Pegem Journal of Education and Instruction, Vol. 13, No. 4 , 2023 (pp. 447-453)

RESEARCH ARTICLE WWW.PEGEGOG.NET

Safeguarding Power Grids against Electricity Theft by Data-Driven Analysis

[1]Mrs. M. Pragathi, [2]L. Sravani, [3]P. Kaarthika Reddy , [4]B.Niharika

[1]Assistant Professor, Department of Computer Science and Engineering, Sridevi Women's Engineering College, Hyderabad, India.

Email: yeddula.pragathireddy@gmail.com

[2,3,4,]B.Tech Student, Department of Computer Science and Engineering, Sridevi Women's Engineering College, Hyderabad, India.

Abstract:

In order to keep smart grids cost-effective, energy theft detection (ETD) plays a crucial role. Problems with missing values, large variation, and non-linearity mean that current ETD approaches are inadequate for dealing with the massive amounts of data that are currently accessible. Synchronising various steps in power theft categorization also requires an integrated infrastructure. In order to tackle these issues, a new ETD framework is suggested that integrates three separate components. The first section deals with data that isn't standardised, such as power usage, outliers, or missing information. Module 2 handles datasets with extreme imbalance by using a recently suggested hybrid class balancing method. To effectively and reliably forecast instances of power theft, the third module makes use of a classification engine based on an enhanced artificial neural network (iANN). To make conventional ANN better at handling more complicated classification jobs using smart metre (SM) data, we provide three unique mechanisms: hyper-parameters tweaking, regularisation, and skip connections. In order to enhance the final classification's generalisation and function fitting capabilities, we also study several iANN topologies. Results from numerical analyses of real-world energy use datasets show that the suggested ETD model outperforms state-of-the-art ML and DL approaches and works well for industrial use cases.

INTRODUCTION

There is a serious, widespread, and seemingly perpetual energy crisis. Neither do you want this to happen nor is it inevitable. Technical losses (TL) and nontechnical losses (NTL) are two categories of losses that power system networks experience while transferring energy. TL are an inevitable part of energy transmission and may be found in transformers, cables, and long-distance transmission lines. Electricity theft and non-payment of utility bills are the

Corresponding Author e-mail: yeddula.pragathireddy@gmail.com
How to cite this article: 1Mrs. M. Pragathi, 2L. Sravani, 3P. Kaarthika Reddy , 4B.Niharika. Safeguarding Power Grids against Electricity Theft by Data-Driven Analysis.Pegem Journal of Education and Instruction, Vol. 13, No. 4, 2023, 447-453. **Source of support:** Nil **Conflicts of Interest:** None.

DOI: 10.48047/pegegog.13.04.53

Received: 12.10.2023

Accepted: 22.11.2023 **Published:** 24.12.2023

two main components of non-transferable liabilities (NTL), which have long been a problem for utilities. Compared to nonpayment due to power waste and power quality issues, the multi-faceted cost of electricity theft is often substantial for utilities. Power theft has always been an issue for electric power providers, and no utility is safe from it. The current worldwide estimate puts the annual cost of energy theft at $96 billion. This percentage is much greater in poor nations, where it is expected to cost $60 billion annually. This massive loss cripples energy firms worldwide, pushes up costs for end-users, and raises the need for expensive government subsidies. Reducing power system losses is a primary goal of the smart grid in order to close the energy demand-supply imbalance. Power utilities now have sufficient instruments to fight energy theft and fraud, thanks to the recognition of internet of things (IoT) technologies and data-driven techniques focused on single-level data collecting. Numerous Internet of Things (IoT) sensors track a variety of real-time data points, including transmission, weather (wind, sun, temperature), and customers' power consumption records, and the data is updated on a regular basis to account for the unpredictable nature of energy use. For instance, the amount of historical data is large and challenging to examine during the uncertain COVID-19 periods, when individuals may choose to stay inside more.

RELATED WORK

Overview, problems, prevention, and a smart meter-based strategy for controlling electricitytheft

Utilities in underdeveloped nations have a hard time detecting and prosecuting those responsible for non-technical loss (NTL) during electrical energy transmission. Much of the NTL is comprised of stolen electricity. Because of these losses, the producing station is under more strain, and the rate that legitimate consumers pay rises. In this study, we'll look at what drives people to steal power. These negative consequences prompt a discussion of several approaches to theft detection and estimate. An external control station, filter circuit, harmonic generator, and smart metre are all suggested in

this paper's architectural design. The goal of this programme is to reduce energy waste and illicit use. Additionally, smart metres are designed to provide data on several metrics pertaining to real-time power use. Based on data sent by the distribution feeder's sending end, an external control station calculates the NTL in the feeder. In the event that a significant quantity of NTL is identified, the harmonic generator is turned on at that feeder in order to introduce an extra harmonic component, so ruining the appliances of the unauthorised users. A cost-benefit analysis for introducing the suggested system to India is given as an example.

