Optimizing Electricity Theft Detection using Neural Turing Machines and Machine Learning

1st Erdal Büyükbiçakci
Department of Computer Technologies
University of Applied Sciences
Information Technologies Vocational Highschool
Sakarya, Turkey
erdal@subu.edu.tr

2nd Selcuk Bulat
Department of Nanoscience and Nanoengineering, Institute of Natural Sciences, Sakarya University, Turkey 2
Turkish Standards Institution
Istanbul, Turkey
sbulat@tse.org.tr

3rd Senthilkumar C
Department of Artificial Intelligence and Data Science, Erode Sengunthar Engineering College
Erode, India
scsenthilkumar@gmail.com

4th Rajesh A
Department of CSE, Erode Sengunthar Engineering College
Erode, India
rajeshari@gmail.com

5th Malatthi Sivasundaram
Department of Computer Science and Design
KSR College of Engineering
Tiruchengode, India
malathi.gurunathan@gmail.com

6th Rama Raju SVSV Palla
Department of CSE
Vignan's Institute of Information Technology
Duvvada, India
psvsvramaraju@gmail.com

2024 First International Conference on Software, Systems and Information Technology (SSITCON) | 979-8-3503-5293-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/SSITCON62437.2024.10797213

***Abstract*—Electricity theft detection (ETD) is crucial for smart grids to maintain their cost-effectiveness. Current approaches for identifying theft struggle with large datasets of electricity use owing to problems with missing values, data variance, and nonlinear patterns in the data; furthermore, an integrated infrastructure to coordinate the study of electricity load data is not currently in place.The suggested method trains models after feature extraction and stage preprocessing. Data normalization, missing value imputation, and data cleansing are all components of data preprocessing. Utility data sets often suffer from inconsistent and missing data as a result of manual and unsynchronized meter reading and registering processes; however, this can be mitigated through the feature extraction process. All the way through training, the proposed approach stuck with the NTM paradigm. The average accuracy of this state-of-the-art approach is 94.22%, which is higher than RNN and ELM.**

Keywords—neural turing machine (NTM), Electricity theft detection, extreme learning machine (ELM).

I. Introduction

Technical losses and non-technical losses (NTLs) are the two main categories of power loss that affect electricity systems. Dissipation of heat during electricity transmission is one source of technical losses, but theft of power by malicious customers is the leading cause of non-technical losses. This results in significant financial losses for the utility companies. An innovative step towards a smart grid, advanced metering infrastructure (AMI) has just been integrated into the power networks. Smart meters installed in the customer's premises are essential for AMI to monitor energy usage and generate bills. This strategy can decrease the prevalence of more conventional types of electrical theft, like tampering with meters and line hooking[1]. Numerous software and hardware-based methods have been suggested in the literature as potential means of detecting cyber-attacks, such as power theft. Presently, data-driven approaches are more attractive than others due to the abundance of data regarding energy usage made available by smart meters.The AMI network relies on smart meters put in customers' homes or places of business to provide consistent reporting on energy consumption. Since unscrupulous consumers can alter their own consumption figures by hacking their smart meters, this data on energy consumption is vulnerable to these attacks. Even though it's not easy to discover these types of breaches, automated systems that detect electricity theft could benefit from customers' extensive energy consumption data. Several automated theft detection approaches have been proposed in the literature[2]. Power plants generate and transmit electricity to consumers across extensive networks, with losses occurring at both ends of the chain. The efficient and safe utilization of energy resources should be a component of every nation's social and economic development goals due to their scarcity and high cost. An alternative to outdated methods of energy monitoring, the smart grid (SG) is designed to withstand the test of time[3]. The most telling features are likely to be rapid drops in power usage (flat spots), unexpected surges during off-peak periods, and unpredictable consumption spikes. These features were most likely identified using feature importance analysis methods, such as feature ranking approaches (e.g., SHAP values or permutation importance) or by analyzing model performance metrics when specific features were removed from the analysis. The second section discusses previous research that has addressed the issue of electricity theft in the literature. The methodology used in this work is detailed in Section III. It begins with an analysis of the dataset and efforts to improve its quality. Then, characteristics are extracted and classified based on an analysis of the customers' load profiles. The results are presented and discussed in Section IV.Section V serves as the paper's conclusion.The requirement to deal with complicated, non-linear correlations between electricity usage patterns and stealing tendencies drives the use of particular approaches, such as neural networks. Neural networks, particularly recurrent designs such as NTMs, are ideally adapted to processing time-series data, which is critical for power usage research. This approach was also chosen for its capacity to learn temporal connections and capture long-term patterns in consumption data. Furthermore, convolutional layers or other techniques may be chosen based on their ability to handle spatial dependencies in time-series data, such as repeating patterns or anomalies across time.

