|| Volume 4 || Special Issue 9 || ICCEME 2019-2020 2020 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

ELECTRICITY THEFT DETECTION USING MACHINE LEARNING

Raghwendra Shukla[1], Mayur Vidhwani[2] ,Prof. V.R. Ghule[3]

Department, Computer Engineering, SKN Sinhgad Institute of Technology and Science (SKNSITS), Lonavla-Pune

Abstract: Power robbery is one of the serious issues of electric utilities. Such electricity theft produce financial loss to the utility companies. It is not possible to inspect manually such theft in large amount of data. For detecting such electricity theft introduces 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 one machine learning algorithm which gives high accuracy in less time. In this we apply pre-processing on smart meter data then does feature selection. Practical application of the proposed GBTD for theft detection by 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*

application ---------- ----------

I INTRODUCTION

Many electric utilities have financial loss due to electricity theft. Here are various types of electrical power theft, including Tapping a line or bypassing the energy meter As per an investigation [citation needed], 80% of overall burglary happens in private abodes and 20% on business and modern premises.. If we try to detect the theft manually then it's not possible as large amount of data will be there. So here we are applying machine learning algorithm to detect the theft. Theft can be detected by checking for abnormalities in the user's electricity consumption patterns. From user fundamental data it is easy task to analyze user behavior. 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 goal of this project is to find electricity theft and solve this issue..

1.2 Scope

Using this system its help to mseb for secure calculation.

1.3 Existing System

In existing system utilities have to send their employee to check the smart meter of users and when employee goes to check and that time if he get any users meter is off then only they get the theft.

1.4 Proposed System

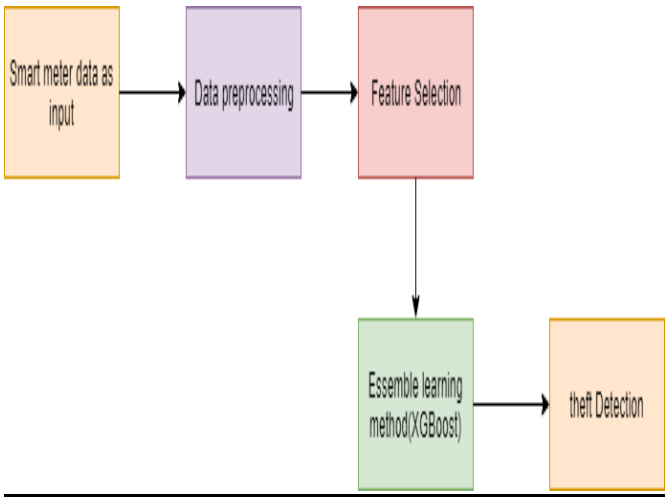
India loses more money to theft than any other country in the world. The state of Maharashtra which includes Mumbai—alone loses $2.8 billion per year, more than all but eight countries in the world. In this proposed system we use dataset having electricity usage of a smart grid (SG) meter (or simply smart meter).Using this dataset we does feature selection and preprocessing on dataset. When we have large number of features in dataset then feature selection is very important part in our Machine Learning. As we use feature selection it gives us most important feature and this feature selection gives us more accuracy. Then we perform the preprocessing on that data. After that 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.

II MOTIVATION

India 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 dataset having electricity usage of a smart grid (SG) meter (or simply smart meter).Using this dataset we does feature selection and preprocessing on dataset. When we have large number of features in dataset then feature selection is very important part in our Machine Learning. As we use feature selection it gives us most important feature and this feature selection gives us more accuracy. Then we perform the preprocessing on that data. After that 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

III SYSTEM ARCHITECTURE



Explanation:

(a)SG meter data as input

(b) Next step is pre-processing on dataset.

(c) Feature selection method to select features.

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

(e) Result theft detection

IV CONCLUSIONS

This proposed system detects the electricity theft using xgboost machine learning method. However, LightGBM appeared to be the fastest classifier. This proposed system helps to electricity utilities to detect electricity theft and they will not have to bare loss. This is most important application of this project.

V LITERATURE SURVEY

1. Machine Learning Algorithm for Efficient Power Theft Detection using Smart Meter Data.

This paper proposes an approach to identify the suspect customers using the customer power usage pattern. Machine learning algorithm is used for this purpose. The trustworthiness of customers is verified and is selected for theft program. This analysis is carried out by tweaking the actual Smart Meter data to create fraudulent data. The ANN classification model is developed using supervised learning algorithm that helps to discriminate the customers profile based on their genuine activity and fraudulent activity in electricity power usage. Simulation result shows that the proposed system is efficient in identifying the suspects with high accuracy.

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

For the smart grid energy theft identification, this letter introduces a gradient boosting theft detector (GBTD) based on the three latest gradient boosting classifiers (GBCs): extreme gradient boosting (XGBoost), categorical boosting (CatBoost), and light gradient boosting method (LightGBM). While most of existing ML algorithms just focus on fine tuning the hyperparameters of the classifiers, our ML algorithm, GBTD, focuses on the feature engineering-based preprocessing to improve detection performance as well as time-complexity. GBTD improves both detection rate (DR) and false positive rate (FPR) of those GBCs by generating stochastic features like standard deviation, mean, minimum, and maximum value of daily electricity usage. GBTD also reduces the classifier complexity with weighted feature-importance (WFI) based extraction techniques. Emphasis has been laid upon the practical application of the proposed ML for

theft detection by minimizing FPR and reducing data storage space and improving time-complexity of the GBTD classifiers. Additionally, this letter proposes an updated version of the existing six theft cases to mimic real world theft patterns and applies them to the dataset for numerical evaluation of the proposed algorithm ***Automated***

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