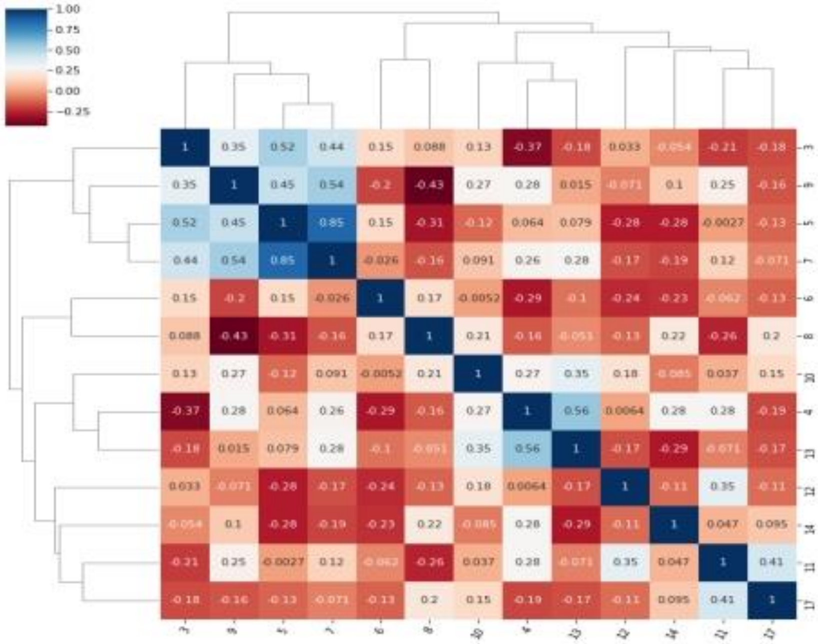

***Figure 2**: EV correlation metrics*

The EV feature correlation matrix of the dataset for our suggested system is displayed in Figure 2, highlighting the highly associated feature clusters that have a major impact on the predictions. The relevant EV dataset was arranged using a cluster map-plot function in Python, which resulted in a tree-like dedogram and graphic representations of connected attributes. Figure 1 illustrates the presence of highly correlated features at the top-left, major diagonal, and other areas. Few features with low correlation exist in close proximity of the plot area.

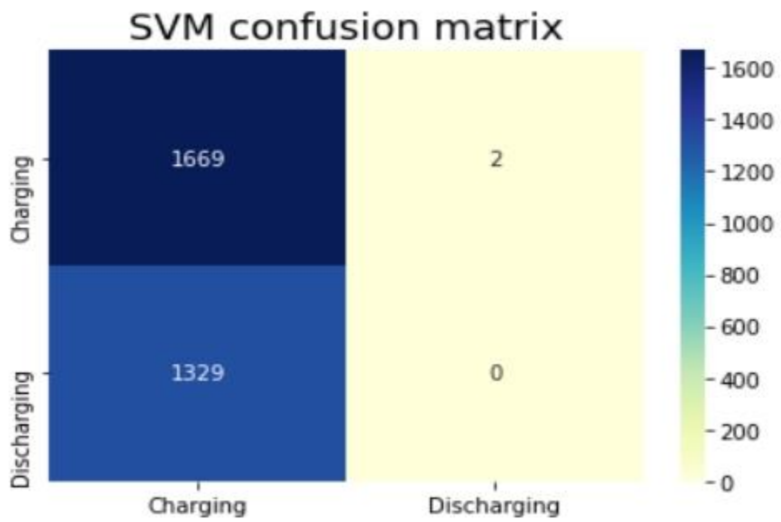

***Figure 3**: Confusion matrix of SVM*

The SVM confusion matrix, which presents a table structure of the numerous predicted outcomes, is seen in Figure 3. The primary diagonal displays the total number of correctly predicted values, while the secondary diagonal shows the total number of incorrectly predicted values. Out of a total of 1671 Charging cases, the confusion matrix showed that 2 were wrongly predicted while 1669 were correct. With no accurate forecasts, the Discharging cases provided 1329 wrong predictions.