II. Literature Survey

The use of ETD algorithms allows for the detection of power theft incidents. The existing ETD algorithms found in published works are categorized according on whether the system uses unsupervised or supervised learning techniques. Unsupervised learning is a method for teaching algorithms to complete tasks without the use of labels[4]. Then, the

979-8-3503-5293-1/24/$31.00 ©2024 IEEE

proposed approach sends the encrypted aggregated data to the system operator. The data imbalance between regular and unusual users is a big problem that hinders the effectiveness of ML algorithms when it comes to identifying theft[5]. The K-nearest neighbor technique is used to fill in the missing values. A technique that integrates oversampling and under sampling, to rectify the data imbalance [6]. However, the suggested theft detection system's automatic hyper parameter adjustment is not taken into consideration[7]. Using ensemble methods, able to detect cases of power theft in SGs. Light boosting, adaptive boosting, XGBoost, extra trees (ET), random forest (RF), and CatBoost are the top classifiers among the ones mentioned above. When data imbalances occur, SMOTE is the tool to use[8]. Using SMOTE for data balance unfortunately leads to the over fitting problem. For the purpose of detecting power fraud, propose an XGBoost classifier. The power business has used AI in a variety of applications to address real-world problems. When it comes to adding more renewable energy sources to the smart grid, AI really shines because of its ability to optimize electricity price and make it responsive to power supply swings caused by unpredictable weather[9]. In order to lessen the impact of cyber-attacks on power theft, machine learning (ML) has been used in recent years[10]. Evidence of theft in SGs has been detected using DL techniques and supervised and unsupervised ML methods. Over fitting can occur as a result of DL models being trained on a fixed dataset. Thus, kids stop learning to generalize and start learning to recognize specific traits and patterns. Secondly, in order to address changes in consumption patterns and new cyber-attacks, retraining the models with both old and new data is not an appropriate solution[11]. Especially for large datasets, this process is computationally heavy and time-consuming. The use of artificial intelligence to tackle difficult problems has received a lot of attention in recent years in Google's AlphaGO and AlphaGO Zero. Both AlphaGO and AlphaGO Zero demonstrated that reinforcement learning (RL), a type of emerging ML, is similar to human learning in that it can adapt to its surroundings and learn by exploration and exploitation mechanisms [13]. Equally impressive is its ability to mimic an agent and, given limited data, determine the optimal action to take. When it comes to making decisions, RL is quite skilled and possesses many admirable traits [13]. Just like the human brain, RL experiments with its surroundings to find out what works best for making decisions. A number of factors contribute to the suboptimal performance of supervised ML-based ETD models, including consecutive missing values in EC datasets, problems with data class imbalance, incorrect hyper parameter tuning of ML models, and so on[14]. ETD machine learning models can be assessed with complicated performance measures including ROC-AUC, Precision, Accuracy, and Matthew's correlation coefficient (MCC). A plethora of methods for identifying and eradicating power theft via ML-based classifiers have been developed as a result of extensive study in this field; a few relevant works are cited below[15]. A hybrid deep learning (DL) method is used to create a model that can detect instances of electricity theft. Near-miss under-sampling (NM) was used to handle data imbalance, and bee colony optimization was used to tweak the hyper parameters of the adaptive boosting (ADB) classifier[16]. A machine-learning pipeline including SMOTE, KPCA, and SVM is proposed by the authors. An SMOTE-based ETD pipeline to address data class imbalance. Then, it employs kernel function and KPCA to glean additional relevant characteristics from the dataset. Support vector machines (SVMs) exhibit the following characteristics. The power theft detection deep learning model proposed, integrates a convolutional neural network (CNN) with a long short-term memory (LSTM)[17]. Using the local values related to the missing data point, this technique determines the missing instances. The proposed ETD model uses a combination of over-sampling, under-sampling, the kNN imputer, and the SMOTE Tomek algorithm to handle data class imbalance[18]. The proposed approach trained the model using six simulated theft attacks to make sure the data samples were balanced. When the number of theft and non-theft events in an updated dataset are balanced, performance assessments with a decreased percentage of false positives are achieved[19]. The proposed model performed admirably even when the pattern of power use altered as a result of seasonal fluctuations. Geographic data sources and weather data are examples of non-electricity consumption datasets from which model features are derived[20]. The Proposed study recommends developing a very efficient ensemble model using NTM for discovering smart meter power theft incidents. Unlike standard neural networks, TMs feature an external memory component that allows them to store and retrieve information over long sequences. This capacity allows NTMs to handle more complex tasks that necessitate long-term memory retention and retrieval, making them suitable for evaluating time-series data in which patterns do not immediately follow one another but are spaced out over time.A new method for detecting electricity theft utilizing NTM classification and detailed time-domain characteristics is proposed based on the literature. To simplify the training process going forward and make sense of the findings, the proposed approach use Principal Component Analysis (PCA) to conduct classification on a smaller feature space and compare it to the results from classification using all input features. In order to get better performance overall, the proposed system use a Bayesian optimizer to tweak the model's hyperparameters. To get the most out of the model training time while still getting decent results, it employs an NTM to find the ideal ranges for the other critical parameters.The work addresses power theft by creating a detection system using machine learning algorithms and intelligent computing approaches. Traditional solutions, such as manual inspection or basic rule-based systems, can be time-consuming and ineffectual in detecting advanced theft schemes. The method uses patterns in consumption data to detect anomalies and suspicious activities that may indicate theft. It uses data analytics and machine learning models to continuously monitor and analyze electrical use in real time, making it more reliable and scalable than older techniques.

