Behavioral Diagnosis on Individual Electricity Consumption: Formulation Using a Neural Network Based on Adaptive Resonance Theory

Salvador Falcón Canillas , Reginaldo José da Silva , Carlos Roberto Minussi *

Posted Date: 28 July 2025

doi: 10.20944/preprints202507.2247.v1

Keywords: non-technical electrical losses (NTL); neural network; fuzzy-ART; adaptive resonance theory; fraud detection; consumer behavior; supervised and unsupervised classification; continual learning

 doi:10.20944/preprints202507.2247.v1

Article

Behavioral Diagnosis on Individual Electricity Consumption: Formulation Using a Neural Network Based on Adaptive Resonance Theory

Salvador Falcón Canillas, Reginaldo José da Silva and Carlos R. Minussi *

Electrical Engineering Department, UNESP - São Paulo State University, Av. Brasil 56-15385-000, Ilha Solteira (SP), Brazil

* Correspondence: carlos.minussi@unesp.br

Abstract

This research aims to study the daily consumption behavior of individual customers connected to the electricity distribution network and, extending it to longer periods, seek evidence of fraud, classified as non-technical losses. It should be noted that current Brazilian legislation authorizes distribution companies to pass non-technical losses on to electricity tariffs, consequently increasing the tariff for consumers who comply with their contractual obligations. In contrast to this practice, this research aims to develop a system for studying consumer behavior collaboratively and in complement to existing techniques, thereby mitigating or eliminating these losses. To achieve this objective, we propose the development of an inference system based on ANNs from the adaptive resonance theory (ART) family of [1] and Grossberg [2,3]. Specifically, a Fuzzy-ART network, known for its ability to learn reliably and in real time, was employed. The customer consumption data used to develop this detection system comes from real customers of the Commission for Energy Regulation (CER) of Ireland, utilizing data from only one year to extract different consumption patterns across various seasons. Each sample, or input vector, corresponds to a customer's daily consumption in 30-minute intervals, allowing for the capture of information about the customer at different times of the day. Given the difficulty of obtaining real data, seven types of fraud were generated to represent, as closely as possible, the various types of fraudsters that might be encountered in real life. To avoid biasing the model due to the typical predominance of benign data, the database was balanced, consisting of 3,500 days of benign customer data and 3,500 days of fraudulent customer data.

Keywords: non-technical electrical losses (NTL); neural network; fuzzy-ART; adaptive resonance theory; fraud detection; consumer behavior; supervised and unsupervised classification; continual learning

1. Introduction

The purpose of this research is to understand the behavior of each consumer's electricity consumption. In this case, several parameters will be investigated, ultimately providing crucial technical support to complement and assist in identifying suspect points of non-technical losses [4], which are intentional actions involving the use of electricity.

Electric Power Systems (EPSs) must be planned and operated to meet the electricity demand of their customers (residential, commercial, and industrial) with quality (including voltage, waveform, and frequency, among others, within pre-established variations) and continuity, in addition to covering a portion of electrical losses. These losses are classified as: (1) technical electrical losses; and (2) non-technical losses. Technical electrical losses are primarily a consequence of the circulation of electric current in all conductive elements due to electronic collisions (Joule effect) between electrical charges, resulting in increased temperature released in the form of heat, as well as leaks (electric current leakage points) due to imperfect insulation, among other factors. Non-technical losses result

 doi:10.20944/preprints202507.2247.v1

from fraudulent actions by regulated consumers, as well as by outside agents in the context of EPSs. The elimination, or at least minimization, of non-technical losses is necessary because they compromise electrical power quality, mainly when they result from "workarounds" that are difficult to identify.

The study of non-technical electrical losses has been addressed in the literature using various statistical techniques, and recently, there has been a considerable increase in methods based on machine learning [5]. An extensive approach to this topic is discussed later in the "Related Research" section. This doctoral research aims to develop an inference system for identifying evidence of consumers with atypical behavior, characterized by practices that result in non-technical losses. These losses, which do not represent commercial revenues for electricity companies, are ultimately paid for by consumers, particularly those who are part of [6]. Therefore, it is necessary to mitigate these non-technical losses by converting them into benefits of social value, primarily with the desired objective of reducing regular electricity rates.

The proposed inference system utilizes an artificial neural network (ANN) [5] with continuous training [7], which distinguishes it from most ANNs used in the specialized literature; in other words, it is an innovative proposal. The ANN, the subject of this research, refers to the Fuzzy-ARTa plastic architecture [1–3] designed for unsupervised training and is appropriately suited to the problem under study (diagnosis of non-technical losses in EPS).

