Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Smart Meter for Power Theft Detection using Machine Learning

[1]Prof. Rohit Bamane, [2]Mervyn Vinod, [3]Jahnavi Shah, [4]Shweta Ahuja, [5]Aditya Sahariya

Department of Computer Engineering, S.P.P.U. /Dr. D. Y. Patil Institute of Technology, Pimpri.

---***---

Abstract *Power theft is one of the serious issues of electric utilities.* Such electricity theft produces a financial loss to the utility companies. It is not possible to manually inspect such a theft in large amounts. For detecting such electricity theft we have introduced a gradient boosting theft detector (GBTD) based on the three latest gradient boosting classifiers (GBCs): extreme gradient boosting (XGBoost), categorical boosting (Cat Boost), and light gradient boosting method (LightGBM). XGBoost is a machine learning algorithm which gives high accuracy in less time. In this algorithm, we perform preprocessing on the smart meter data and then perform feature selection. Practical application of the proposed GBTD for theft detection is minimizing FPR and reducing data storage space and improving time-complexity of the GBTD classifiers which detect nontechnical loss (NTL) detection.

Keywords: *electricity data, machine learning, XgBoost*

1. INTRODUCTION

Many electric utilities have financial losses due to electricity theft. There are various types of electrical power thefts, including tapping a line or bypassing the energy meter. As per surveys, 80% of the overall burglary happens in private abodes and 20% on business and modern premises. If we try to detect the theft manually, it would not be possible as large amount of data would be present. So, here we have applied machine learning algorithm to detect the theft. Theft can be detected by checking for abnormalities in the user's electricity consumption patterns. From the user's fundamental data, we can analyse the user behaviour. We implement a supervised ML-based theft detection model that identifies whether an abnormal/fraudulent usage pattern has occurred in the SG (smart grid) meter. We use superiority of XGBoost, a gradient boosting classifier (GBC), over other ML algorithms for nontechnical loss (NTL) detection.

1.1 Goals & Objectives

The main objective is to detect the theft based on fraudulent usage pattern.

Our goal is to detect abnormalities in the usage pattern & prevent electricity theft.

1.2 Scope

Usage of this model to detect fraudulent activities in the electric utility sector.

1.3 Existing System

In the existing system, electricity providers have to send their employees to check electricity meters of the users. Once the employee physically reaches the location, only then can he check the meter and find out for a theft.

- ***1.4 Proposed System***

  Our country loses more money to theft than any other country in the world. The state of Maharashtra which includes Mumbai- alone loses $2.8 billion per year, more than all but eight countries in the world. In this proposed system we use a dataset having the electricity usage of a smart grid (SG) meter. Using this dataset we perform feature selection and do pre-processing on the dataset.

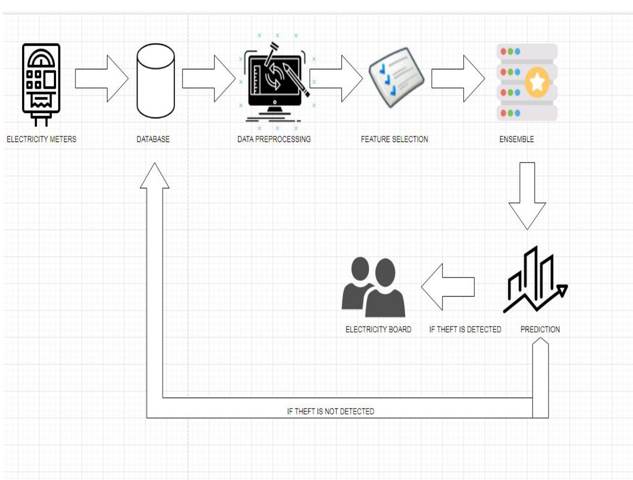
When we have large number of features in a dataset, then feature selection is a very important part in Machine Learning. As we use feature selection, it gives us the most important feature and this feature selection helps us achieve more accuracy. Then we perform preprocessing on the data. Next, we use the superiority of XGBoost, a gradient boosting classifier (GBC), over other ML algorithms for nontechnical loss (NTL) detection. Gradient boosting is called gradient boosting because it uses a gradient descent algorithm to minimize loss when adding new trees. This approach supports both regression and classification predictive.