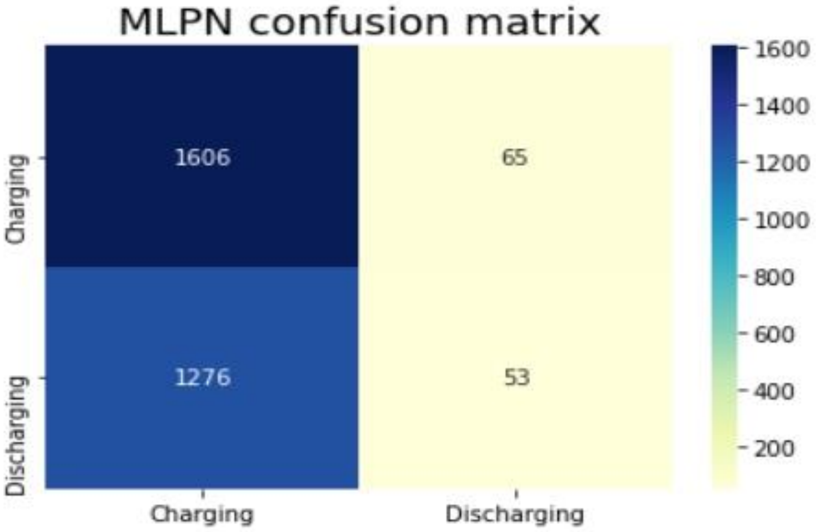

***Figure 4**: Confusion matrix of MLPN*

Figure 4 illustrates a confusion matrix (4×4) used to evaluate the performance of multilared neural network model. The results show that Charging cases reported 1606 accurately predicted cases with 65 mistakenly classified classes and the PST, which produced53 correct predictions with 1276 incorrectly classified data classes. The MLPN confusion matrix, for which the true values are known, is used in Figure 4 above to illustrate how well a classification method works on a set of experimental data.

iter	target	C
1	0.9998	9.674
2	0.9999	5.518
3	0.9998	9.73
4	0.9999	7.177
5	0.9999	7.008
6	0.9999	5.017
7	0.9999	4.313
8	0.9998	8.171
9	0.9999	4.313
10	0.9999	6.096
11	0.9998	0.1
12	0.9998	2.393
13	0.9998	8.885
14	0.9999	1.082
15	0.9999	5.233

***Figure 5**: Results of optimizer*

Figure 5 depicts resultps of proposed optimization system and the optimizer deduced that the model with serial number 6 performs optimally when the hyper-parameter value C = 5.016 are applied. An acquisition function named bayes_opt parameter was applied, along with a utility function designated as the upper confidence boundaries, or "ucb".

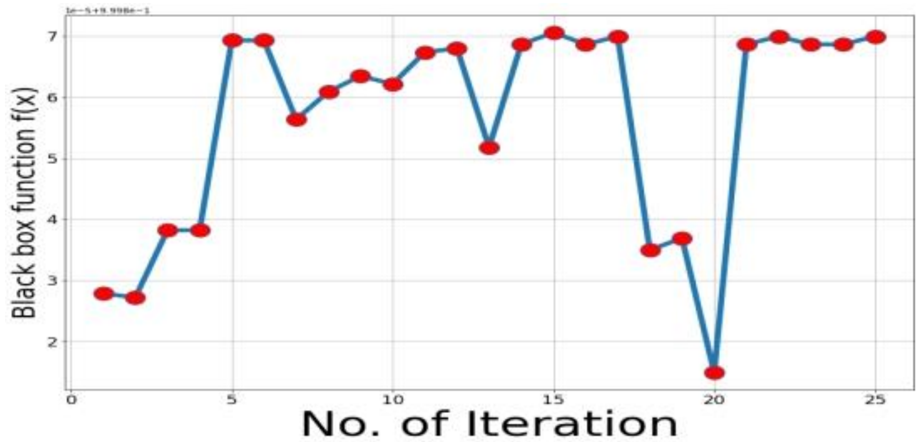

Figure 6*: Graph of optimizer*

The optimizer graph shown in Figure 6 features a blue line connecting each point in the search space for optimal performance, while the red dots represent rending points. The optimizer deduced that the optimal SVC model utilizes hyperparameter values of C = 5.277 and degree = 1.0 based on the earlier findings.