This ANN is part of a family of ART neural networks proposed by [1]. The great advantage of using this family of ANNs is seen in its intrinsic characteristics: stability, plasticity, simplicity, and speed of training. Stability represents the full guarantee of convergence in the ANN training phase. Plasticity is a promising quality, as it allows continuous (incremental) training. That is, it is an ANN that, while performing diagnostics, enables the inclusion of new knowledge without the need to restart the entire training process, avoiding the well-known idealized cognitive condition characterized as a complete void (tabula rasa) (Aristotle). With due regard for its specificities, in this case, the execution is similar to human action, which is continuous learning. Considering the set of qualities of the ART-descendant ANN, the aim is to provide an inference system with high-speed, incremental, and reliable training. By incorporating the concepts of fuzzy sets [8], Carpenter and Grossberg [1] reformulated the original proposal for the Fuzzy-ART architecture (unsupervised training), and the ANN Fuzzy-ARTMAP (supervised training), to expand the capacity to work with more general data, mixing analog and binary information in a pre-processed universe considering values within the interval [0,1]. Note that this represents efficient coding and facilitates the treatment, with appropriate adjustments, of any real-world problem, such as the problem addressed in this research. Indeed, the applications can be extended to solving other issues, such as fraud detection in today's vast universe.

This formulation aims to offer a series of resources with the purpose of better understanding the problem of non-technical losses: learning and interpreting new modus operandi strategies, continuously (endemic event), which have been practiced by people with the habit of defrauding electricity consumption.

2. Related Works

2.1. Categorization of Non-Technical Loss Detection Methods

Identifying non-technical losses in electrical systems is critical to ensuring the operational efficiency and economic sustainability of energy distribution entities. These losses often linked to fraudulent activities or inconsistencies unrelated to the grid's physical infrastructure require meticulous detection approaches. In this context, we aim to classify and understand various detection schemes based on their methodologies and data sources.

Data-driven methods focus on analyzing specific consumer-related information. These algorithms, exploring details such as consumption patterns and demographic characteristics, seek to identify discrepancies that may be indicative of losses. Within this category, methods can be divided

 doi:10.20944/preprints202507.2247.v1

into supervised and unsupervised. Supervised methods utilize a predefined set of categories or labels, training models to distinguish between fraudulent and non-fraudulent situations. On the other hand, unsupervised methods operate without this prior classification, focusing on identifying anomalies based on the intrinsic structure of the data. Network-Oriented Methods, on the other hand, emphasize data from the electrical infrastructure itself. These schemes, utilizing parameters such as network topology and performance metrics, aim to identify losses resulting from physical or technical irregularities. These methods can employ a range of techniques, including state estimation, load flow analysis, and the utilization of specialized sensors to detect irregular activity.

Finally, Hybrid Methods attempt to combine the strengths of the two previous approaches. These algorithms can, for example, begin with a technical assessment of the network to identify potential loss zones and subsequently apply data-driven methods to refine detection at the consumer level [9]. In summary, the parameters typically found in all non-technical loss detection articles are:

- **Category and Concept**: Refers to the specific classification and related subcategories of a research study.
- **Algorithms**:Conceptual indications of the computational methods employed to identify non-technical losses. The nature of the data determines the selection of the appropriate algorithm, the application context, and the specific performance requirements. Certain algorithms may be more effective in environments with large datasets, while others may be optimized for greater accuracy in more limited datasets.
- **Data Type**: Specifies the data types required for each detection approach. This dimension is crucial when developing a new method or selecting between existing approaches.
- **Dataset Size**: Relates to the volume of data required for practical analysis, generally determined by the number of consumers involved.
- **Features**:In many scenarios, before any analysis, raw data is processed to extract essential features that will serve as input for classification techniques.
- **Metrics**: Represents the performance metrics adopted to evaluate the effectiveness of detection systems in different contexts, facilitating comparison between different approaches.

2.1.1. Categorization and Definitions of Data Types

The proposed categorization aims to ensure that researchers aren't limited to specific data types when selecting an algorithm. The central idea is to provide flexibility in selecting non-technical loss detection systems based on the available data.

2.1.2. Raw Data Used in NTL Detection

Data are meticulously categorized based on their physical source. Data originating from individual consumers, such as active energy measurements, is designated as "Consumer-Level" data. Conversely, information originating from a broader geographic range, such as details of the network topology, is classified as "Area-Level" data. Interestingly, these datasets, regardless of their primary classification, can be subdivided into time-series and static data, allowing for a more detailed and specific organization. A thorough analysis of the current literature reveals a notable gap in the use of static data associated with the "Area-Level," particularly those that are not intrinsically linked to the network topology. Illustratively, [10] stands out as one of the few works that incorporates "Area-Level" data in NPD. Additionally, it is noteworthy that high-precision energy data, as well as environmental information such as temperature, are rarely explored in the existing literature.

