

Detection of Non-Technical Loss in Power Utilities using Data Mining Techniques

Rajesh. R

PG Student

*Department of Computer Science and Engineering
Government College of Engineering, Tirunelveli, Tamilnadu,
India*

Mrs E.Siva Sankari

Assistant Professor

*Department of Computer Science and Engineering
Government College of Engineering, Tirunelveli, Tamilnadu,
India*

Matheswaran.P

PG Student

*Department of Computer Science and Engineering
Government College of Engineering, Tirunelveli, Tamilnadu, India*

Abstract

This paper presents Non-Technical loss (NTL) in power utilities and it describes how to handle. Non-technical loss has been an influential factor on the benefits of electric power utilities. At the same time, with distribute generation extensively installed, the consumption patterns having many similarities between dishonest users and normal customers. Non-Technical Loss may be theft of electricity, illegal connection, fault metering and billing error. Improving the reliability of NTL detection algorithm becomes particularly important. Data mining techniques are used to detect the Non-Technical Loss using classification algorithm. The implementation of a intelligent computational tool to identify the non-technical losses and to select its most close feature, considering information from the database with consumers profiles. This work presents using the weka software to the proposed objective, comparing various classification techniques and optimization through intelligent algorithm.

Keywords: Multilayer Perceptron, Data Mining, Non-Technical Loss

I. INTRODUCTION

The losses of electric power are constituted by the difference between the generated or bought energy and the billed ones, and can be divided into two distinct types: technical and nontechnical losses. The former are related with problems in the system through the physical characteristics of the equipment, that is, the technical losses are the energy lost in the transport, the transformation and the equipment of measurement, becoming a very high cost to the electric power companies. The non-technical losses are those associated with the commercialization of the supplied energy to the user and refer to the delivered and not billed energy, propitiating a loss in the profits. They are also defined as the difference between the total losses and the technical losses, being strongly related to illegal connections in the distribution system.

In recent years, the problem of detecting non-technical losses in distribution systems has been paramount. Theft and adulteration of power meters, with the purpose to modify the measurement of the energy consumption, are the main causes that lead to non-technical losses in power companies. Since that to perform periodic inspections to minimize such frauds may be very expensive, it is a hard task to calculate or measure the amount of losses, and in most part of the cases it is almost impossible to know where they occur . Aimed at reducing fraud and energy theft, several electric power companies have been concerned that the illegal connections should be better profiled. The minimization of such losses may guarantee investments in energy quality programs, as well as it may enable a reduction in its price to the consumer. Currently, some advances in this area can be observed with the use of various artificial intelligence techniques in order to automatically identify non-technical losses, which is a real application in Smart Grids. Importantly, the solution to this problem is not trivial, nor of regional character. Despite the widespread use of machine learning techniques for the identification of non-technical losses in power systems, the problem of selecting the most representative features has not been widely discussed in the context of nontechnical losses.

In this sense, it is also believed that identifying the best characteristics that describe possible non-technical losses is as important as identifying such losses. In 2004, electricity theft in India was roughly US\$4.5 billion and 1.5 percent of GDP. The case of India differs in many aspects from the case in Jamaica. Electricity theft in India is primarily political and has its roots in the agrarian sector of the society . State Electricity Boards (SEB) have grown increasing negligent of the problem and stealing electricity has almost become a way of life for some. Some provinces in India have taken several measures to mitigate against the widespread practice of electricity theft in the 4 region. In Andhra Pradesh, they showed that although they have an SEB, they could still reduce the volume of electricity theft in their region. Prior to the mechanisms employed, the implications of electricity theft were great. Subsidies to the electricity sector imposed on budgetary allowances to other just as needy sectors. The subsequent strategy targeting electricity theft had a four prong approach. This approach dealt with strengthening the law and

enhancing enforcement of it; there was also an effort to identify the corruption involved within the utilities and a better management control and customer service system for the utility company.

A. Technical Losses:

- Copper losses those are due to $I^2 R$ losses that are inherent in all inductors because of the finite resistance of conductors
- Dielectric losses that are losses that result from the heating effect on the dielectric material between conductors
- Induction and radiation losses that are produced by the electromagnetic fields surrounding conductors. Technical losses are possible to compute and control, provided the power system in question consists of known quantities of loads. The following are the causes of technical losses:
 - Harmonics distortion
 - Improper earthing at consumer end
 - Long single phase lines
 - Unbalanced loading
 - Losses due to overloading and low voltage
 - Losses due to poor standard of equipments.