Table 2: Time complexity of model

Model	Training Time	Testing Time
SVM	4.81	3.97
MLPN	0.51	0.00

Table 2 shows the SVM and MLPNN training and testing time complexity. The SVM required more time for testing and training than the MLPNN. The SVM reported the longest training time, 4.81 seconds when compared to MLPNN, which provided 0.51 seconds.

Conclusion

The MLPN technique performed better than the SVM technique when trained using a black-box function for forecasting the duration of EV sessions and energy consumption. The Python programming language was utilized to make the implementation process easier, and this model is scalable and can serve as a norm for other professionals. It contains several deployable deep learning classes and libraries that are accessible via a limited set of performance-optimized scripts. The Bayesian optimizer was used with a black-box function throughout each training cycle (iteration) which reduced the computational cost. These findings indicate a significant opportunity to optimize prediction techniques for charging stations, enabling EV-based mobility service providers to create intelligent charging scheduling plans. Likewise, the successful implementation of these novel approaches would necessitate rapid exchange of data and a fully integrated, low-latency technology network, which may involve decentralized network problems or those unique to the research on wireless

Refrences

Agrawal, R., and Imielinski, T. (2019).A Swami: Mining Association Rules between Sets of Items in Large Databases, *Paper presented at the SIGMOD Conference*.

Amara-Ouali, Y, Goude, Y, Massart, P, Poggi, J. M, Yan H.(2021) A review of electric vehicle load open data and models. Energies, 14(8), 2233.

Bae, S. and A. Kwasinski, A.(2019) Spatial and temporal model of electric vehicle charging demand, IEEE Transactions on Smart Grid, 3(1), 394–403.

Brynant, P. C. & Allen, D. G. (2013). *Compensation, Benefits and Employee Turnover: HR strategies for Retaining Top Talent, Compensation and benefit review,* 45(3), 171-175.

Byun, H. and Lee, S. W. (2003). *A survey on Pattern Recognition Applications of support Vector Machines.* International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), 17(03), 459-486.

Chen, T. and Guestrin, C. (2020) XGBoost: A scalable tree boosting system, in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 785–794.

Chung, Y. W., Khaki, B., Li, T., Chu, C. and Gadh, R. (2019) Ensemble machine learning-based algorithm for electric vehicle user behavior prediction, Appl. Energy, 254(113732), 1-10.

Darmawan, M. F., Jamahir, N. I., Saedudin, R. D. R. and Kasim, S.(2018) Comparison between ANN and Multiple Linear Regression Models for Prediction of Warranty Cost," Int. J. Integr. Eng., 10(6), 2-20.

Dasgupta, D., and Majumdar, N. (2019). Anomaly Detection in Multidimensional Data using Megative Selection Algorithm, Paper presented at the IEEE Conference on Evolutionary Computation. Hawaii.

Ding, T., Zeng, Z., Bai, J., Qin, B., Yang, Y. and Shahidehpour, M. (2020) Optimal electric vehicle charging strategy with Markov decision process and reinforcement learning technique. IEEE Trans. Ind. Appl., 56, 5811–5823.

Ferreira, R. P.(2018) Artificial Neural Network for Websites Classification with Phishing Characteristics, Social Networking. 7(2), 97–109.

Fujimaki, R.,Yairi, T., and Machida, K. (2020). An Approach to Spacecraft Anomaly Detection Problem using Kernel Feature Space, Paper Presented at the In Proceeding of the Eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, NY, USA.

Gruoss, G., Mion, A., and Gajani, G. S.(2020) Forecasting of Electrical Vehicle impacton infrastructure: Markov chains model of charging stations occupation. E Transportation 2020; 6:100083. https://doi.org/10.1016/j.etran.2020.10008.

