Performance Evaluation of Different Machine Learning Techniques for Detection of Non-technical Loss

Adyasha Banajyoti and C. N. Bhende

Abstract The percentage of losses in India in transmission and distribution sector of electricity has been fairly high. In distribution system, a considerable amount of energy is dissipated which can be categorized into technical and non-technical losses. It is possible to control and compute technical losses, provided the load quantities are known for the given power system, whereas the non-technical losses do not have any recorded information as it is difficult to track energy theft due to the act of meter tampering or bypassing the measurement system. Generally, sudden or surprise checking is done in localities, where the electricity theft is suspected by the distribution companies. However, these operations alone are not enough to identify the miscreants or to reduce the energy losses. Moreover, manual inspection is quite tedious and can be very costly. Thus, certain advance technologies like machine learning techniques need be used to counter the electrical theft more effectively. In this paper, various machine learning techniques are discussed and their performances are compared for the detection of power theft in power system.

Keywords Machine learning · Power loss · Theft detection · Long short term memory · Boosting

1 Introduction

The economy of the country gets affected because of the losses in the power system. All energy that is generated and sent to the distribution and transmission system fails to arrive at the end consumer. We are losing a considerable amount of energy in our distribution system, which can be split off into the technical and non-technical losses. The distribution system is responsible for highest non-technical and technical losses.

The electrical components of the power system are responsible for the technical losses in power system. They occur naturally and it is mostly due to the power lost in

A. Banajyoti (✉) · C. N. Bhende
Indian Institute of Technology Bhubaneswar, Bhubaneswar, India
e-mail: Aab45@iitbbs.ac.in

S. K. Sabut et al. (eds.), *Proceedings of International Conference on Communication, Circuits, and Systems*, Lecture Notes in Electrical Engineering 728,
https://doi.org/10.1007/978-981-33-4866-0_11

measurement equipments, power transformers, most importantly transmission lines, etc. It is possible to compute and manage technical losses when we know load quantities and the system parameters.

However, non-technical losses are mostly because of the electrical theft, billing errors, defective energy meters, etc. Hence, these losses are more difficult to measure and there is no recorded information regarding the same.

Installation of smart meters can help till some extent but people have found advanced methods of the energy theft like attacking the operating system of the metering device.

The service providers usually perform unexpected visits to inspect in some suspected areas of increased electrical theft now a days. However, to capture the miscreants or to minimize the losses, these steps alone are not sufficient. Also, the manual inspection is a quite tedious and can be costly. Hence, to tackle the electrical theft more successfully, utility companies need to look into some of the latest technologies available in market. Using smart meters, we can very easily collect the electricity data of customers and we can use advanced data mining techniques to study vast data to find the electrical theft and hence the study of various machine learning techniques for the detection of energy theft. With the developments in the domain of machine learning, many strategies are put forward to identify and lessen the theft of energy, which are researched more in the literature survey.

In this paper, the performance of various machine learning schemes is analyzed for the detection of power theft. Section 1 contains the brief introduction on the power theft scenario in India. Section 2 consists of the researches on different type of energy theft detection techniques. There is a detailed working flow of the theft detection models and their results in the Sect. 4. Section 5 concludes the paper along with the comparison of all the models for the evaluation of their performance.

2 Various Machine Learning Techniques

Before adding details to the theft detection models, we need to be well acquainted with these machine learning techniques. The briefings of various machine learning schemes are given below.

2.1 Decision Tree Coupled Support Vector Machine (SVM)

SVM is used for classification problem may that be binary classification or multi-class classification. Many approaches based on SVM used the load profile of customers to expose irregularities or anomalous behavior which is highly related to NTL activities [1]. But in this model, the decision tree (DT) and SVM have been combined to detect non-technical loss [2]. At first, all the input features such as number of persons, temperature, electricity consumption, time of day are given to the regression DT, which predicts the expected electricity consumption of the costumers. Then, along with the previously defined input features, this newly generated expected electricity consumption data (output of DT) is supplied as input to the SVM. Finally, the SVM helps in detecting the energy theft in power system.

