International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An Efficient Power Theft Detection Using Modified Deep Artificial Neural Network (MDANN)

G.P. Dimf[1*], P. Kumar [2], and V.N. Manju[3]

Submitted: 05/10/2022 **Revised:** 20/12/2022 **Accepted:** 30/12/2022

Abstract*:* Electricity theft becomes a major concern for utilities in this new era of high tech, self-sufficient dwellings. Finding and reducing energy losses or theft has proven challenging due to insufficient inspection methods. In terms of energy, both technical and non-technical losses (NTL) are included in distribution. Energy theft is a significant factor in NTL that can strain the finances of service providers. Wireless data transmission is used in modern smart metres. It follows that hi-tech dwellings can be easily hacked to steal power. Many new technologies have been implemented into Advance Metering Infrastructure (AMI) to combat energy theft. It is necessary to derive the consumption pattern in order to identify illegal energy customers. Using data mining methods, a computational system is designed for examining and identifying energy consumption patterns. Through the use of machine learning, we are able to improve our customers' energy consumption statistics and provide them with early warning of any irregularities. Multiple supervised learning techniques are examined and contrasted in relation to their predictive accuracy, recall, precision, AUC as well as F1 score. These include the decision tree (DT), ANN, Deep ANN, Modified ANN and AdaBoost. Based on the results of the study, MDANN is superior to alternative classifiers for supervised learning including ANN AdaBoost as well as DT according to recall, F1 Score along with AUC. The upcoming research should focus on testing different supervised learning algorithms using various datasets and including appropriate pre-processing procedures to boost performance.

Keywords: *include power theft detection; deep learning; smart grid; accuracy; recall; precision; area under the curve (AUC); F1 score;*

1. Introduction

Power companies around the world face a big difficulty in the form of energy waste during electricity's distribution and transmission. The loss of energy are often classified as either nontechnical losses (NTLs) or technical losses (TLs) [1, 2]. A term "net total losses" (NTL) refers to the sum of all losses minus all theft losses (TLs). In fact, the vast majority of power theft [3] is the result of physical operations like tapping lines, damaging metres and tampering with metre readings. Potential revenue losses for utilities may occur from fraudulent electricity practises. About $5.5 billion in ann[1]ual losses are attributed to electricity theft alone. In many developing nations, electricity theft is seen as a major barrier for economic development. One recent study, for instance, found that fraud and theft account for nearly 20% of total electricity produced in India [38]. Concerning circumstances exist in numerous Asian and African nations, including, but not limited to, Nepal, Pakistan, Lebanon, Kenya, Tanzania and Uganda. Not only emerging markets are affected by the issue. [9] Energy theft is thought to be responsible for about $6 billion in financial losses annually in the United States alone, or about 80% of all commercial losses.

American citizens [4]. The US. Additionally, electricity theft costs utilities upwards than $25 billion annually [5]. Also, there is concern that electricity stealing habits could compromise the safety of the grid. Electricity theft, which can lead to overloaded electrical networks, poses a threat to public safety in several ways. For this reason, it is crucial for the security and reliability of the power grid that electricity theft be detected accurately. Power companies were able to acquire massive amounts of frequent electricity usage data from smart metres thanks to an advanced metering infrastructure (AMI) in smart grids [6, 7]. While there is always two sides to a story, it's important to note that the AMI network opens up fresh avenues for electricity theft. Various ways, including digital tools and cyber-attacks, can be used to launch such attacks on the AMI. Human inspection of

[1]*Research Scholar, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, 627012, Tamil Nadu, India.*
[2]*Associate Professor, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, 627012, Tamil Nadu, India.*
[3]*Assistant Professor, Department of CSE, CMR Institute of Technology, Bengaluru, India.*
[*]*Corresponding Author Email: dimfgp02@gmail.com*

unauthorised line diversions, comparison of malicious metre records with benign records and inspection of malfunctioning equipment or hardware are the fundamental techniques for identifying power theft. However, these steps take a long time and cost a lot of money to complete during the entire metre verification process for a system.

Not even resorting to manual processes can safeguard against cybercrime. In recent years, many different approaches have been offered for overcoming the aforementioned problems. You can broadly categorise these approaches as either game theory, state or artificial intelligence (AI) based models [8].

In [12, 17] multiple deep learning architectures are assessed for detecting electricity theft. These designs include convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short term memory (LSTM) and stacking auto encoders. Conversely, functionality of the detectors are evaluated utilising simulated data that prevents an accurate comparison with shallow systems [39]. Moreover, a customer specific deep neural network (DNN) detector was proposed by the authors of [14, 19], which might be used to successfully thwart such cyber-attacks. Recently, CNN has been employed for a number of purposes [20-22] owing to its ability for extracting beneficial as well as distinguishing features from raw data. For the purpose of detecting power theft, these usage cases utilize CNN extraction of features over high resolution smart metre data. Smart grid electricity theft was analysed using a large as well as deep convolutional neural network (DCNN) framework developed in [12].

The conventional methods for detecting theft of electricity primarily depend heavily on the project schedule for technicians, who are employed for power distribution companies. A process involves reading an electricity metre, followed by manual analysis and calculation, recording, counting and analysis [21]. It is possible to prevent energy theft on a hardware level through a variety of actions, including the installation of a specialised watt hour metering box, connection of a conductor to reduced voltage outlet as well as its closure to metering device, addition of an anti-thief function to watt hour metre [30] along with an increase in electrical acquisition system's application rate. On the other hand, the majority of such conventional anti-theft detection techniques concentrate on power device advancement. [26, 27] It is challenging for identifying power consumption traits of power stealing users as well as detecting stealing of electricity behaviour carried out via sophisticated assault means because there aren't enough anti-power stealing methods for analysing large historical data sets on power consumption [31, 32]. The advancement of novel information, automation and AI technologies must therefore be bolstered for supporting the growth of power industry. Studying an intelligent anti power theft approach based on big data of power consumption for identifying power theft behaviour is of great engineering significance with a continuous enhancement of dynamic monitoring as well as acquisition technology of energy consumption data for power grid consumers [33, 34].