- The methods, with their data-driven foundation, primarily rely on time series and static data related to the consumer. It is notable that such methods unanimously incorporate data related to energy consumption in a temporal sequence, and approximately half of them also integrate static data.
- In turn, grid-focused methods adopt energy consumption data, whether high or medium resolution, complemented by measurements related to voltage and current. It is worth

 doi:10.20944/preprints202507.2247.v1

highlighting the importance attributed to measurements from other devices, such as observer meters or RTUs, and the essential need for a detailed understanding of the topology of networks.
- In line with what could be anticipated, the methods classified as hybrid explore a broad spectrum, encompassing both of the two categories previously mentioned (data-oriented and network-oriented).

2.1.3. Features Used in NTL Detection

It is common for data-driven (sometimes hybrid) NTL detection methods to utilize not only the raw data described above, but also features extracted from that data. Commonly used features are listed here. It is crucial to emphasize that these features are predominantly derived from consumer-centric time series, with particular emphasis on curves related to active energy consumption. The temporal resolution of these features intrinsically mirrors the resolution of the raw data. Therefore, it is challenging to detail all the features employed in each study accurately. Conversely, many researchers prefer to delineate feature sets, providing a broader and more generalized understanding of the topic. Relying on distinctive features, rather than strictly focusing on the time series, is a common practice in data mining activities such as classification and clustering.

The main features used for NTL detection are presented below:

- Basic Statistics: Comprising the average, maximum/minimum values, and standard deviation, calculated over a given interval.
- Power Factor: Established as the ratio between active (kW) and reactive (kVAr) power, this requires instantaneous power measurements and high-resolution data (preferably up to 15 minutes) for an accurate estimate.
- Load Factor: Describes the relationship between average consumption and peak active energy over a set period.
- Streaks: Denotes the frequency with which the consumption curve crosses a defined moving average.
- Consumption to Contracted Power Ratio: Relates total active energy consumption over a period to the contracted power value.
- Pearson's Coefficient [11]: This coefficient assesses the adequacy of a linear regression between active energy consumption and time.
- Billed/Consumed Energy Ratio: Reflects the discrepancy between billed and consumed energy, normalized by contracted power.
- Consumption Projection: An estimate of future consumption or the discrepancy between a projection and observed values.
- Wavelet Coefficients [12]: Measure the discrepancy between the wavelet coefficients of a current consumption curve and those of previous periods.
- Fourier Coefficients: Similar to wavelet coefficients, but focused on Fourier analysis. The phase of the first coefficients can also be taken into consideration.
- Polynomial Coefficients: Contrast the coefficients of polynomials fitted to the current consumption curve with those of previous periods.
- Distance to the Average Consumer: Corresponds to the calculation of the Euclidean distance between an individual consumption curve and the average consumption of all consumers.
- Consumption Curve Slope: Measure of the slope of the best-fit line to the consumption curve time series.
- Principal Component Analysis (PCA): Derived from Principal Component Analysis or its "kernelized" counterpart (most recent publications). A selection of these components can be used.
- Fractional Order Dynamic Errors: These characteristics reflect the variations between a meter and real-time consumption records.

 doi:10.20944/preprints202507.2247.v1

- Miss adjustment Rate: Quantifies the divergence between measurements at the MV/LV transformer and the sum of measurements from smart meters and estimated technical losses, all normalized by the nominal power of the substation.
- Seasonal Consumption Ratios: Compare energy consumption in different seasons or consumption relative to the average of consumers at the same substation in a specific season.
- Discrete Cosine Transform Coefficients: Involve the first k coefficients of this transformation.
- Percentage Change in Consumption: Represents an x% reduction in consumption during a period T compared to a previous interval or relative to the average.
- Estimated Records: Refer to the number of records made by estimate, in the event of inaccessibility to the meter.

2.2. Performance Metrics Used in Non-Technical Loss Detection

The first seven metrics (accuracy, detection rate, and precision, false positive rate (FPR), true negative rate (TNR), false negative rate (FNR), F1 score) are frequently used in classification tasks, calculated from the confusion matrix. In the literature on NTL detection, the most common metrics are accuracy and detection rate, which are commonly used in almost all data-driven methods. An increase in accuracy indicates that the system generally performs well, correctly classifying both positive and negative samples. However, this is not sufficient when dealing with an imbalanced dataset (typical in NTL detection), where one class (negatives, i.e., benign users) is excessively larger than the other (positives, i.e., fraudulent users).