2.2 Probabilistic Neural Network (PNN)

Along with the input and output layer, the PNN has two more layers namely pattern layer and summation layer. The output layer generates a vector of probabilities for each input. In the final layer, the highest of these probabilities are picked and are labeled as 1 (positive, i.e., done theft) and others are labeled as 0 (negative, i.e., No theft). This inspiration has been taken from a paper, where a probabilistic neural network (PNN) classifier along with S transform is used for the identification of different power quality problems [3].

2.3 Stacked Long-Short Term Memory (LSTM)

Long-short term memory network is an advancement of recurrent neural network (RNN), which has the capability to remember very old information. It is usually done by explicitly introducing a memory unit into the network. Our electricity consumption data is time series data, hence the use of LSTM. Different papers have used LSTM or gated recurrent unit (GRU), which is another advancement of RNN, for predicting the power consumption and theft identification [4, 5].

2.4 Convolutional Neural Network with Long-Short Term Memory (CNN-LSTM)

Like other deep neural networks CNN also has input and output layer along with many hidden layers. In the CNN, the convolutional and pooling hidden layers are subsequently accompanied by the dense layer for dimension alteration of output of hidden layer. In this model, the CNN and LSTM are combined for detecting non-technical loss. The CNN has the capability to extract feature from a given data set, which made it useful for our model [6].

2.5 Gradient Boosting Algorithms

Boosting algorithms combine multiple weak models to create a strong model. Over the last few years, boosting algorithms have become immensely popular in the machine learning field because they are less affected by the over-fitting problem. Gradient boosting machine (GBM) uses multiple decision trees to generate the final prediction. In GBM, the errors done by the previous trees are considered in the each new decision tree generated. In quite a number of papers, recently developed gradient boosting classifiers were used for the detection of NTLs [7]. GBM has two variants, one is extreme gradient boosting and other is light gradient boosting which are given below.

Extreme Gradient Boosting (XGBoost) It is an advancement of the GBM algorithm. XGBoost follows the same working procedure as GBM. In this algorithm, the tree is split level-wise or depth-wise. XGBoost implements parallel preprocessing, which in turn makes it faster than GBM. The introduction of variety of regularization techniques in XGBoost helps in reducing over-fitting and finally improves the overall performance. Also, the XGBoost model is capable of handling the missing value cases in the dataset on its own.

Light Gradient Boosting (LightBoost) LightBoost algorithm is also an alteration of the GBM. The speed of execution and efficiency of the LightBoost model is the reason of its popularity. It can easily handle bulk data, but it does not perform well with a small dataset. Instead of a level-wise growth, the trees in LightBoost grow in a leaf-wise manner. LightBoost uses binning technique for continuous data, which results in lower memory usage and better efficiency. The leaf-wise split enables the LightBoost to perform good with bulk datasets with a notable reduction in execution time as compared to XGBoost.

3 Data Processing and Model Evaluation Metrics

3.1 Data Preprocessing

To be able to relate more with the real-time scenarios, we have selected various features for our dataset. These features are number of persons, temperature, time of day, season, electricity consumption. To validate proposed schemes, we have taken the hourly energy consumption data of six months for the residential area of IIT Bombay [8]. From various other sources, we have collected other features of our dataset and later they are linked together. From the historical meteorological dataset, we have extracted the details about season and temperature with respect to time and date. As we couldn’t find the number of persons in a specific house at a specific time, an average number of persons are reckoned in every house at that specific time.

3.2 The Unbalanced Data Problem

When the ratio of observations in each class are not in proportion that is called as an imbalanced dataset, which is a familiar issue in machine learning classification. Same is the case for our dataset, where the percentage of theft cases is merely a 10% of the whole dataset. So, the accuracy can be misleading in this case. For boosting methods, this problem was quite evident, hence data balancing technique called as synthetic minority over-sampling technique is used for boosting methods [9].

Synthetic Minority Over-Sampling Technique (SMOTE) Removing random observations from the majority class is called under-sampling, and adding duplicates of minority instances to the dataset is called over-sampling. SMOTE is one type of over-sampling method. While randomly removing data in random under-sampling, it may discard some important information and in random over-sampling duplicates may cause over-fitting. So, in SMOTE instead of taking duplicates, a subgroup of the dataset is considered from the minority class and then new synthetic similar instances are created, which solves the problem of over-fitting and useful information are not lost like under-sampling case.