III. Proposed System

Figure 1 depicts each stage, which may involve data collection, pre-processing, feature extraction, model training, and detection. Brief descriptions for each step could include: Data Collection: The long-term collection of electricity usage data from smart meters. Pre-processing involves cleaning the data, filling in missing values, and normalizing the dataset. Feature Extraction: Identifying important patterns, such as peaks in demand or regular drops (flat spots), that are associated with thieving habits. Model training is the process of using labeled historical data to train a machine learning model (such as an NTM). Detection entails using the learned model to real-time or new data to identify probable theft cases. Each stage contributes by gradually

refining the data, preparing it for accurate analysis, and eventually enabling reliable theft detection.

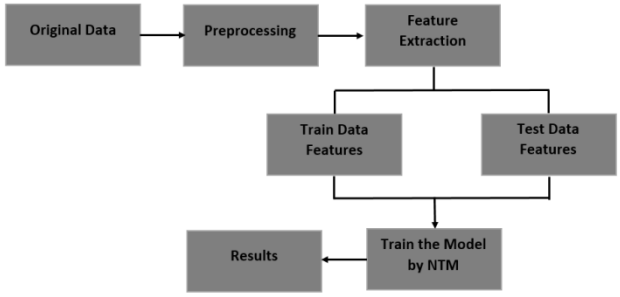

Fig. 1. Flowchat for the Identification of Electricity Theft

A. Data Preprocessing

At time interval p, the energy consumption for an honest customer k and a malevolent customer q are defined as $q_{k,p}$ and $q^*_{k,p}$, respectively. A threshold of ten missing values may be crucial because missing data in power consumption records can affect the detecting process. If there are too many missing values, the dataset may lack the continuity required for accurate detection. Setting a restriction guarantees that the data is resilient enough for accurate analysis, preventing errors caused by large gaps.What follows is a description of the primary methods used for preparing raw electricity data:

1) Data Cleaning: Raw data may contain inaccurate values, also called outliers; these correspond to periods of high electrical demand, such as holidays and special occasions like birthdays and celebrations. High electrical demand periods are described as time periods when electricity consumption exceeds a set threshold, usually during peak usage hours, such as evenings when most households and businesses consume considerable amounts of power. These times are critical for detecting theft since they can disguise unusual consumption patterns. This proposed employs the "three-sigma rule of thumb" to retrieve the outliers using the following formula:

$$Z(q_{k,p}) = \begin{cases} avg(q_{k,p}) + 2\sigma(q_{k,p}) & q_{k,p} > q^*_{k,p} \\ q_{k,p} & else \end{cases} \quad (1)$$

For every time interval that comprises the weekday/time pair every month, $q^*_{k,p}$ is computed using the mean $avg(>$)and standard deviation $\sigma(.)$.