The second most commonly used metric is the detection rate (DR), which has already been mentioned above, and is also known in the literature as recall, true positive rate, NTL detection success, or hit rate. This metric expresses the proportion of samples classified as NTLs relative to the total number of NTLs in the dataset. Typically, high DR values indicate a well-performing detection system; however, this is not always the case. Other metrics must also be considered to ensure this is true. In general, both DR and accuracy should be considered when evaluating system performance. The following two most commonly used metrics are precision and false positive rate (FPR). Precision, also known as positive predictive value (PPV), assertiveness, or confidence, is calculated as the number of detected NTLs divided by the total number of NTL alarms.

Appropriate metric selection is crucial, especially in scenarios characterized by class imbalance, as observed in NTL detection. In these contexts, it is essential to consider a combination of different metrics, encompassing accuracy, DR, FPR, and TNR. Notably, a less prevalent metric in the NTL detection-specific literature is the Bayesian detection rate (BDR) [13]. This case is predominantly applied in data-driven NTL detection systems, and its formulation depends on variables such as the fraud probability P(I), DR, and FPR. In essence, the BDR quantifies the chance of a false alarm occurring under operational conditions. Both DR and FPR are intrinsic to the classifier in use, while the fraud probability is an exogenous metric. In the fields of fraud and intrusion detection, this probability tends to have low values, reflecting the infrequent nature of fraud. With a small value for P(I), such as 1%, to obtain a high BDR (i.e., reduce the incidence of false alarms), the FPR must reach extremely low levels, even if the DR is high [14].

- Accuracy: (TP + TN) / (TP + TN + FP + FN) (2.3.1)
- Detection Rate (DR): TP / (TP + FN) (2.3.2)
- Precision: TP / (TP + FP) (2.3.3)
- False Positive Rate (FPR): FP / (FP + TN) (2.3.4)
- True Negative Rate (TNR): TN / (FP + TN) (2.3.5)
- False Negative Rate (FNR): FN / (FN + TP) (2.3.6)
- F1 Score: 2TP / (2TP + FP + FN) (2.3.7)
- AUC (Area Under the Curve): The area under the ROC (Receiver Operating Curve) of the binary classifier; (2.3.8)
- Recognition Rate: 1 - 0.5 x (FP/N + FN/P) (2.3.9)
- Bayesian Detection Rate: P(I) x DR / (P(I) x DR + P(¬I) x FPR) (2.3.10)

 doi:10.20944/preprints202507.2247.v1

- Support: Refers to rule-driven systems. Represents the proportion of data to which a specific rule is applicable, relative to the total data set.
- Training Time (s): Represents the amount of time, in seconds, required to train an NTL detection model.
- Classification Time (s): Denotes the time, in seconds, it takes a NTL detection system to classify a single instance.
- Cost of Undetected Attack: Quantifies the financial impact of the most damaging attack that was not identified by the system.
- Energy Balance Mismatch: Corresponds to the discrepancy between the total energy consumed at the user level and the energy recorded at the substation.
- Average Bill Increase: Indicates the increase in the average bill if NTL losses were shared among all consumers.
- Normalized Labor Cost: Estimates the costs associated with inspecting all instances categorized as NTL by the system.
- Anomaly Coverage Rate: Defines the fraction of anomalous consumers under the supervision of an RTU compared to the total number of anomalous consumers.
- RTU (Remote Technical Unit) Cost: Represents the total costs for implementing and maintaining an RTU.
- Minimum Deviation Detected: Specifies the smallest deviation from a predetermined standard that can be identified by the system.
- Reduction in Stolen Electricity: Quantifies the reduction in the volume of illicitly appropriated electricity when implementing a specific FDS.

2.3. Algorithms Used in Non-Technical Loss Detection Systems

Each fraud detection approach has its own peculiarities, based on different data sets and specific algorithms. The complexity of these systems can vary substantially. Some may be based on simple structures, such as using a single Support Vector Machine (SVM) [15] to classify consumers. In contrast, other approaches may adopt more complex systems, which include preliminary steps of data cleaning and clustering, use sets of classifiers, and even perform an in-depth analysis of the energy system.

Despite the specific nuances of each technique, it is possible to identify, at its core, a limited set of algorithms that outline the basis of each fraud detection method. Many of these algorithms are already widely recognized and detailed in the scientific literature. Therefore, we will not delve into their descriptions in this context.