B. Non-Technical Losses:

- Tampering with meters to ensure the meter recorded a lower consumption reading
- Errors in technical losses computation
- Tapping (hooking) on LT lines
- Arranging false readings by bribing meter readers
- Stealing by bypassing the meter or otherwise making illegal connections
- By just ignoring unpaid bills
- Faulty energy meters or un-metered supply
- Errors and delay in meter reading and billing
- Non-payment by customers.

II. METHODOLOGY

The computer used for testing is a Dell PowerEdge 840 workstation with Windows XP, a 2.40 GHz Intel Quad-core Xeon X3320 Processor with 4 GB of RAM. In our work WEKA tool has to been used. WEKA aims aggregating algorithms from different approaches or paradigms in the subfield of artificial intelligence dedicated to the study of learning by machines . This work uses WEKA to identify and to select features because the implementation and manipulation of intelligent techniques are simple and fast.

A. Data Collection:

The data collection for this research study was performed two times. In the first data collection phase, historical customer billing and consumption data was collected for training, i.e., to make the intelligent system learn, memorize and differentiate between normal and suspicious consumption patterns. In the second data collection phase, similar data was collected as in the first stage, but this data was used for testing the fraud detection system, i.e., to detect and identify suspicious customers.

Two types of customer data were collected, which are :

- Enhanced Customer Information Billing System (e-CIBS) data
- High Risk Data.

B. Customer Filtering and Selection:

As the e-CIBS data acquired from TNBD is in raw format, therefore, in order to extract relevant and useful information, only customers with complete and useful data were selected from uci repository, for the MLP model Development. Since, the data acquired is in the form of a database, therefore data mining techniques using the Structured Query Language (SQL) were applied to satisfy the four criteria, as follows:

- Remove repeating customers in the monthly data.
- Remove customers having no consumption (i.e., 0 kWh) throughout the entire 12 month period.
- Remove customers not present within the entire 12 month period

C. Data Preprocessing:

Data preprocessing is the first major stage involved in the development of the fraud detection system. Data preprocessing involves data mining techniques in order to transform raw customer data into the required format, to be used for detection and identification of fraud load consumption patterns. Different data preprocessing techniques have been applied to the CSV dataset.

Weka filter act as an important role in preprocessing technique. Based on its operation the filter may be classified into supervised and unsupervised filter.

D. Classification:

Classification is using to predict the unknown value .The prediction will gain knowledge from training data. The development of the classification, namely the MLP model, is the main focus of this project and research study. Development of the classification involves: load profile inspection for detection of normal and fraud customers, training and development of the MLP classifier, and MLP testing.

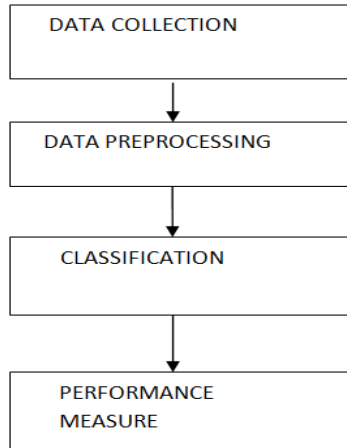


Fig. 1: Execution Flow

E. Multilayer Perceptron:

ANNs consist of a method of solving problems related to engineering and science through simple circuits that mimic the human brain, including their behavior, i.e., learning, making mistakes and making discoveries. An ANN is ormed by small modules which simulate the operation of a neuron.

These modules operate according to the elements that were inspired, receiving and transmitting information. A simple artificial neuron model has the main features of a biological neural network, parallelism and high connectivity, a neuron is composed of a linear combination and a transfer function. ANN is characterized by having an input layer and having one or more intermediate or hidden layers and the information is received by the input layer, processed by hidden layer, and displayed by the output layer. Generally, this topology is applied in pattern recognition and also to approximate non-linear functions. Figure 2 consists of a multilayer neural network and normally is chosen architecture empirically. In this work, the topology or architecture of ANN is designed by standard WEKA, i.e., there are no changes.