*2) Missing Value Imputation:*Determining values when absent Electricity consumption statistics contain missing values for a variety of causes, such as storage problems and smart meter malfunction. One kind of missing data is continuous missing of many values; when this happens, the solution is to remove users from the system if the number of missing values is more than 10[21]. A second type, which is handled by the formula (2), is missing single data. This is found out by looking at the raw data. Because of this, the proposed approach may recover the following values:

$$Z(q_{k,p}) = \begin{cases} \frac{q_{k,p-1}+q_{k,p+1}}{2} & q_{k,p} \in NaN \\ q_{k,p} & else \end{cases} \quad (2)$$

in where $q_{k,p}$ represents the consumers' electricity usage over a period (e.g., an hour); if $q_{k,p}$ is null, the proposed approach display it as NaN.

*3) Data Normalization:*Data normalization Due to the sensitivity of neural networks, it is necessary to clean up the data before using it.One of the numerous methods employed for this objective is the min-max normalization calculation, which is

$$Z(q_{k,p}) = \frac{q_{k,p} - min(q_{k,p})}{max(q_{k,p}) - min(q_{k,p})} \quad (3)$$

Where $\min(.)$ represents the lowest value and $\max(.)$ represents the highest value for a specific day.

B. Feature Extraction

The primary characteristics of a dataset can be extracted via features extraction. Because of the reduction in processing resources and data dimensionality, the process of implementing algorithms becomes easier. The feature extraction procedure can help improve utility data sets that have inconsistent or missing data due to manual or unsynchronized meter reading and registering. The current analysis extracts a minimum of fourteen new features from each user's consumption data using the R-language packages anomalous and Factoextra. This visual representation shows not just the retrieved features but also their interrelationships. Dark blue represents a very positive correlation between two features, while dark red represents a very negative correlation. As an example, consider the "entropy" and how it is favorably correlated with "lumpiness" and negatively correlated with the "trends". The statistical context and importance of each produced feature are detailed here; however, the two most crucial features, the Canberra distance measure and the flat spots, will be discussed in the parts that follow, as they were extracted for this particular inquiry[22].A "flat spot" is a period of abnormally constant or low power consumption, which may not be typical in normal family or corporate usage. It may indicate manipulation, in which electricity is pulled without being consumed. These flat patches are worrisome, especially when they occur during times of projected power usage unpredictability.As a statistical metric, the Canberra distance measure $EG_{measure}$ is computed using Equation 4.

$$EG_{measure} = g(q,b) = \sum_{k=1}^{t} \frac{|q_k - b_k|}{|q_k| + |b_k|} \quad (4)$$

Given a set of data points, $EG_{measure}$ can spot any noticeable deviations. In order to find outliers in the consumption data, the proposed system calculated $EG_{measure}$ by comparing each consumer's current and past monthly consumption records. The average $EG_{measure}$ score of most consumers falls somewhere between zero and four, indicating that the system are in the healthy consumption range. Having a larger number of $EG_{measure}$, however, is indicative of inconsistent consuming behavior, as demonstrated by their purchasing habits. After going over all the proposed approach have discussed thus far, it is reasonable to conclude that $EG_{measure}$ is an important statistic for evaluating client status and buying behaviors without a tedious computing process[23]. One major problem with this metric is that it doesn't automatically set a cutoff value that distinguishes between two groups. To find flat spots, or fspots, in a time series, one usually finds the greatest run length inside

each equally sized interval in the given sample space. Specifically, this metric can be employed to ascertain the occurrences of low consumption that occur most frequently within a dataset of consumption records. These areas with low usage are usually revealed when consumers hook up directly to the mainline or employ specialized gadgets to interfere with the metering process.