As mentioned previously, NTL detection methods can be categorized according to their primary approach: whether they are data-driven, network-focused, or adopt a hybrid strategy, combining elements of both approaches. It is worth noting that there are many publications addressing NTL, using various classical methodologies. More recently, publications based on machine learning techniques have emerged, e.g., Silveira [16] conducted a study on NTLs using an Fuzzy-ARTMAP ANN, among other publications. This research will focus on NTLs based on ANNs from the ART family, incorporating continuous learning. The foundation of this approach is the behavioral analysis of electricity consumers, contemplating the discovery of new fraud modus operandi. It also includes other areas of interest in this type of problem, for example, the distribution of drinking water to the population, the banking network, credit card operations, etc.

2.3.1. Data-Oriented Methods

Data-driven methods focus primarily on the thorough examination of datasets, making extensive use of techniques from the field of Machine Learning [5]. These approaches can be subdivided into two essential categories: supervised and unsupervised, which will be detailed in subsequent chapters.

 doi:10.20944/preprints202507.2247.v1

- Data processing and model selection: From a raw dataset, the appropriate model for NTL detection needs to be selected. The presence (or absence) of previously labeled data influences the decision between supervised and unsupervised methods. Additionally, the quality and diversity of the data influence algorithm selection. This selection may eventually disregard parts of the raw dataset, as configured in the data selection step. Subsequently, there is the data cleaning step, a common practice in knowledge discovery, followed by feature extraction, if applicable.
- Modeling: The modeling approach varies depending on whether it is supervised or unsupervised. Unsupervised models do not use labeled data during training, using it only for evaluation purposes. Supervised methods, on the other hand, segment the dataset into training and testing sets. Once the training set is established (usually through cross-validation), feature selection is often employed in the training phase. Simultaneously, parameter optimization uses metrics that can be determined based on the availability of labels in the data.
- Application: New data, which are not part of the original "Raw Data" set, are used to evaluate the effectiveness and operability of the model in question. The classification results are then processed to generate a list of potential offenders—that is, a list with the associated probability of each consumer committing fraud. This step can be related to the testing phase of the NTL detection model or its simulation.

The development of classifiers, whether supervised or unsupervised, is based on established Artificial Intelligence (AI) techniques [5]. Such approaches are widely discussed in the literature. The subsequent sections will focus specifically on the application of these methods to the challenge of NTL detection. Several recent studies have explored approaches based on artificial intelligence and data mining techniques for detecting non-technical losses. These studies examine various supervised and unsupervised methods, applying them to real or synthetic datasets to detect anomalous consumption patterns. Below, we present relevant contributions from the literature that illustrates the practical application of these methodologies in different contexts.

In the publication by Quinde, Rengifo, and Vaca-Urbano [17], the authors employed three distinct methods (hierarchical, K-means, and K-medians) to detect NTLs using the daily consumption curves of advanced metering instruments (AMI). These methods were applied to artificial (synthetic) data produced using a Gaussian hidden Markov model. This data represents a typical pattern of residential demand among users in the city of Guayaquil, Ecuador. The performance of these algorithms was analyzed based on their ability to identify incongruent consumption profiles, resulting in the detection of around 68% of NTLs.

Another experiment was conducted using different clustering methods in the context of NTLs. In this case, algorithms based on centroids, density, spectra, and statistical distributions were employed to analyze data related to Type I and Type II line losses. Three main markers were removed: mean line loss, line loss coefficient of variation, and ammeter tripping records. The results indicated that the algorithm used presented the best performance for detecting NTLs in 10kV distribution systems [18].

Badawi et al. [19] proposed a solution to the NTL problem based on data-driven techniques, which is divided into five distinct models. Among the proposed algorithms, the "Gradient Boosting Machine" [20] stands out. This method transforms time series and extracts features from a database provided by the State Grid Corporation of China, comprising finite difference sums, autoregressive integrated moving averages, and the Holt-Winters model, to enhance detection performance.

There is a methodological proposal in the context of Industry 4.0 aimed at detecting NTLs through the use of techniques from machine learning [5] and Big Data. Other algorithms such as decision trees, random forests, support vector machines, and neural networks can also be used. These algorithms utilize the extensive data collected in smart grids to identify patterns and markers of fraud, such as unauthorized connections or meter tampering, thereby enhancing detection and prevention efforts [21].

 doi:10.20944/preprints202507.2247.v1

Finally, [22] proposed the combined use of convolutional ANNs (CNNs) [23] with data retrieval techniques to solve the problem of detecting NTLs. The experiment is performed using state-of-the-art pre-trained CNNs, such as VGG16, to obtain detailed features of electricity consumption time series represented as images.