For effectively identify electricity theft, it has to be expected that MDANN will automatically collect several aspects of clients' consumption patterns using smart metre data. The RF is used in place of the MDANN classifier, which identifies consumer patterns using extracted attributes, to improve detection performance. All of the electricity customers in Tirunelveli and Tuticorin were used in training and validation of the suggested framework.

2. The Literature Survey

Bhat et al. compared CNN, RNN, LSTM [36] as well as loaded auto encoders (LAEs) for the purpose of detecting electricity theft. However, the performance of the detectors in shallow architectures was tested using generated data, which is not a credible metric [23]. An elaborate CNN model ETD was introduced for SGs by Zheng et al. They found that the majority of existing methods are not very effective at detecting electricity theft because they failed of preventing the periodicity of electrical consumption and also rely on one dimensional (1-D) statistics of electricity usage. [24].

To alleviate these problems, this work aims to develop a reliable ETD system. Specifically, we introduce a model for detecting electricity thieves using CNNs and a Meta heuristic optimization approach [25, 37] that takes cues from nature. There are a number of convolutional, pooling and fully linked layers, which make up CNN a part of the system. In particular, CNN is well suited to recording the regularity of information like electricity usage. This is the first study that we are aware of that proposes and uses deep algorithm model (combining CNN with MDANN algorithm) to investigate electricity theft in smart networks. Additionally, a vast accurate dataset of energy use was subjected to rigorous testing.

3. Supervised Learning Algorithms

Each supervised learning method for detecting electricity theft is broken down here, along with the underlying theory and equation. The fundamental steps of supervised learning algorithms are depicted in Figure 1. For constructing prediction model, ML algorithm initially employs training data, feature vectors as well as label data as inputs.

3.1. DT

As a supervised learning technique, Decision Tree (DT) has the capacity of solving both classification as well as regression issues [12]. This article discusses the Decision Tree (DT) [13], a method for categorising instances based on the values of their attributes. Each node in the DT algorithm's tree stands in for a different quality of an instance that needs to be classified

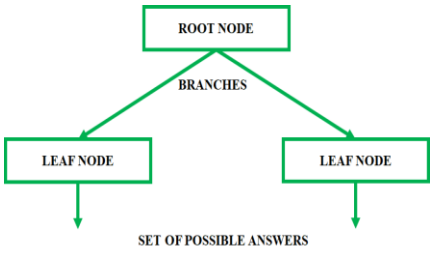

Fig.1. The Decision Tree

In addition, the tree presupposes that each node represents a numerical value. Figure 1 is an easy to understand example of DT.

3.2. ANN

A single input layer, one or more hidden layers and one or more output layers are shown in Fig. 2 as the three layers that comprise an ANN's architecture [10]. It's worth noting that a multilayer feed forward neural network or multilayer perceptron are other names for ANN. The brain's network of neurons served as inspiration for this approach. Dendrites (which are found in the human brain) are used to represent the inputs of an ANN as they transfer electrochemical signals received from neurons to the cell body. There is a weight associated with each input that determines which hidden layer it transmits signals to. Often, the sigmoid function, an activation function, is what drives a neuron. In addition to the step function, the Gaussian function and a linear function, the hyperbolic tangent function is also mentioned as a possible activation function [16].

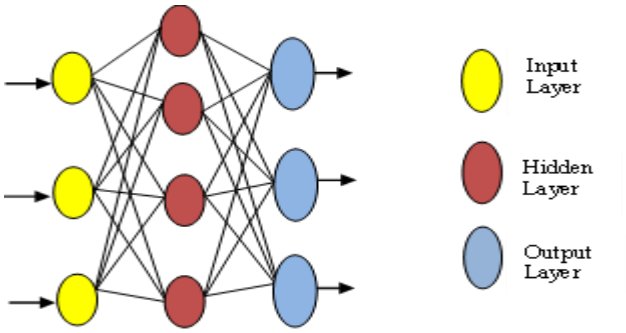

Fig.2. The structure of an ANN

An ANN's axon, which connects two neurons and continues to the synapse, is its very last layer. Typical architectures for ANNs have one output, two inputs and a hidden layer. Neural Network (NN) periods are defined by the neurons' bidirectional traffic between input and output. (The optimal epoch for NN training is determined by the amount of error that can be tolerated during training and the ANN output equation looks like this:

$$y_i = \Psi \sum_{i-0}^{n} \omega ji\, xi + \theta i$$

3.3. Deep Artificial Neural Network (DANN)

Artificial neural networks with two or several hidden layers are recognised as deep neural networks (DNN) [13]. It takes a lot of power and data for deep learning (DL) for capturing a lot of information from original input data of additional layers of a neural network (NN). The terms "deep learning" and "deep neural network" are synonymous; however, "deep artificial neural network" is a somewhat different phrase (DANN). DANN's several layers allow it to categorise attributes in their various manifestations. Understanding the myriad ways in which data from lower levels is combined to yield higher level features is the key to unlocking all of these layers. See Figure 3 for a visual representation of the MDANN structural design.