C. Model Training

*1) Neural Turing Machine:*The approach's automation and precision provide the most significant improvement. Using advanced algorithms like as neural networks or particular machine learning approaches, the system can detect tiny patterns of aberrant power use that would otherwise go undetected using traditional methods. Furthermore, technique certainly increases detection accuracy and decreases false positives by incorporating advanced anomaly detection models that learn from prior data to forecast normal consumption patterns. The introduction of Neural Turing Machines (NTMs) or other complicated designs improves the ability to manage sequential data, allowing it to adapt to changing consumption patterns.It was in the book Neural Turing Machines by Ivo Alex Graves, Danihelka, and Greg Wayne when the idea was initially proposed. For its simplicity, the feed-forward neural network was used as the controller. None of the NTM parameter restrictions were updated in our solution. A controller, a memory and an arbitrary number of heads are all up for grabs[24].The controller and memory implementation has remained unchanged from the suggestions made in the previous research. An in-house feed-forward neural network controller was already mentioned. There are four stages to the addressing procedure. Implementation of content addressing and interpolation follows the guidelines laid out in the system. Implementation of convolutional shift only allows us to change one cell to the left or right. When dealing with raw shift addressing weight, the following formula is employed:

$$\sigma = \frac{2}{1+exp(n_p)} - 1 \qquad (5)$$

drawing on the raw addressing vector of the controller and its raw shift weight, n_p. Here is the process for applying the convolutional shift to the addressing vector a^d following interpolation σ:

$$a^n(k) = \left(1 - |\sigma| a^d(k) + a^d(k+e)\right) \qquad (6)$$

where [c] is initialized to 1 when σ is non-zero. Under no circumstances is e equal to 1. σ is a positive integer between 1 and 1. When the value of σ is 1, the addressing vector shifts to the left by one cell. On the other hand, a rightward shift of one cell is carried out when σequals 1. A value of 0 σ does not signify a change. The sharpening scalar δ can be calculated using the sharpening weight $u$and the following formula:

$$\delta = ln(exp(a) + 1) + 1 \qquad (7)$$

It uses the identical formula proposed in the original paper to sharpen it. The proposed learning algorithm of choice for weight optimization was rmsproposed, which involves back-propagation across time. The use of back-propagation with any weight optimization method is made possible by the proposed methodology. The proposed approach have verified that NTM works by applying the suggested learning exercises. Replicable results were obtained on the n-gram, copy, and repeat copy tasks.On top of that, the proposed system developed and evaluated NT Monsequences. The proposed approach put the models through their paces to see how well the system could generalize.

For sequences with $s < 20, s + h < 20$ and up to $s = 60$, the proposed approach conducted tests of NTM. All of the deterministically predictable properties were accurately predicted by the model with this report ratio of s.Any model with a performance greater than 0.34 can generalize to any observed sequence length.Natural language processing (NTM) can learn any sequence. Even if sequences longer than 60 can be learned, the system need to use the minimum training duration. With a modestly sized controller and a little memory, the best generalization was achieved. The results were obtained using a 4×4 memory NTM, a 100-neurons controller, and a single head. The process by which NTM acquires the sequence w^s, y^h, e^{s+h} is readily apparent. The convolutional shift addressing mechanism enables the NTM to shift focus over its memory in a continuous, differentiable manner, rather than making discrete jumps. This allows the model to capture gradual changes in the power consumption pattern, improving its ability to track slight but significant shifts, which might signal theft.Adding extra memory cells is unneeded for an NTM because it learns to utilize just two cells and ignore the others or to use all of the memory as a single cell. In the former case, the other elements of the head addressing vector are zero during the full sequence. Raising the capacity of a memory cell does nothing to improve the precision of generalizations; all it does is change the internal representation of the character count. The capacity of an NTM to store memory over longer periods can aid in the discovery of subtle patterns or recurring anomalies that regular neural networks may overlook. NTMs provide more dynamic analysis of time-based data, which is crucial in detecting theft that occurs sporadically or intermittently.

IV. Result and Discussion

Utilities are facing challenges in properly meeting customer demand for electricity due to an increase of power thieves. Present approaches fail to accurately detect energy theft (ETD) because to their high false positive rate (FPR), propensity for overfitting issues, and inaccurate classification of imbalanced power consumption data. Power theft has cost utilities a lot of money, thus finding the perpetrators of this crime will need further research.