Hanumantha K. R., Srinivas, G., Damodhar, A., and Vikas M. K. (2020). Implementation of Anomaly Detection Technique Using Machine Learning Algorithms, *International Journal of Computer Science and Telecommunications (IJCST)*, 2(3), 25-31.

Jain, A. K, Yadav, S. K. and Choudhary, N.(2020) A novel approach to detect spam and smishing SMS using machine learning techniques, International Journal of E-Services and Mobile Applications (IJESMA), 12(1), 21-38.

Jupin, J. A., Sutikno, T., Ismail, M. A., Mohamad, M. S., Kasim,S. and Stiawan, D.(2019), Review of the machine learning methods in the classification of phishing attack, Bulletin of Electrical Engineering and Informatics, 8(4), 1545- 1555.

Lazarevic, A., Ertoz, L., Kumar, V., Ozgur, A. and Srivastava, J. (2019). *A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection,* Paper presented at the SIAM International Conference of Data Mining.

Lee, C. Lee, M. and Han, D.(2019) Energy efficient location logging for mobile device,‖ in SAINT, 84, 1-20.

Lee, Z. J., Li, T. and Low, S. H.(2019) ACN-data: Analysis and applications of an open EV charging dataset, in Proc. 10th ACM Int. Conf. Future Energy Syst., New York, NY, USA, 2019, 139–149, doi: 10.1145/3307772.3328313.

Li, Y., Hu, Y., Zhu, H., Wang, H., & Zhang, X. (2022). IEEE Transactions on Industrial Informatics, 18(1), 196-204.

Liaw, A. and Wiener, M. (2020) Classification and regression by random forest, R News, 2(3), 18–22.

Lu, Y., Li, Y., Xie, D., Wei, E., Bao, X., Chen, H.and Zhong, X.(2019) The application of improved random forest algorithm on the prediction of electric vehicle charging load, Energies, 11(11), 3207.

Majidpour, M., Qiu, C., Chu, P., Gadh, R. and Pota, H. R.(2019) Fast prediction for sparse time series: Demand forecast of EV charging stations for cell phone applications, IEEE Trans. Ind. Informat., 11(1), 242–250.

Majidpour, M., Nah, Q. John, A. and dozh, H. R.(2020) A novel forecasting algorithm for electric vehicle charging stations, in Proc. Int. Conf. Connected Vehicles Expo (ICCVE), 1035–1040.

Mohammad K. H., and Doreswamy, H. (2019). Machine Learning Based Network Anomaly Detection, *International Journal of Recent Technology and Engineering (IJRTE), Blue Eyes Intelligence Engineering and Sciences Publication,* 8(4), 542-548.

Mies, J. Helmus, J. and Hoed, V. D.(2019) Estimating the charging profile of individual charge sessions of electric vehicles in The Netherlands, World Electr. Vehicle J., 9(2), 17.

Sarker, M. R. Pandžic, H. Sun, K. and Ortega-Vazquez, M. A.(2018) Optimal operation of aggregated electric vehicle charging stations coupled with energy storage," IET Generation, Transmission & Distribution, 12(5), 1127– 1136.

Shi, J., Huang, Y., Guo, X., Zhang, Z. (2022). App Energy, 308, 117648.

Talwar, A., and Kumar, Y. (2013). Machine Learning: An artificial intelligence methodology. *International Journal of Engineering and Computer Science, 2(12)*, 3400-3404.

Wang, Y. D., and Pan, V. (2021). Anomaly Detection in Traffic using L1-norm Minimization Extreme Learning Machine, *Neurocomputing, 149*, 415-425.

Zhang, X. & Jiang, S. A. (2012). A Splitting Criteria Based on Similarity in Decision Tree Learning, *JSW,* 7(8), 82-1775.

Zamir, A., Khan, H. U., Iqbal, T., Yousaf, N., Aslam, F., Anjum, A. (2020). Phishing web site detection using diverse machine learning algorithms, The Electronic Library. 45.