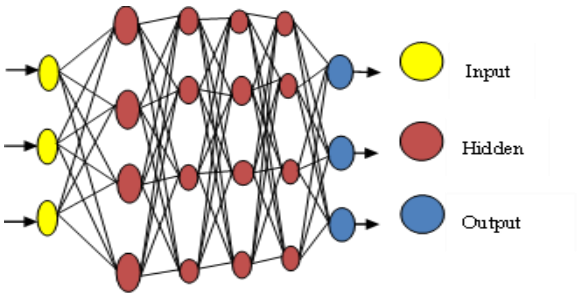

Fig.3. DANN model

3.4. Adaboost

AdaBoost, a method of ensemble learning, as presented by Freund and Schapiro [11]. It excelled in classifying data to an unprecedented degree. To a greater extent than other learning approaches, AdaBoost can handle complex prediction problems without becoming overly specialised. The strategy aids the development of incompetent learners by keeping a set of weights from the training dataset. Then, after every cycle of ineffective learning, it will make an adaptive adjustment to the students' performance. Misclassified weights tend to grow and correctly classified weights tend to decrease in the training dataset [8].

4. Existing System

Prior efforts to identify electricity thieves relied on customers' power use profiles. The spot where electricity is not being properly billed for [35]. Every single customer in that area is being treated as a possible criminal. One drawback of the preceding efforts is that they have been based on the unproven supposition that certain consumers are engaging in fraudulent activity

when it comes to stealing electricity [28, 29]. Possibilitative clients might be exposed as scammers in this case. The motivation for this study comes from the need to combine the real energy consumption statistics with the hypothetical data based on our clients' hypothetical behaviour. Data clusters are analysed with machine learning techniques, while clients are categorised with deep learning. Based on how often and where a customer's information is used, we can tell if it's legitimate or fake.

4.1. System Modeling

This new approach to preventing energy theft in smart homes is both cutting edge and effective. The energy monitoring device was really put to use in a real home because of a non-invasive method of data collection. Time arrangement details and power consumption in an uncontrolled home environment are among the data sets collected. Keen Houses are built by integrating IoT and smart metres. Advanced Metering Infrastructure monitoring and regulation (AMI),

Reconciling the underlying foundation was the primary goal of the Energy Management System (EMS). The DSMS (Request Side Management System) is built into the Enterprise Management System (EMS). As such, its primary value is in the administration of interest responses and commitments. It gathers demand information to guide load-shifting and other forms of optimal force application that maximise profits from the power market at both peak and off-peak times. Within the safety of the user's home, they can use their smartphones to manage their IoT devices. Data generated from smart homes could be analysed using newly discovered and designed frameworks to inform judgments that are both smart and energy-efficient. Energy theft is problematic in the dense matrix network. Damages in the billions of dollars range have been incurred by several countries.

5. Methodology

Theft of electrical power occurs when someone uses electricity without paying the appropriate tax on the amount of power used. The suggested framework has been depicted in Figure 4.

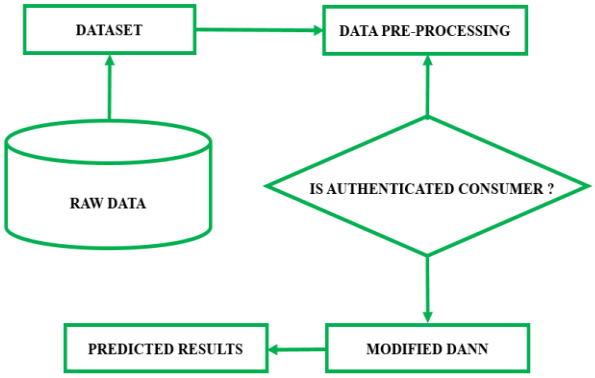

Fig.4. Suggested Model for Identifying Electrical Theft

5.1. Data Collection

For this study, data was gathered twice. To train the intelligent system to recognise normal usage patterns from those that might be cause for concern, initial data collecting focused on billing and consumption information from past clients. In the second phase of data collection, the same kinds of details were acquired as in the first, but they were used to put a fraud detection system through its paces, looking for and identifying potentially suspicious customers.

There are two types of customer information gathered:

- The data from the Enhanced Customer Information Billing System (e-CIBS)
- Information with a Very High Potential for Harm.

5.2. Information on Power Useage

Data sets of actual customer power consumption from State Grid Corporation of China are used for the study (SGCC). Totalling 42,372 rows and 1,035 columns, this data set is quite substantial. The customer's ID is in the first column, the "Flag" prediction indicator is in the second and the third through the last columns are the days of the week (1,035). The Metadata types in the dataset consist of a mishmash of alphabetic characters, numeric values and blanks (NaN). The amounts of power (electricity signals) used by each consumer over a two year period are represented numerically, along with any missing or inaccurate information.

Additionally, the flag column's metadata is (zero and one) as well as refers to the type of consumers (normal or thief), whereas a number of zeroes in "Flag" column denotes the number of typical power consumers and sum of zeroes represents a total number of normal electricity users (38,757). There are two thieves, represented by the character one in "Flag" column as well as one flag, shown by the number one in that column (3,615). Ultimately, it shows that a number (42,372) signifies information about users' power consumption patterns over the past 1,035 days (from Jan. 1, 2018 to Oct. 31, 2020).

5.3. Modifying the Database

A provided power use dataset has been altered at multiple phases using numerous methods for use in developing ETD blueprints. Since the neural network only allows numerals and these values aren't specified, a new dataset has been generated via substituting all the null and non-values in the prior dataset using zeros until the neural network recognises them. Next, a new dataset is divided into training set (consisting of 80% of the data) and testing set, which comprises 20% of data using 80/20 rule.