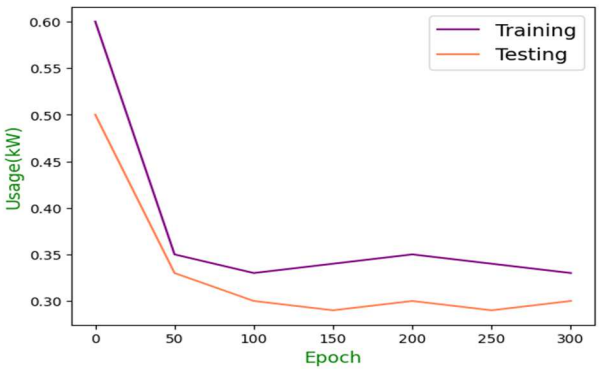

Fig. 2. Loss Curve of Training and Testing

Data on normalized consumption is used to base the experiment. When using a batch size of 100, the training process ends after 300 epochs. Training and testing losses leveled out gradually, as seen in Figure 2.

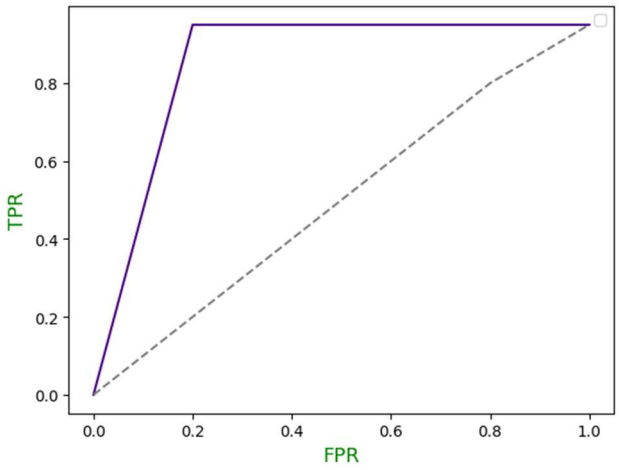

Fig. 3. ROC Curve of NTM Models

When compared to the baseline model, the computed AUC value for the NTM is 0.95, indicating significant improvement. Figure 3 shows that the suggested approach can accurately classify both classes.

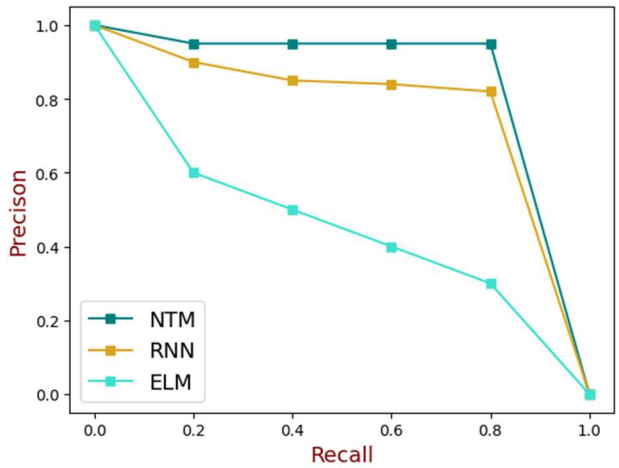

Fig. 4. PR Curve of the Different Model

The PR curve is shown in Figure 4. When compared to the other benchmarks, the PR curve makes it quite evident that the various models perform far better.

TABLE I. PERFORMANCE EVALUATION(%)

Models	Precision	Recall	F1-Score	MAP
ELM	0.8716	0.8541	0.8631	0.8768
RNN	0.8964	0.8734	87.63	0.8991
NTM	0.9219	0.9040	91.24	0.9260

Table I also displays comparisons based on recall, F1-score, mean average precision (MAP), and precision. The results of table I show that the suggested technique outperforms the others on MAPscore.

Figure 5 shows that when comparing the accuracy, recall, precision, F1score, and MAP of the many offered ways, the suggested strategy clearly performs better, demonstrating its efficacy and importance. On the same performance evaluation measures, the proposed approach also compare the suggested strategy to some of the more well-known conventional approaches. Convolutional shift addressing enhances model performance by allowing it to better manage the trade-off between short-term and long-term dependencies in data. This method allows for smoother transitions while focused on different areas of the input data, improving the accuracy of the theft detection system by spotting nuanced trends over time.