Losses in the Power System, Both Technical and Non-Technical, and Their Impact on the Indian Economy

Electrical energy and peaking shortages are persistent problems in India. The country's economic development has been severely stunted as a result of these shortages. Because the sum of technical and nontechnical losses in the distribution system is equal to the overall losses. Extensive rural electrification, inadequate T&D capacity, an excessive number of transformation stages, incorrect load distribution, and other factors are listed as causes of such large losses. Put simply Over a certain time period, losses may be thought of as the disparity between the amount of power that enters the distribution network and the amount that leaves, as paid for by electricity accounts (estimated or metered). When it comes to electrical systems, technical losses are those brought on by factors like network impedance, current flows, and auxiliary supplies. Technical losses may originate from either investments in the network or operations inside the network. Theft, unbilled accounts, projected customer accounts, inaccuracies caused by estimating usage with unmetered supplies, and metering faults are all sources of nontechnical losses. In this research, we will use a case study and MATLAB simulation to examine power system technical and nontechnical losses.

An Advanced Metering Infrastructure Energy Theft Detection Framework Utilising Multiple Sensors

One of the most important parts of the smart grid is the enhanced metering infrastructure, which uses computers to replace older, analogue metres. Smart metres have made it easier to keep track of a large number of consumers, but they have also made AMI a prime target for hackers looking to steal energy via remote attacks or local physical damage. Although smart metres

include several sensors and data sources that may detect energy theft, the individual approaches often provide false positive results when used in reality. The article introduces AMIDS, an AMI intrusion detection system that integrates smart metre usage data with sensor readings for a more precise identification of energy theft. With the use of metre audit records of both physical and cyber activities, together with consumption data, AMIDS is able to better predict and identify theft-related actions. The testing findings on both typical and unusual load patterns demonstrate that AMIDS is capable of accurately detecting attempts at energy theft. Moreover, AMIDS accurately detected valid changes to the load profile that were mistakenly labelled as harmful by simpler studies.

Improving KBS for detecting nontechnical losses using statistical methods, text mining, and neural networks

Several issues pertaining to energy losses are now plaguing power distribution businesses. Illegal manipulation or a malfunction in the customer's measuring equipment might cause the energy consumed to go unbilled. These losses, known as nontechnical losses (NTLs), often outweigh losses caused by the distribution system, which are known as technical losses. While several studies have relied on data mining to identify NTLs in the past, none have yet included a Knowledge-Based System (KBS) built on the in-depth understanding of the inspectors. This research developed a KBS for NTL identification using statistical methods, text mining, and neural networks, all of which are based on the inspectors' prior knowledge and experience. In order to combine the rules created by the inspectors' expertise with the rules extracted from samples using text mining, neural networks, and statistical approaches, the rules were first translated into rules. The system was evaluated using actual data samples obtained from Endesa databases. With over 73 million clients, Endesa is a major player in the Spanish distribution industry and a major player in the global markets of Europe and South America.

METHODOLOGY

1. **Upload Electricity Theft Dataset**: Our dataset will be uploaded to the application using this module.

2. **Preprocess Dataset**: Because ML algorithms cannot handle character values, this module will be used to clean the dataset by eliminating missing values. Then, we will process the dataset to convert all nonnumeric characters to numeric characters by giving an integer ID to each unique string of data. Following processing, the dataset will be divided into two parts:

the train set will include 80% of the dataset, and the test set will comprise 20%.

3. Generate CNN Model: Train convolutional neural networks (CNNs) using datasets by using this module.

4. CNN with Random Forest: this module is used to get accuracy

5. CNN with SVM: this module is used to train dataset with CNN and SVM

6. Run Random Forest Random forest: is a supervised ML algorithm that trains RF independently on datasets; it finds extensive usage in classification and regression issues.

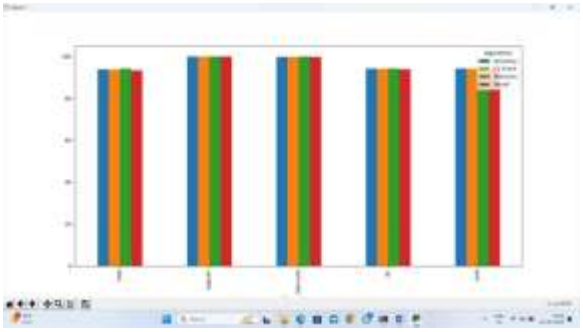
7. Run SVM Algorithm: Support vector machines (SVMs) are a kind of supervised machine learning algorithms that may be used to regression and classification issues. And here is the module that trains the SVM by itself using the datasets mentioned before.