5.4. Customer Selection and Filtering

Only customers with extensive and relevant data were selected from the UCI repository for the building of the MLP model because the e-CIBS data collected from TANGEDCO is in raw format and therefore requires pre-processing to extract important and meaningful information. Since the information was collected in a database, we employed Structured Query Language (SQL) data mining methods to ensure that we met the following four requirements:

- It is recommended that monthly data be cleaned of repeat customers.
- Customers who use no energy at all over the course of a year should be cut off.
- Customers that aren't there during the whole year should be removed.

5.5. Data Preprocessing

Metrics and user profiles are aggregated for a certain area. Every single customer in that area has been authentically profiled. The percentage of reliable customers has been calculated employing a Deep Learning system. After splitting the data in two ways, the deep neural network is utilised to determine which customer profile displays the more realistic pattern of energy use. According to the data from their smart metres, the customers are separated into distinct groups. The readings from a representative customer's smart metre are used here; they are taken every 30 minutes and expressed in kilowatt hours over the course of 28 days.

The normal consumers are the negative class and the out of the ordinary ones are the positive class. Data extracted from the confusion matrix includes the following:

- TP: Anomalies in consumer behaviour are correctly predicted as outliers.
- TN: Normal consumers are expected to be normal with high confidence.
- FP: An ordinary shopper is viewed as suspicious.
- FN: Consumer Anomaly Considered Typical

5.6 Theft detection using Modified Deep Artificial Neural Network

The normal multi-layered neural networks, which are also called modified deep artificial neural networks (DANN), consist of input layers, hidden layers and output layers. The discrete convolution is the key operation in convolutional layers. We use a 2 × 2 kernel as an example to illustrate the discrete convolution. The input I has a value in each grid. Then, a two-dimensional kernel function K ∈ R2×2 is used to extract features. The output S of the convolution is:

$$s(i,j) = \sum_{k_i=0}^{1} \sum_{k_{j=0}}^{1} 1(i + k_i, j + k_j) k(k_i, k_j) \qquad (1)$$

Above Equation (1) illustrate that convolutional kernels map the neighbouring information of the input into the output

We construct the power consumption recordings matrix y and the electrical thefts matrix nu assuming that there are xn historic electricity usage records and that each record has T interval. Meanwhile, assume that xr's influencing external elements are absent. The task of detecting power theft is classified as a discrete two-class task, and each power consumption record (x1, x3) should be placed into one of the specified classes (abnormal or normal), as shown below.

$$y_1 = \begin{cases} 1 & if\ record\ is\ abnormal; \\ 0 & otherwise \end{cases} \qquad (2)$$

In the following paper, we examine whether power consumption records x have displayed aberrant behaviours in the past by using these records along with external information sequences such as days and types as time series xb, f, respectively.

5.7. Performance Metrics

5.7.1. Accuracy

The accuracy of classifiers is measured as the percentage of correctly labelled examples relative to a total number of examples. This is a format used for the calculation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3)$$

Here TP, TN, FP and FN symbolises "True Positive," "True Negative," “False Positive” as well as "False Negative”.

5.7.2 Precision

Equation (4) defines accuracy as the fraction of correctly labelled positive classes (TP) relative to the full set of positive classes (TP + FP). It was hypothesised that the FP rate was low if the value was extremely precise.

$$PRECISION = \frac{TP}{TP + FP} \quad (4)$$

5.7.3. Recall

In statistics, recall refers to the rate at which positive class (TP) observations are properly labelled as such, as a percentage of all class observations (TP + FN).

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

5.7.4. F1-Score

For classes with an uneven distribution of members, the F1 score (F-measure) is preferable since it considers a recall and precision weighted average (as demonstrated in equation (6). From a scale of 0 (worst) to 1 (best), this is the estimated value (the best). It is advised to keep an eye on both recall and precision measurements if it is found that the classes are significantly imbalanced. F1 score, on the other hand, combines the two into a single measure that is more appropriate for assessing the given form of data set:

$$F1 - Score = \frac{2 * Recall * Precision}{Recall + Precision} \quad (6)$$

6. Results and Discussion

The results of every type of comparison strategy are presented here. As instructed, training set and testing set were created from the dataset. As, ratios has been modified in the algorithm, so did algorithm's recall, accuracy, AUC, precision and F1 score. AUC along with F1 score were utilized owing to dataset's uneven classifications. Tabulated in Table 1 below are the results of several evaluation procedures used to assess various comparison methods. We bolded the top performers in the table. On average, ANN performed at 93.78%, followed by MDANN (94.6%), Ada Boost (92.86%) and DT (93.3%). Both DT and MDANN achieved 92.83% and 93.58% accuracy with respect to the 60/40 split percentage. As an added bonus, accuracy was maximised using ANN and AdaBoost for 60/40 and 70/30 ratios, respectively.

The AUC was more than 0.5 for every classification ratio. These classifiers proved to be effective in classification problems. Using a 70/30 split percentage, MDANN achieved a higher AUC (10.95) than DT (0.6524) and AdaBoost (0.6875). Related to the AUC was another measure of quality, the F1 score. The highest F1 scores can correspond to the highest AUC values. Table 3 reveals that MDANN, DANN and AdaBoost performed best in terms of AUC at the 70/30 split percentage, with values of 11.03, 0.9211 and 0.6875, respectively. F1 score also yielded the highest success rates (56.04, 55.02 and 16.58 percent).