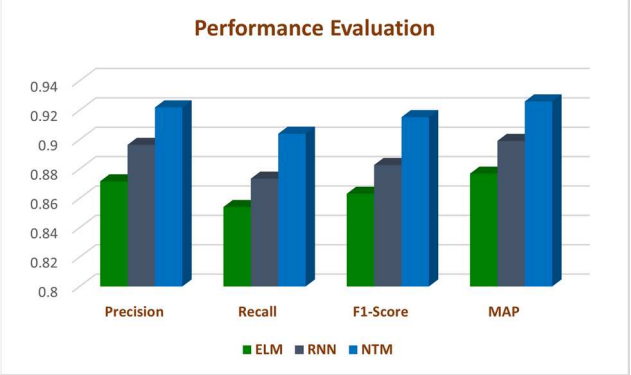

Fig. 5. Performance Evaluation of the Model

V. CONCLUSION

Electricity theft is a global pandemic that affects utilities and power customers equally. Electric dangers are on the rise, utility firm economic development is halted, and energy costs for customers are driven up. One important part of building smart grids is collecting massive amounts of data, like customer consumption data, which may be utilized to identify instances of power theft through the application of deep learning and machine learning algorithms. Data preparation includes data cleansing, missing value imputation, and data standardization. Because utility meter reading and registration are typically done manually and out of sync, utility data sets are prone to missing and inconsistent data. Fortunately, feature extraction can help with this. For the model's training, the NTM is sufficient. The proposed approach found that our suggested model achieved an average accuracy of 94.22% when compared to the RNN and ELM models.

REFERENCES

[1] M. Nabil, M. Ismail, M. Mahmoud, M. Shahin, and K. Qaraqe, "Deep Recurrent Electricity Theft Detection in AMI Networks with Random Tuning of Hyper-parameters," 2018 24th Int. Conf. Pattern Recognit., pp. 740–745, 2018.

[2] M. Nabil, M. Mahmoud, M. Ismail, and E. Serpedin, "Deep recurrent electricity theft detection in ami networks with evolutionary hyper-parameter tuning," in Proceedings - 2019 IEEE International Congress on Cybermatics: 12th IEEE International Conference on Internet of Things, 15th IEEE International Conference on Green Computing and Communications, 12th IEEE International Conference on Cyber, Physical and So, 2019, no. November, pp. 1002–1008. doi: 10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00175.

[3] A. A. Almazroi and N. Ayub, "A Novel Method CNN-LSTM Ensembler Based on Black Widow and Blue Monkey Optimizer for Electricity Theft Detection," IEEE Access, vol. 9, pp. 141154–141166, 2021, doi: 10.1109/ACCESS.2021.3119575.

[4] I. U. Khan, N. Javaid, C. J. Taylor, and X. Ma, "Robust Data Driven Analysis for Electricity Theft Attack-Resilient Power Grid," IEEE Trans. Power Syst., vol. 38, no. 1, pp. 537–548, 2023, doi: 10.1109/TPWRS.2022.3162391.

[5] R. Xia, Y. Gao, Y. Zhu, D. Gu, and J. Wang, "An attention-based wide and deep CNN with dilated convolutions for detecting electricity theft considering imbalanced data," Electr. Power Syst. Res., vol. 214, no. PA, p. 108886, 2023, doi: 10.1016/j.epsr.2022.108886.

[6] M. I. Ibrahem et al., “Efficient Privacy-Preserving Electricity Theft Detection with Dynamic Billing and Load Monitoring for AMI Networks,” pp. 1–14, 2020.

[7] S. Hussain, M. Wazir, T. A. Jumani, and S. Khan, “A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection,” Energy Reports, vol. 7, pp. 4425–4436, 2021, doi: 10.1016/j.egyr.2021.07.008.

[8] D. Yao, M. Wen, and X. Liang, “Energy Theft Detection With Energy Privacy Preservation in the Smart Grid,” IEEE Internet Things J., vol. 6, no. 5, pp. 7659–7669, 2019, doi: 10.1109/JIOT.2019.2903312.

[9] M. Nabil, M. Ismail, S. Member, and M. M. E. A. Mahmoud, “PPETD : Privacy-Preserving Electricity Theft Detection Scheme With Load Monitoring and Billing for AMI Networks,” IEEE Access, vol. 7, 2019.