The parameters of ANN are:

- Training algorithm: backpropagation
- Transfer function: sigmoid;
- Number of epochs: 500;
- Momentum: 0.2; and
- Learning rate: 0.1.

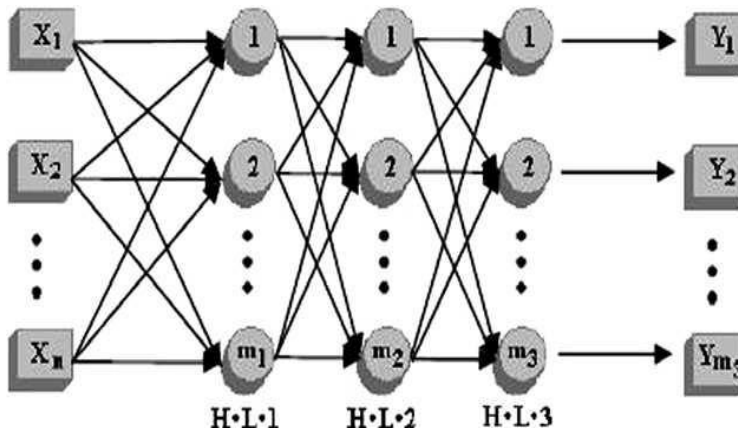


Fig. 2: Multilayer Perceptron model.