To determine how widely the outlier group was represented during the classification process, precision evaluation was used in this research. The experiments showed that a 90/10 split percentage ratio yielded the best results for three classifiers (ANN, MDANN and AdaBoost), with 58.33%, 60.40% and 58.46% precision, respectively. DT's accuracy at the 70/30 split percentage was the highest at 54.99%. The overall average precision was 66.73 percent, with ANN coming out on top. MDANN had the highest average recall at 42.97 percent, followed by ANN at 51.83 percent, AdaBoost at 8.82 percent and DT at 0.00 percent (3.9 percent). As expected, ANN and MDANN outperformed other methods in terms of recall, with 51.83 and 63.0 percent, respectively, for a 60/40 split. Despite having 80 and 20 samples for training and testing, respectively, DT, another classifier, had the highest recall (5.49%). AdaBoost's recall rate was best when the ratio was set to 70/30. The average accuracy and precision produced by ANN was 92.54% and 64.05%, respectively, however the average results for the other three assessment metrics (recall, F1 score and AUC) were 42.97%, 47.98% and 10.95%, respectively, for MDANN. In conclusion, the splitting% or the ratio between training and testing had a significant part in achieving an optimum outcome.

7. Assessment and Discussion Based On Comparisons

Several measures of performance including precision, accuracy, recall, AUC as well as F1 score are depicted in Figures 6(a)-6(e) for DT, DANN, ANN, MDANN and AdaBoost. Theoretically, the model was fitted using a training dataset, while its efficacy was measured using the testing dataset. The percentage of the dataset to be split was chosen so that its application and performance on fresh data could be evaluated. With a 70/30 splitting %, MDANN outperformed AdaBoost, ANN and DT in terms of training and prediction model accuracy (Figure 6). (a). It's also worth noting that when the splitting proportion was 80/20 rather than 70/30, the trained models' performance vastly improved in both AdaBoost and DT. Most classifiers' efficacy plummeted when the split% reached 90/10, with AdaBoost being an exception.

Table 1. We can see a comparison of the supervised learning outcomes.

Techniques	T :T	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	AUC
DT	90/10	92.32	48.39	3.51	5.68	0.6563
	80/20	92.34	51.79	5.49	9.18	0.6655
	70/30	93.01	54.99	4.36	7.23	0.6606
	60/40	92.83	49.86	2.59	3.99	0.6524
Average		92.62	51.26	3.99	6.52	0.6587
ANN	90/10	93.78	80.05	15.04	24.8	0.7136
	80/20	93.68	63.68	34.81	44.95	0.8045
	70/30	93.66	58.23	44.74	50.63	0.8486
	60/40	93.98	58.33	51.83	54.94	0.882
Average	-	93.98	65.07	36.61	43.83	0.8121
DANN	90/10	93.71	66.73	20.72	31.33	0.7391
	80/20	93.68	61.93	39.13	47.93	0.824
	70/30	94.28	63.66	46.22	53.57	0.8587
	60/40	92.53	49.26	62.16	66.02	0.9211
Average	-	93.55	60.40	42.06	46.96	0.8357
Ada Boost	90/10	92.91	64.66	8.79	14.82	0.682
	80/20	92.78	56.15	7.05	11.84	0.6704
	70/30	93.4	56.51	10.09	16.58	0.6875
	60/40	92.86	56.52	8.82	14.65	0.6813
Average	-	92.99	58.46	8.69	14.47	0.6875
MDANN proposed	90/10	94.76	67.76	21.63	32.35	10.85
	80/20	94.73	62.95	40.04	48.95	10.93
	70/30	95.33	64.68	47.13	54.59	10.97
	60/40	93.58	50.28	63.06	56.04	11.03
Average	-	94.6	61.42	42.97	47.98	10.95