3.3 Model Evaluation Metrics

In the field of machine learning, a confusion matrix depicts the results of the prediction by a classification problem. The number of right and wrong predictions are computed for each class and entered in the confusion matrix. From the confusion matrix, following performance measures can be found out,

Accuracy The number of cases correctly predicted to be doing theft or not doing theft divided by the total number of cases gives the accuracy. But sometimes accuracy can be misleading because it gives equal weightage to the errors of both the classes (theft and non-theft). Accuracy can be high even if our model is not detecting theft cases and only predicting the non-theft cases. That is why need other measures along with accuracy.

F-Measure As we cannot fully be dependent on accuracy, we compute F-measure. The total number of positive cases, who are also correctly predicted, divided by the total number of positive cases gives recall. The ratio of total number of positive cases to the cases who are predicted to be positive is called precision. It is the harmonic mean of both the measures. The F-measure takes into account the worst case between precision or recall.

Table 1 Comparison of different machine learning models

Model no.	Model name	Accuracy (%)	F-measure (%)
1	DT coupled SVM	87.91	63.93
2	Probabilistic neural network	92	70.7
3	Stacked LSTM	90.78	79.85
4	CNN with LSTM	89.8	72.8
5	XGBoost with SMOTE	92.18	82.44
6	LightBoost with SMOTE	95.124	88.33

4 Results and Discussion

Different machine learning models are analyzed; the same consumer data set is provided to them and each model is trained and evaluated. Confusion matrix for each model is generated and accuracy, recall, precision, and F-measure is determined.

To get a clear picture of the performances of the models, a comparison study is performed. The comparison of percentage of accuracy and F-measure between all the models have been given in the Table 1. From Table 1, it can be observed that although the SVM alone and DT coupled SVM have accuracy more than 85%, their F-measure is not up to the mark. The neural networks performed better in detecting non-theft but the detection rate of theft (evident from F-measure) is still not acceptable. The stacked LSTM performed quite well with acceptable accuracy and F-measure. The boosting methods outperformed all other methods, especially the LightBoost with SMOTE model which has more than 95% accuracy and nearly 90% F-measure.

5 Conclusion

In this paper, the performance of different machine learning models are analyzed. From the results, it is concluded that the LightBoost method outperformed all the other presented methods with the accuracy of more than 95% which makes the method suitable for practical use.

References

1. J. Nagi, A.M. Mohammad, K.S. Yap, S.K. Tiong, S.K. Ahmed, Non-technical loss analysis for detection of electricity theft using support vector machines, in *IEEE 2nd International Power and Energy Conference* (2008), pp. 907–912
2. A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, S. Mishra, Decision tree and SVM-based data analytics for theft detection in smart grid. IEEE Trans. Industr. Inf. **12**(3), 1005–1016 (2016)
3. S. Mishra, C.N. Bhende, B.K. Panigrahi, Detection and classification of power quality disturbances using s-transform and probabilistic neural network. IEEE Trans. Power Deliv. **23**(1), 280–287 (2008)
4. J. Goh, S. Adepu, M. Tan, Z.S. Lee, Anomaly detection in cyber physical systems using recurrent neural networks, in *IEEE 18th International Symposium on High Assurance Systems Engineering (HASE'2017)* (2017), pp. 140–145
5. X. Wang, T. Zhao, H. Liu, R. He, Power consumption predicting and anomalydetection based on long short-term memory neural network, in *IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA'2019)* (2019), pp. 487–491
6. R.R. Bhat, R.D. Trevizan, R. Sengupta, X. Li, A. Bretas, Identifying nontechnical power loss via spatial and temporal deep learning, in *15th IEEE International Conference on Machine Learning and Applications (ICMLA'2016)* (2016), pp. 272–279
7. R. Punmiya, S. Choe, Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. IEEE Trans. Smart Grid **10**(2), 2326–2329 (2019)
8. Open energy Info. https://seil.cse.iitb.ac.in/residential-dataset. Last accessed 13 Nov 2019
9. H. Lee, S. Jung, M. Kim, S. Kim, Synthetic minority over-sampling technique based on fuzzy c-means clustering for imbalanced data, in *International Conference on Fuzzy Theory and Its Applications (iFUZZY'2017)*, Pingtung (2017), pp. 1–6