III. EXPERIMENTAL RESULT

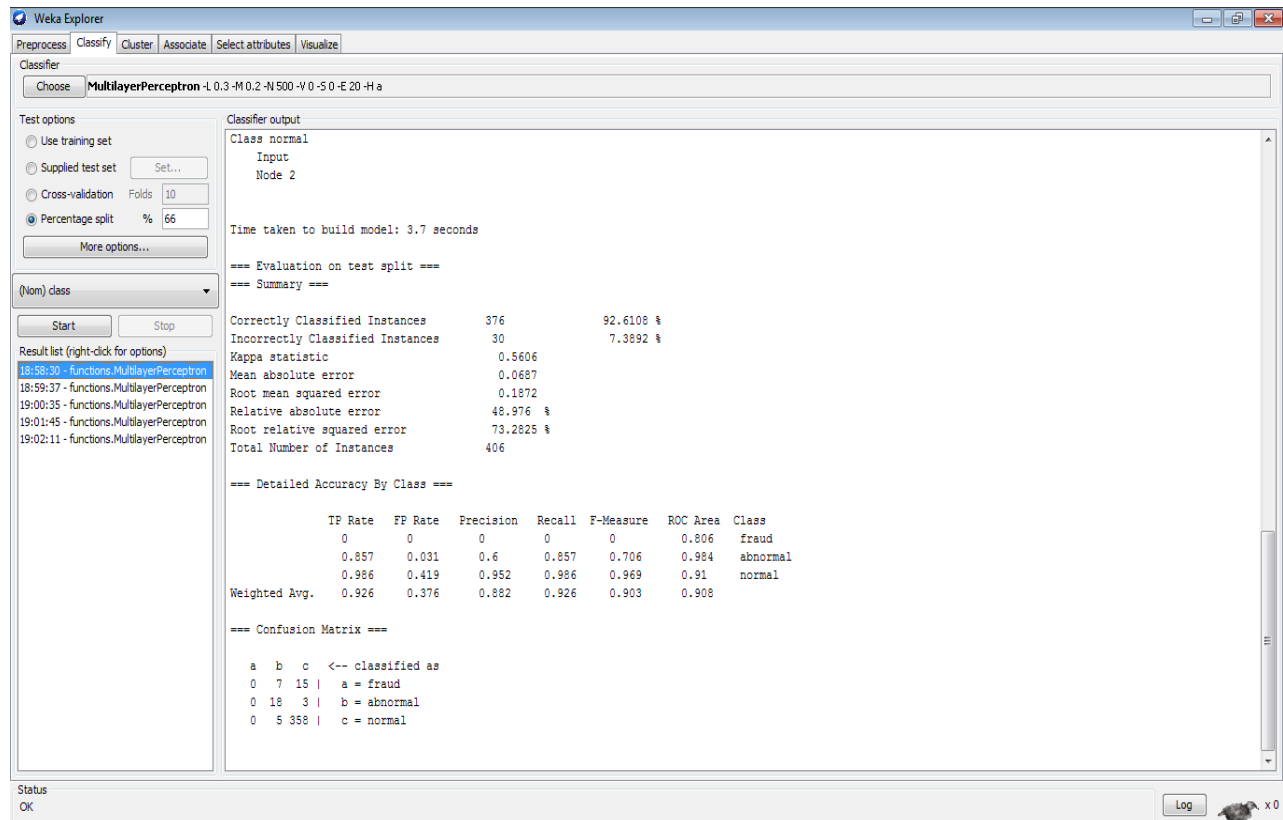


Fig. 3: classification output based on multilayer perceptron

Figure 3 describes the classification accuracy. In this project the percentage of correctly classified instances are 92.6108 %. Incorrect percentages are 7.3892 %. The confusion matrix shows classification accuracy in matrix format. In this matrix the diagonal element shows correctly classified instance. Other than diagonal element shows incorrect result.

IV. CONCLUSION

Non-technical losses in power distribution systems are related to the energy consumed and that is not billed, representing a problem for power companies, since these losses are strongly related to fraud and theft of energy, where the impact is huge on consumers industrial. Although several works have studied this subject, few of them addresses the problem of selecting the most relevant features to identify potential consumers that present non-technical losses. The problem of selecting a subset of features that maximizes the accuracy rate of the classifiers can be modeled as an optimization procedure. The proposed method was validated on a set of labeled data with industrial profiles that it was obtained a Brazilian power company. Therefore, in order to recognize potential consumers that present non-technical losses, it is also important to find the best features that can differentiate it. In this paper, it is shown that a subset of features can increase the accuracy rate of a classifier and also present an alternative and simple way to classify and select the most representative features through WEKA software. In future, it is intended to use different techniques for intelligent classification and optimization.

REFERENCES

- [1] J. Nagi, A. Mohammad, K. Yap, S. Tiong, and S. Ahmed, "Nontechnical loss analysis for detection of electricity theft using support vector machines," in Proceedings of the 2nd IEEE International Power and Energy Conference, 2008, pp. 907–912.
- [2] K. S. Yap, Z. Hussien, and A. Mohamad, "Abnormalities and fraud electric meter detection using hybrid support vector machine and genetic algorithm," in Proc. 3rd IASTED Int. Conf. Advances in Computer Science and Technology, Phuket, Thailand, Apr. 2–4, 2007.
- [3] Kim, M, and Kim T., "A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection", Proceedings of IDEAL, pp.:378-383, 2002.
- [4] Carlos León, Félix Biscarri, Inigo Monedero, Juan Ignacio Guerrero, Jesús Biscarri, and Rocio Millán, "Variability and Trend-Based Generalized Rule Induction Model to NTL Detection in Power Companies" IEEE Trasaction on Power system, VOL. 26, NO. 4, NOVEMBER 2011.
- [5] A. H. Nizar, Z. Y. Dong, J. H. Zhao, and P. Zhang, "A Data Mining Based NTL Analysis Method", IEEE Power Engineering Society (PES) General Meeting, 2007
- [6] M. Kantardzic, "Data Mining: Concepts, Models, Methods, and Algorithms", Hoboken, NJ: Wiley-Interscience: IEEE Press, 2003.
- [7] T. B. Smith, "Electricity Theft: A Comparative Analysis", Energy Policy, vol.32, no. 18, Dec. 2004,

- [8] J. Filho, "Fraud identification in electricity company costumers using decision tree," in Proc. IEEE/PES Int. Conf. Systems, Man and Cybernetics, The Hague, The Netherlands, 2004.
- [9] Nanlin Jin, Peter Flach, Tom Wilcox, Royston Sellman, Joshua Thumim, Arno Knobbe "Subgroup Discovery in Smart Electricity Meter Data" IEEE Transaction on industrial informatics, vol ,10,no.2,may 2014.
- [10] A. H. Nizar, Z. Y. Dong, M. Jalaluddin, and M. J. Raffles, "Load Profiling Non-Technical Loss Activities in a Power Utility", in Proc. of the 1st International Power and Energy Conference (PECON), November 28-29, Putrajaya, Malaysia, 2006.