2.3.2. Network-Oriented Methods

Grid-oriented methods rely on data collected from distribution network sensors, particularly smart meters. They capitalize on the physical laws that govern the electrical grid with the primary goal of identifying fraud. In other words, by understanding how the electrical grid is expected to operate under normal conditions, it is possible to identify anomalies or activities that deviate from this expected operation, indicating potential fraud. These methods are not limited to sensor data alone but also incorporate information related to the grid structure, such as the topology and connectivity between transformers and consumer phases.

Several studies employ energy flow tools to size NTLs and locate their origin, providing a counterbalance to the energy balance when an observer meter is available. Additionally, several techniques are based on estimating the distribution state and identifying atypical or inconsistent data. Although these approaches tend to offer greater accuracy, their implementation is not always feasible due to various limitations.

Some proposals suggest the use of specific sensors for fraud identification. Within this scope, algorithms have been developed to determine the minimum number of sensors and their optimal locations on the network, thereby maximizing the effectiveness of irregularity detection.

The following are several proposals from the literature that exemplify the application of network-oriented methods for detecting non-technical losses:

This article proposes a network-oriented approach to detect non-technical losses in distribution systems using smart meter (SMI) data. The method combines power flow studies and state estimation to identify and locate NTLs, allowing for the determination of inspection zones. The methodology processes network and topology data to recreate power flows, detecting inconsistencies via voltage drops to locate illegal connections or faulty meters [24].

Another proposal uses smart meters to develop a new data analytics technique for detecting and locating non-technical losses. This technique is based on analyzing anomalous data, similar to state estimation methods. It employs the Weighted Least Squares (WLS) formulation and concepts from anomalous data analysis, based on power grid modeling and analysis, to identify measurement inconsistencies [25].

The following article presents an approach for estimating non-technical losses (NTPs) in distribution feeders. The methodology integrates data from customer billing statements and active and reactive power measurements at the substation coupling point. The proposed method employs a modified version of the Power Summation Method (PSM) for load flow, which adjusts the data to ensure power balance for any operating condition. It involves incorporating the calculation of NTPs into operational planning routines via a load flow algorithm, enabling the assessment of their impact on the voltage and reactive power profile of the power grid [26].

An additional load flow-based method for detecting and locating non-technical losses (NTPs) in distribution systems uses data from smart meters. This approach, called the QV method, focuses on identifying illegally connected loads and requires measurements of voltage magnitude, active power, and reactive power. For its application, load buses are modeled as QV buses, and the differences between measured and calculated active power indicate possible NTL locations. The method has been tested on unbalanced distribution systems, showing promising results, especially in secondary distribution systems [27].

Finally, this approach is notable for being exclusively sensor-based (data-driven) and "model-less," as it estimates grid parameters directly from smart meter data, without requiring a complete pre-existing grid model. The process involves estimating voltage sensitivity coefficients, calculating actual customer consumption based on voltage measurements and estimated coefficients, and

 doi:10.20944/preprints202507.2247.v1

flagging fraudulent customers by comparing measured and estimated consumption against a threshold [28].

3. Methodology Based on the Use of The Fuzzy-Art Neural Network

This section outlines the methodology developed and applied to identify non-technical losses (NTLs) in actual electricity consumption data. This approach uses the Fuzzy-ART ANN [1]. Subsequently, the authors proposed a series of alternative algorithms to solve specialized real-world problems. This proposal constitutes an embryonic version, focusing on exploring the potential of this neural network in this type of problem. It is an unsupervised clustering technique (pattern recognition search) based on the principles of adaptive resonance theory (ART) [1]. The choice of this methodology is justified by its intrinsic capacity for adaptive and incremental learning (continuous, incremental training), allowing the detection of atypical behavior patterns without the need for supervision or prior knowledge of the types of fraud.

3.1. Construction and Characterization of the Database

The electricity consumption data used in this study come from the Electricity Smart Metering Technology Trials (ESMIT), a robust and comprehensive initiative led by ESB Networks under the auspices of the Commission for Energy Regulation Smart Metering Project in Ireland. This project was conceived and established by the Commission for Energy Regulation with the primary objective of facilitating the practical learning and validation of intelligent metering systems in a real-world environment.

The technology trials, in particular, sought to gain a deeper understanding of the provision of infrastructure and support systems for smart metering.

This experiment used exclusively a database of consumers characterized as type E for essential methodological reasons. It is observed that the tariff does not exhibit any hourly or seasonal variations. The energy price per kWh remains fixed throughout the day and the week. Therefore, the price is not included in Table 1. The primary objective of this work is to detect NTLs based on electricity consumption profiles. A-D tariffs lead to behavioral changes through the pricing structure, producing artificial and heterogeneous profiles that prevent the identification of factual anomalies.