- ***2. Motivation***

Electricity theft is increasing day by day and it has become difficult to detect fraudulent activities. Due to this, many citizens of the country are being cheated.

By creating a theft detection model, we can aim to reduce fraudulent activities in and around us and make our cities a better place to live in.

- *3 **System Architecture***



Explanation:

(a)SG meter data as input

(b) Next step is preprocessing on dataset.

(c) Feature selection method to select features.

(d) Ensemble learning method (Xgboost) applied on data.

(e) Result theft detection

4. CONCLUSIONS

The proposed system detects electricity theft using XgBoost machine learning method. However, LightGBM appeared to be the fastest classifier. This proposed system helps electricity providers to detect electricity theft and prevent loses. This is most important application of this project.

5 Literature Survey

1. Power Utility Nontechnical Loss Analysis with extreme machine learning method.

This paper proposes a new approach to identify the non-technical losses using a few emergent machine learning techniques called the extreme learning machine (ELM). Non-technical losses are caused by actions external to the power system and consist primarily of electricity theft. The ELM technique uses the data from the electricity board which consists of the customer's id and their meter readings. Data mining is used to extract patterns of customer's behavior depending on their meter readings. Any abnormal behavior that is recognized with the help of a major change in the pattern can be related to the NTL. ELM and online sequential ELM (OS-ELM) techniques are used. A comparison between ELM and other ML techniques such as SVM was undertaken and the results showed that ELM had higher accuracy and better performance which led to an increase in the efficiency of the theft detection results.

2. Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing.

This paper has introduced a Gradient boosting theft detector (GBTD) for smart theft identification. Gradient boosting theft detector consists of three latest gradient boosting classifiers (GBCs): extreme gradient boosting (XGBoost), categorical boosting (CatBoost), and light gradient boosting method (LightGBM). To improve the efficiency and the performance of the model, GBTD not just focuses on the preprocessing of the hyper parameters but also on the feature engineering-based preprocessing. The detection rate (DR) and false-positive rate (FPR) of those GBCs are improved by generating stochastic features like standard deviation, mean, minimum, and maximum value of daily electricity usage. To reduce the complexity occurring due to the number of classifiers, GBTD uses the weighted feature-importance technique. The main emphasis is on minimizing FPR, reducing data storage space and improving the time-complexity of the GBTD classifiers. Additionally, this letter proposes an updated version to automate the proposed algorithm by mimicking the real-world theft patterns and applying them to the dataset.

REFERENCES

[1] P. Jokar, N. Arianpoo, and V. C. M. Leung, “Electricity theft detection in AMI using customers’ consumption patterns,” IEEE Trans. Smart Grid, vol. 7, no. 1, pp. 216-226, Jan. 2016

[2] . 2. M. Buzau, J. Aguilera, P. Romero, and A. Expósito, “Detection of Non-Technical Losses Using Smart Meter Data and Supervised Learning,” IEEE Trans. Smart Grid, Feb. 2018. [DOI: 10.1109/TSG.2018.2807925]

[3] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 2016, San Francisco, CA, USA [DOI 10.1145/2939672.2939785].

[4]A. Dorogush, V. Ershov, and A. Gulin, “CatBoost: gradient boosting with categorical features support,” Workshop on ML Systems at Neural Information Processing Systems (NIPS), 2017. 5. G. Ke, Qi Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Liu. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree." 31st Conference on Neural Information Processing Systems (NIPS), 2017

[5] A. H Nizar, Z. Y. Dong IEEE Transactions on Power

Systems (Volume: 23, Issue: 3, Aug. 2008), 25 July 2008

ISSN Information: Print ISSN: 0885-8950 Electronic

ISSN: 1558-0679 INSPEC Accession Number: 10117920

Publisher: IEEE