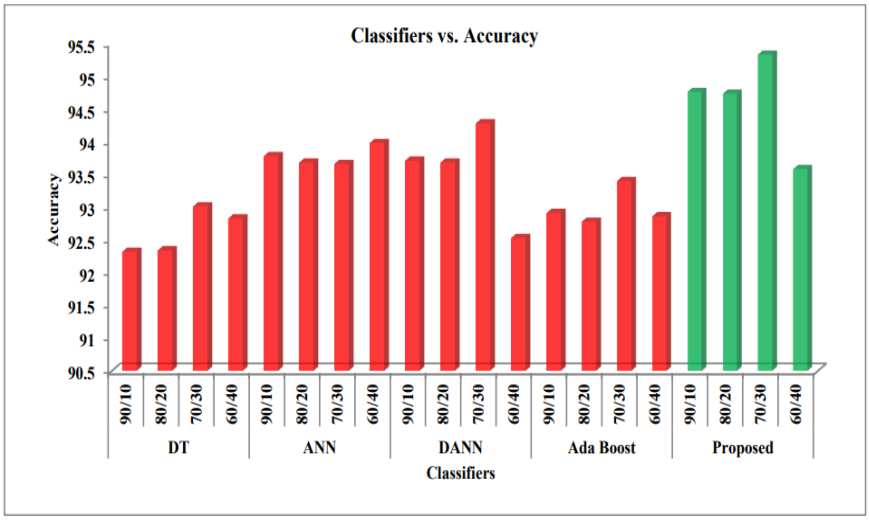

Fig. 5(a). Comparison of various Classification method vs Accuracy

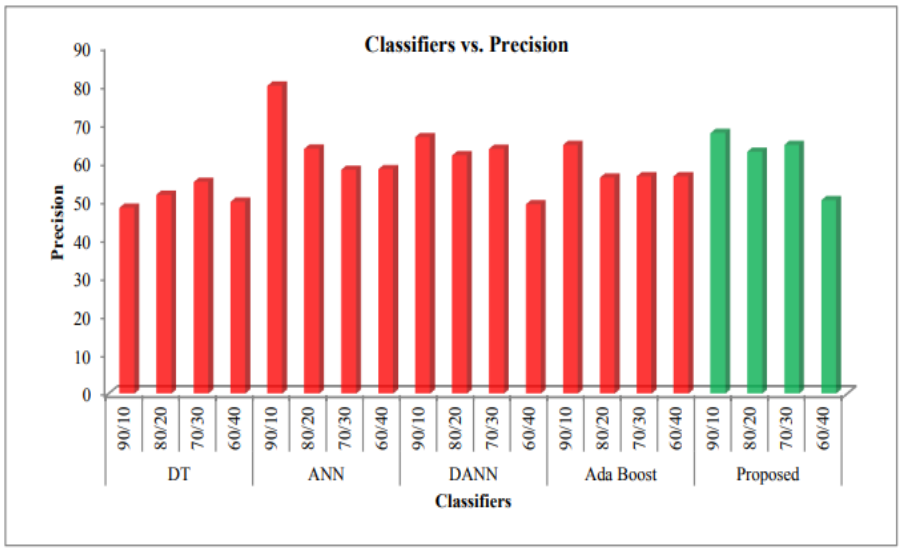

Fig. 5(b). Comparison of various Classification method with precision

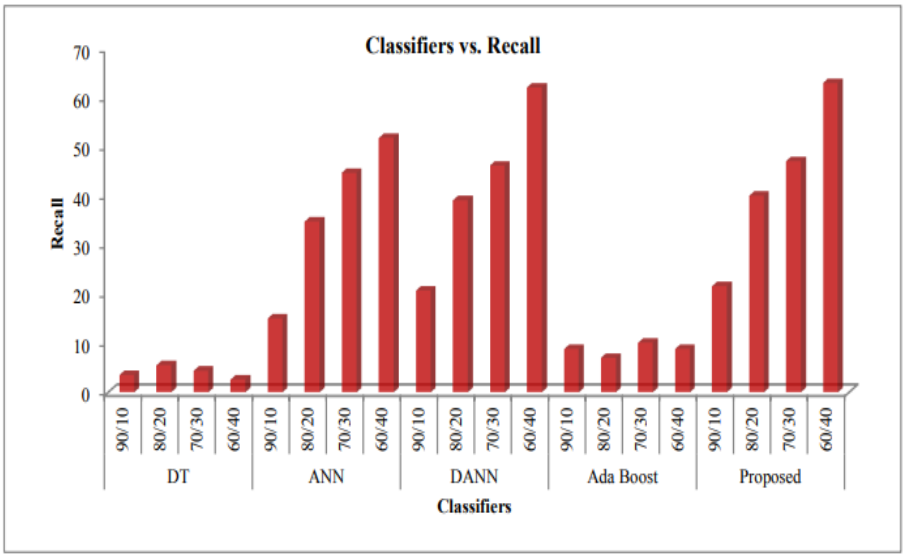

Fig. 5(c). Comparison of various Classification method vs Recall

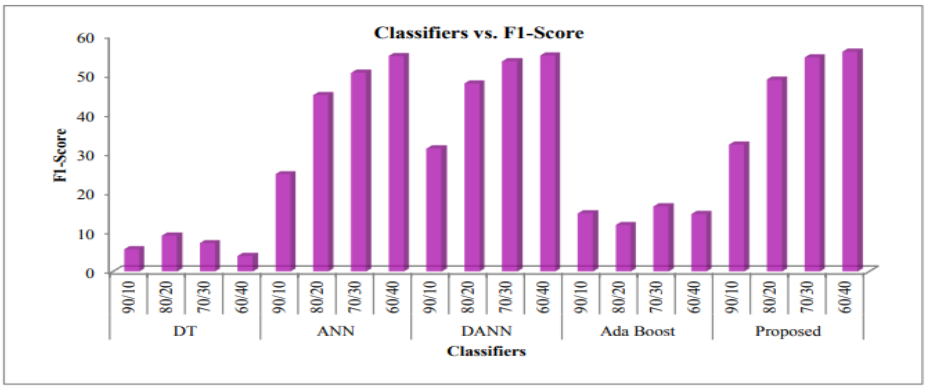

Fig. 5(d). Comparison of various Classification method vs F1 score

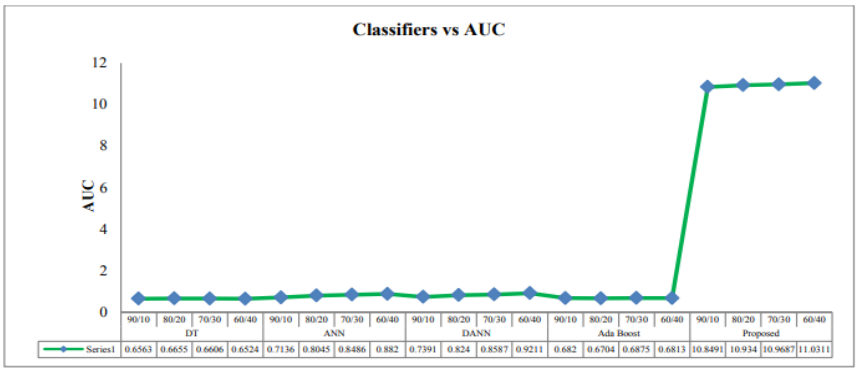

Fig. 5(e). Overall comparison of various Classification method vs Accuracy

Figure 5(b) shows that the accuracy of ANN improved noticeably at the 90/10 split, while it declined at the 70/30 split. AdaBoost achieved its best accuracy, 64.66%, at the 90/10 split. AdaBoost's performance was roughly 64% precise across three different ratios of splitting percentage. At first, it seemed that the 70/30 splitting% considerably increased DT in the precision test. On the other hand, its accuracy gradually dropped around the 60%/40% split. At a 90/10 split, MDANN's accuracy was almost as high as 63.06%.