As part of the Smart Metering Trials conducted in Ireland, residential consumers were allocated to different tariff structures to assess the impacts of dynamic pricing on consumption behavior. The residential tariffs tested were:

- Tariff A, B, C and D:– with different prices depending on the time of day;
- Tariff E:– fixed tariff, used as a control group.

Tariffs A, B, C and D show significant hourly variation, as illustrated in Table 1

Table 1. Time-of-use tariffs (cents per kWh).

Tariff	Nocturnal (23h – 08h)	Diurnal (08h–17h / 19h – 23h)	Peak (17h–19h)
A	12	14	20
B	11	13,5	26
C	10	13	32
D	9	12.5	38

Note: The peak period applies only Monday to Friday, excluding holidays.

In this experiment, a database of consumers classified as type E was used exclusively for essential methodological reasons. It can be observed that the tariff does not exhibit any hourly or seasonal variations. The energy price per kWh remains fixed throughout the day and the week. Therefore, the price is not included in the previous table. The primary objective of this work is to detect NTLs based on electricity consumption profiles. Tariffs A-D lead to behavioral changes through the price structure, resulting in artificial and heterogeneous profiles that hinder the

 doi:10.20944/preprints202507.2247.v1

identification of factual anomalies. By separating only consumers with tariff E, a database with the following advantages is obtained:

- Tariff homogeneity: no hourly or weekly differences;
- Natural consumption profiles: not affected by economic incentives;
- Temporal consistency: essential for detecting fraud or genuine anomalous patterns.

The data set used for each selected consumer comprised a full year of measurements. A one-year period enables the model to capture not only daily and weekly consumption patterns, but also seasonal variations inherent in energy consumption (e.g., increased consumption in winter due to heating systems, lower consumption during vacation periods or extended summer holidays). These variations are natural components of a consumer's load profile and must be incorporated into the model to prevent confusion with degeneracies. Including a complete annual cycle of data also allows the algorithm to "learn" about consumption peculiarities across different days of the week and holidays.

The fundamental unit of analysis adopted in this study corresponds to consumption days. Each sample, therefore, was organized as a vector composed of 48 values. These 48 values represent electricity consumption records collected every 30 minutes throughout a single day. Shorter time intervals can be considered perfectly by including the relevant adjustments. This vector representation of the daily consumption profile is rich in information, allowing the algorithm to capture the subtleties and fluctuations in consumption over 24 hours, from peak demand to periods of lower consumption. The choice of 30 minutes is an optimal balance between the need for sufficient detail to identify fine anomalies (e.g., a power outage for a few hours or a subtle change in load over a specific period) and the computational feasibility of processing large volumes of data. However, much smaller measurement fractions (e.g., 15 or 5 minutes) would generate high-dimensional input vectors, exponentially increasing computational complexity.

3.2. Data Processing and Pre-Processing

The phase prior to data input to the Fuzzy-ART ANN. This is a critical and multifaceted preprocessing step that was rigorously implemented, which is vital to ensuring the integrity, consistency, quality, and proper formatting of daily consumption vectors.

3.2.1. Data Cleaning and Inconsistency Treatment

The initial preprocessing step involves systematically cleaning the data to remove incomplete or inconsistent records. In field-collected consumption data, such as those from the CER Ireland trials, quality issues are common and expected. Such errors would result in daily vectors with fewer than 48 consumption points, indicating gaps in the time series. Removing these defective records is crucial because the presence of noise or corrupted data in the input vectors can lead the clustering algorithm to learn spurious patterns, resulting in:

False positives: Classifying a typical consumption day as fraudulent due to measurement errors, generating unnecessary alarms and costly investigations;

False negatives: Failing to detect actual fraud because the data was so corrupted that the anomalous pattern was masked or mistaken for noise. Therefore, this step ensures that each input vector submitted to Fuzzy-ANN-ART represents a complete, valid, and reliable day of consumption, crucial for analyzing temporal patterns and for the algorithm's effectiveness in identifying anomalies.

3.2.2. Min-Max Normalization

After the cleaning phase, the data were subjected to a Min-Max normalization process. Normalization is a rescaling technique that transforms the values of a feature into a specific standardized range, in this case, between 0 and 1. The formula applied is:

 doi:10.20944/preprints202507.2247.v1

$$X_{normalizado} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3.2.2.1)$$

where:

X: original value of the characteristic,
Xmin: minimum value observed for this characteristic in the entire dataset
Xmax: maximum value.