[10] M. M. Badr, M. I. Ibrahem, M. Baza, M. Mahmoud, and W. Alasmary, “Detecting Electricity Fraud in the Net-Metering System Using Deep Learning,” Res. Gate, no. July 2022, 2021, doi: 10.1109/ISNCC52172.2021.9615628.

[11] N. Javaid and S. Member, “A PLSTM , AlexNet and ESNN Based Ensemble Learning Model for Detecting Electricity Theft in Smart Grids,” IEEE Access, vol. 9, pp. 162935–162950, 2021, doi: 10.1109/ACCESS.2021.3134754.

[12] W. Eberle, “Smart Grid Energy Fraud Detection Using Artificial Neural Networks,” 2014 IEEE Symp. Comput. Intell. Appl. smart grid, no. December, 2014, doi: 10.1109/CIASG.2014.7011557.

[13] Z. Zheng, Y. Yang, X. Niu, H. N. Dai, and Y. Zhou, “Wide and Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids,” IEEE Trans. Ind. Informatics, vol. 14, no. 4, pp. 1606–1615, 2018, doi: 10.1109/TII.2017.2785963.

[14] M. Adil, N. Javaid, U. Qasim, I. Ullah, M. Shafiq, and J. Choi, “LSTM and Bat-Based RUSBoost Approach for Electricity Theft Detection,” Appl. Sci., vol. 10, no. 12, p. 4378, 2020, doi: 10.3390/app10124378.

[15] M. Anwar and N. Javaid, “Electricity Theft Detection using Pipeline in Machine Learning,” 2020 Int. Wirel. Commun. Mob. Comput., no. May, 2020, doi: 10.1109/IWCMC48107.2020.9148453.

[16] A. Arif and N. Alrajeh, “Big data analytics for identifying electricity theft using machine learning approaches in microgrids for smart communities,” Concurr. Comput. Pract. Exp., vol. 33, no. 17, 2022, doi: 10.1002/cpe.6316.

[17] S. K. Gunturi and D. Sarkar, “Ensemble machine learning models for the detection of energy theft,” Electr. Power Syst. Res., vol. 192, no. October, 2021, doi: 10.1016/j.epsr.2020.106904.

[18] X. Kong, X. Zhao, C. Liu, Q. Li, D. L. Dong, and Y. Li, “Electricity theft detection in low-voltage stations based on similarity measure and DT-KSVM,” Int. J. Electr. Power Energy Syst., vol. 125, no. February 2020, 2021, doi: 10.1016/j.ijepes.2020.106544.

[19] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. Shen, “Energy-theft detection issues for advanced metering infrastructure in smart grid,” Tsinghua Sci. Technol., vol. 19, no. 2, pp. 105–120, 2014, doi: 10.1109/TST.2014.6787363.

[20] S. C. Yip, K. S. Wong, W. P. Hew, M. T. Gan, R. C. W. Phan, and S. W. Tan, “Detection of energy theft and defective smart meters in smart grids using linear regression,” Int. J. Electr. Power Energy Syst., vol. 91, pp. 230–240, 2017, doi: 10.1016/j.ijepes.2017.04.005.

[21] S. Li, Y. Han, X. Yao, S. Yingchen, J. Wang, and Q. Zhao, “Electricity Theft Detection in Power Grids with Deep Learning and Random Forests,” J. Electr. Comput. Eng., vol. 2019, 2019, doi: 10.1155/2019/4136874.

[22] S. Hussain, M. W. Mustafa, T. A. Jumani, S. K. Baloch, and M. S. Saeed, “A novel unsupervised feature-based approach for electricity theft detection using robust PCA and outlier removal clustering algorithm,” Int. Trans. Electr. Energy Syst., vol. 30, no. 11, 2020, doi: 10.1002/2050-7038.12572.

[23] L. J. Lepolesa, S. Achari, and L. Cheng, “Electricity Theft Detection in Smart Grids Based on Deep Neural Network,” IEEE Access, vol. 10, pp. 39638–39655, 2022, doi: 10.1109/ACCESS.2022.3166146.

[24] J. Tkačík and P. Kordík, “Neural Turing Machine for sequential learning of human mobility patterns,” Proc. Int. Jt. Conf. Neural Networks, vol. 2016-Octob, no. July 2016, pp. 2790–2797, 2016, doi: 10.1109/IJCNN.2016.7727551.