Compared to 70/30, MDANN recall value is greater at 60/40, as shown in Figure 5(c). MDANN's low recall at the 90/10 split is easily seen. ANN's behaviour was analogous to that of MDANN when it gradually improved recall for each splitting %. As a result of using AdaBoost, we were able to raise that percentage from 90/10 to 60/40. Recall for DT was obviously much lower than it was for alternate methods of comparison. Figure 5(d) shows that when applying MDANN, the F1 score was best achieved at a 60/40 splitting% compared to other possible values (70/30, 80/90 and 90/10). Similarly to MDANN, ANN generated highest F1 score, with a score of 60/40. While AdaBoost's F1 Score was lower than that of MDANN and ANN, it still outperformed DT.

The steady increase of MDANN over the splitting percentage is shown in Figure 5(e). Indeed, as the size of training set shrank, the AUC value for MDANN increased. ANN's behaviour was consistent with that of MDANN, DANN and DT and the AUC was only slightly improved with AdaBoost (80/20 vs. 70/30). The AUC was between0.50 to 0.53 for both. Precision, accuracy, recall, AUC as well as F1 score performance at dissimilar percentages of splitting will vary across classifiers. For the vast majority of them, the 90/10 split would provide the highest degree of accuracy.

8. Conclusion

Important findings from this study include the fact that supervised learning approaches are better to other techniques because of the ease with which high quality model training may be accomplished because to the availability of labelled data. Due to their extensive instruction with massive datasets and high powered computers, pre trained models are very well suited to analysing information on electrical consumption. In this research, we compared four classifiers and four supervised learning algorithms for their ability to identify electrical theft. Measures including accuracy, recall, precision, F1 score and AUC allow for an evaluation of classifier effectiveness. In comparison to alternative classifiers of supervised learning including ANN, DANN AdaBoost and DT, MDANN achieved higher recall, F1 Score along with AUC. Additional supervised learning methods can be tested with different kinds of datasets in the future and appropriate preprocessing techniques can be incorporated to boost performance.

References

[1] W. Li, T. Logenthiran, V.T. Phan and W.L. Woo, "A novel smart energy theft system (SETS) for IoT-based smart home", IEEE Internet of Things Journal, Vol. 6, pp. 5531–9, 2019.

[2] P. Glauner, A. Boechat, L. Dolberg, R. State, F. Bettinger, Y. Rangoni and D. Duarte, "Large-scale detection of non-technical losses in imbalanced data sets", IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Vol. 6 pp. 1–5, 2016.

[3] J. Prasad, R. Samikannu, "Overview, issues and prevention of energy theft in smart grids and virtual power plants in Indian context", Energy Policy Vol. 110 pp. 365–74, 2017.

[4] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar and S. Mishra, "Decision tree and SVM- based data analytics for theft detection in smart grid", IEEE Transactions on Industrial Informatics, Vol. 12 pp. 1005–16, 2016.

[5] 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 Transactions on Industrial Informatics, Vol. 14 pp. 1606–15, 2017.

[6] R. Punmiya, S. Choe, "Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing", IEEE Transactions on Smart Grid Vol. 10, pp. 2326–9, 2019.

[7] 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 Journal of Electrical and Computer Engineering" Oct1 2019.

[8] M. Hasan, R.N. Toma, A.A. Nahid, M.M. Islam and J.M. Kim, "Electricity theft detection in smart grid systems: a CNN-LSTM based approach Energies", Vol. 12, p. 3310, 2019.

[9] K. Zheng, Q. Chen, Y. Wang, C. Kang and Q. Xia, "A novel combined data-driven approach for electricity theft detection", IEEE Transactions on Industrial Informatics Vol. 15 pp. 1809–19, 2018.

[10] M.M. Buzau, J. Tejedor-Aguilera, P. Cruz- Romero and A. Gomez-Exposito, "Hybrid deep neural networks for detection of non-technical losses in electricity smart meters", IEEE Transactions on Power Systems, Vol. 35 pp. 1254–63, 2019.

[11] P. Jokar, N. Arianpoo and V. C. M. Leung, "Electricity theft detection in AMI using customers' consumption patterns‖", IEEE Transactions on Smart Grid, Vol. 7, no. 1, pp. 216–226, 2016.

[12] Y. Wang, Q. Chen, D. Gan, J. Yang, D. S. Kirschen and C. Kang, "Deep learning- based socio-demographic information identification from smart meter datafication from smart meter data‖", IEEE Transactions on Smart Grid, Vol. 10, no. 3, pp. 2593–2602, 2019.

[13] R.R. Bhat, R.D. Trevizan, X. Li and A. Bretas, "Identifying nontechnical power loss via spatial and temporal deep learning", in Proceedings of the IEEE International Conference on Machine Learning and Applications, Anaheim, CA, USA, December 2016.

[14] M. Ismail, M. Shahin, M. F. Shaaban, E. Serpedin, and K. Qaraqe, "Efficient detection of electricity theft cyber attacks in ami networks", in Proceedings of the IEEE Wireless Communications and Networking Conference, Barcelona, Spain, April 2018.

[15] R. Mehrizi, X. Peng, X. Xu, S. Zhang, D. Metaxas and K. Li, "A computer vision based method for 3D posture estimation of symmetrical lifting", Journal of Biomechanics, Vol. 69, no. 1, pp. 40–46, 2018.

[16] J. B. Leite and J. R. S. Mantovani, "Detecting and locating nontechnical losses in modern distribution networks", IEEE Transactions on Smart Grid, Vol. 9, no. 2, pp. 1023–1032, 2018.

[17] W. Xing and D. Du, "Dropout prediction in MOOCs: using deep learning for personalized intervention", Journal of Ed- ucational Computing Research, Vol. 57, no. 3, pp. 547–570,2019.