8. Predict Electricity Theft this module is used to upload test data

9. Comparison Graph using this module we will show accuracy graph between all algorithms

RESULT AND DISCUSSION

The prediction result is shown as "record detected as ENERGY THEFT" or "record NOT detected as ENERGY THEFT" in the square brackets in the above screen, which also contains the test data. The following graph will appear when you click the "Comparison Graph" button.



Each method's name is on the x-axis while accuracy, recall, FSCORE, and precision are on the y-axis in the following graph; CNN-
RF achieves a perfect score for every algorithm.

CONCLUSION

In this paper, we provide a new CNN-RF model for detecting power theft. In this model, the convolutional neural network (CNN) investigates smart metre data in a manner analogous to an automated feature extractor,

while the RF serves as the output classifier. A fully connected layer with a dropout rate of 0.4 is developed during the training phase to reduce the danger of overfitting, which is caused by optimising a large number of parameters. Also, to get over the data imbalance issue, the SMOT algorithm is used. The same challenge is used to test several machine learning and deep learning algorithms; these methods have all been tested on SEAI and LCL datasets. Two features of the suggested CNN-RF model make it an attractive classification approach for usage in the area of energy theft detection, according to the results: 1. The hybrid model can automatically extract features, which is a huge time saver compared to other standard classifiers that rely on the retrieval of welldesigned features, which is a tedious and arduous process. Second, because both the RF and CNN are very effective classifiers for detecting power theft, the hybrid model takes advantage of their best features.Research in the future will focus on determining the potential impact of smart metre data granularity and length on customer privacy as it relates to power theft detection. We should look at the possibility of using the suggested hybrid CNN-

RFmodel to other tasks, such as load forecasting.

REFERENCES

[1] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, "Electricitytheft: overview, issues, prevention and a smart meter basedapproach to control theft," Energy Policy, vol. 39, no. 2,pp. 1007–1015, 2011.

[2] J. P. Navani, N. K. Sharma, and S. Sapra, "Technical and nontechnicallosses in power system and its economic consequencein Indian economy," International Journal of Electronicsand Computer Science Engineering, vol. 1, no. 2,pp. 757–761, 2012.

[3] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, andS. Zonouz, "A multi-sensor energy theft detection frameworkfor advanced metering infrastructures," IEEE Journal on SelectedAreas in Communications, vol. 31, no. 7, pp. 1319– 1330,2013.

[4] P. McDaniel and S. McLaughlin, "Security and privacychallenges in the smart grid," IEEE Security & PrivacyMagazine, vol. 7, no. 3, pp. 75–77, 2009.

[5] T. B. Smith, “Electricity theft: a comparative analysis,” EnergyPolicy, vol. 32, no. 1, pp. 2067–2076, 2004.

[6] J. I. Guerrero, C. Le´on, I. Monedero, F. Biscarri, andJ. Biscarri, “Improving knowledge-based systems with statisticaltechniques, text mining, and neural networks for nontechnicalloss detection,” Knowledge-Based Systems, vol. 71,no. 4, pp. 376–388, 2014.

[7] C. C. O. Ramos, A. N. Souza, G. Chiachia, A. X. Falcão, andJ. P. Papa, “A novel algorithm for feature selection usingharmony search and its application for non-technical lossesdetection,” Computers & Electrical Engineering, vol. 37, no. 6,pp. 886–894, 2011.

[8] P. Glauner, J. A. Meira, P. Valtchev, R. State, and F. Bettinger, “the challenge of non-technical loss detection using artificialintelligence: a surveyficial intelligence: a survey,” InternationalJournal of Computational Intelligence Systems,vol. 10, no. 1, pp. 760–775, 2017.

[9] S.-C. Huang, Y.-L. Lo, and C.-N. Lu, “Non-technical lossdetection using state estimation and analysis of variance,”IEEE Transactions on Power Systems, vol. 28, no. 3,pp. 2959–2966, 2013.

[10] O. Rahmati, H. R. Pourghasemi, and A. M. Melesse, “Applicationof GIS-based data driven random forest and maximumentropy models for groundwater potential mapping: acase study at Mehran region, Iran,” CATENA, vol. 137,pp. 360–372, 2016.