[18] H. Kukreja, N. Bharath, C. S. Siddesh and S. Kuldeep, "An introduction to artificial neural network", International Journal of Advance Research and Innovative Ideas in Education, Vol. 1, no. 5, pp. 27–30, 2016.

[19] D. Yao, M. Wen, X. Liang, Z. Fu, K. Zhang and B. Yang, "Energy theft detection with energy privacy preservation in the smart grid", IEEE Internet of Things Journal, Vol. 6, no. 5, pp. 7659–7669, 2019.

[20] S. Shalev-Shwartz and S. Ben-David, "Understanding Machine Learning: From Theory to Algorithms", Cambridge University Press, Cambridge, UK, 2014.

[21] D. R. Pereira, M.A. Pazoti, L.A. Pereira, D. Rodrigues, C.O. Ramos, A.N. Souza and J.P.D. Papa, "Social-spider optimization-based support vector machines applied for energy theft detection", Computers & Electrical Engi- neering, Vol. 49, pp. 25–38, 2016.

[22] J. Pereira and F. Saraiva, "A comparative analysis of unbal- anced data handling techniques for machine learning algorithms to electricity theft detection,‖ in Proceedings of the 2020", IEEE Congress on Evolutionary Computation (CEC), July 2020.

[23] N. F. Avila, G. Figueroa and C.-C. Chu, "NTL detection in electric distribution systems using the maximal overlap dis- crete wavelet-packet transform and random undersampling boosting", IEEE Transactions on Power Systems, Vol. 33, no. 6, pp. 7171–7180, 2018.

[24] P. Glauner, J. Meira, L. Dolberg, R. State, F. Bettinger and Y. Rangoni, "Neighborhood features help detecting non- technical losses in big data sets, in Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies", Shanghai China, December, 2016.

[25] P. Massaferro, J. M. Di Martino and A. Ferna´ndez, "Fraud detection in electric power distribution: an approach that maximizes the economic return", IEEE Transactions on Power Systems, Vol. 35, no. 1, pp. 703–710, 2020.

[26] M. M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero and A. Gómez-Expósito, "Detection of non-technical losses using smart meter data and supervised learning", IEEE Transactions on Smart Grid, Vol. 10, no. 3, pp. 2661–2670, 2019.

[27] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar and S. Mishra, "Decision tree and SVM-based data analytics for theft detection in smart grid", IEEE Transactions on

Industrial Informatics, Vol. 12, no. 3, pp. 1005–1016, 2016.

[28] K. M. Ghori, R. Abbasi, M. Awais, M. Imran, A. Ullah and L. Szathmary, "Performance analysis of different types of machine learning classifiers for non-technical loss detection", IEEE Access, Vol. 8, pp. 16033–16048, 2019.

[29] G. Figueroa, Y.-S. Chen, N. Avila and C.-C. Chu, "Improved practices in machine learning algorithms for NTL detection with imbalanced data", in Proceedings of the 2017 IEEE Power & Energy Society General Meeting, July 2017.

[30] H. Huang, S. Liu and K. Davis, "Energy theft detection via artificial neural networks", in Proceedings of the 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), October 2018.

[31] Z. Chen, D. Meng, Y. Zhang, T. Xin and D. Xiao, "Electricity theft detection using deep bidirectional recurrent neural network", in Proceedings of the 2020 22nd International Conference on Advanced Communication Technology (ICACT), February 2020.

[32] P. Finardi, I. Campiotti, G. Plensack, R.D. de Souza, R. Nogueira, G. Pinheiro and R. Lotufo, "Electricity theft detection with self-attention", arXiv preprint arXiv: 2002.06219, 2020.

[33] Z. Zheng, Y. Yang, X. Niu, H. N. Dai and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft de- tection to secure smart grids", IEEE Transactions on Industrial Informatics, Vol. 14, no. 4, pp. 1606–1615, 2018.

[34] Z. A. Khan, M. Adil, N. Javaid, M. N. Saqib, M. Shafiq and J.-G. Choi, "Electricity theft detection using supervised learning techniques on smart meter data", Sustainability, Vol. 12, no. 19, p. 8023, 2020.

[35] R. Kurniawan, S. N. H. S. Abdullah, F. Lestari, M. Z. A. Nazri, A. Mujahidin and N. Adnan, "Clustering and correlation methods for predicting coronavirus COVID-19 risk analysis in pandemic countries", in Proceedings of the 2020 8th In- ternational Conference on Cyber and IT Service Management (CITSM), October 2020.

[36] L. J. Muhammad, E. A. Algehyne, S. S. Usman, A. Ahmad, Chakraborty and I. A. Mohammed, "Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset", SN Computer Science, Vol. 2, no. 1, pp. 11–13, 2021.

[37] S.S. Abdullah, F.A Bohani, Z.A. Nazri, Y. Jeffry, M.A. Abdullah, M.N. Junoh and Z.A. Kasim, "Amenities surrounding commercial serial crime prediction at greater valley and kuala lumpur using K-means clustering/penge- caman kemudahan awam sekitar lokasi jenayah kormesial bersiri di lembah klang dan kuala lumpur menggunakan kaedah gugusan K-means", Jurnal Teknologi, Vol. 80, no. 4, 2018.

[38] R. Razavi and M. Fleury, "Socio-economic predictors of electricity theft in developing countries: An Indian case study", Energy for Sustainable Development, Vol. 49, pp. 1–10, doi:10.1016/j.esd.2018.12.006, 2019.

[39] A. Ghazvini, S. N. H. S. Abdullah, M. Kamrul Hasan and Z. A. Bin Kasim, "Crime spatiotemporal prediction with fused objective function in time delay neural network", IEEE Access, Vol. 8, pp. 115167–115183, 2020.