In this study, the Xmin and Xmax values were calculated exclusively from the training set, respecting the principle of separation between training and testing. This choice prevents information leakage from the test set and, consequently, prevents model contamination, which could compromise the validity of the performance evaluation. Since the test set represents unseen data, any statistics calculated on it should be avoided during the training and preprocessing phases. The choice of Min-Max normalization is justified to ensure equitable contribution from features. For distance- or similarity-based algorithms, such as Fuzzy-ART, it is essential that all features contribute equally to the clustering process. Without the normalization phase, features with very different value scales, such as analog consumption values (which can vary significantly in magnitude) and binary temporal attributes (which are strictly 0 or 1), would cause an imbalance. Consumption values, due to their higher magnitude, would tend to unduly dominate the calculation of distance or similarity, underestimating the relevance of the information contained in the binary attributes.

3.3. Generating Artificial Fraudulent Samples for Training and Testing

The ability to train and validate a fraud detection model is intrinsically limited by the availability of data labeled as fraudulent. In real-world NTL detection scenarios in distribution grids, obtaining a robust and diverse database of real fraud is a significant challenge due to its clandestine nature and the difficulty of identifying and confirming it. Fraud data is often scarce, unbalanced compared to normal data, and may not cover all existing manipulation typologies. To overcome this limitation, fraudulent samples were artificially generated from the actual consumption days of regular customers. This approach enabled the creation of a controlled environment to assess the model's ability to identify anomalous patterns. The artificial fraud generation was performed to ensure that the fraudulent samples maintained characteristics inherent to real consumption data (e.g., daily fluctuations, seasonality), but with controlled distortions that mimicked plausible fraudulent behaviors. The process of generating fraudulent samples followed a well-defined sequence, aiming to simulate a variety of behaviors that fraudsters could employ and that the Fuzzy-ART algorithm should be able to identify:

1. **Shuffling and initial division of the benign base**. Previously, all vectors containing daily consumption data that had been classified as benign (after the cleaning and preprocessing steps) were randomly shuffled. Subsequently, the total set of vectors designated as benign was divided into two parts of exactly equal size. This initial division serves to create a "pool" of data that will be kept as benign and another "pool" that will be labeled as fraud.
2. **Definition of benign samples**. The first half of the original benign vectors remained unchanged. These samples were labeled "0," indicating their normal, as well as expected, consumption status. They are essential for the Fuzzy-ART algorithm to build an accurate representation of what is "normal" in the dataset, serving as a comparison base for identifying anomalies.
3. **Fraud simulation base**. The second half of the original benign vectors was reserved exclusively for fraud simulation. These samples were labeled "1", indicating their fraudulent nature after the application of fair manipulations.

Taking into account the methodology described by [29], the seven simulated fraud typologies were carefully designed to encompass a representative spectrum of changes in the consumption profile, aiming to cover different fraud modes and allowing a comprehensive evaluation of the

 doi:10.20944/preprints202507.2247.v1

Fuzzy-ART ANN performance. The generation process for each fraud class was implemented to reflect plausible manipulation behaviors:

Type 1– Constant consumption reduction: In this scenario, the simulation simulates a tampering that uniformly affects all recorded consumption throughout the day. Each value of the daily consumption vector (consumption in each 30-minute interval) was multiplied by a fixed random factor. This factor was selected uniformly within the range of 0.1 to 0.3, meaning that the original consumption was reduced to 10% to 30% of its true value at each measurement point.

Type 2– Partial measurement interruption: This type of fraud simulated a localized and temporary interruption in the consumption record, which could indicate a bypass or disconnection. A continuous segment of the daily vector was randomly selected at any point during the day. The length of this segment varied between 3 and 12 30-minute intervals (corresponding to a period of 1.5 to 6 hours). The consumption values within this segment were then replaced with zeros, while the rest of the vector remained unchanged. This manipulation represents scenarios involving temporary meter disconnection or an intervention aimed at eliminating consumption records for a specific period of the day.

Type 3– Random point-by-point reduction: Unlike constant reduction (Type 1), this class of fraud induced a more irregular and unpredictable distortion in consumption. For each 30-minute interval of the daily vector, its value was individually reduced by a different random factor, also between 0.1 and 0.3. Each factor was sampled independently, introducing variability at each time point. The result is an asymmetric and non-uniform tampering with consumption, where the usual pattern may be preserved, but with erratic variations at each measurement point.

Type 4– Consumption reduction while preserving the original shape of the consumption profile. In this type of fraud, the objective is to simulate a proportional decrease in daily consumption while maintaining the original morphology of the load profile. To do this, the average daily consumption for each vector is calculated, and then a random reduction factor of between 10% and 30% is applied. All 48 values for the day are then adjusted by a correction factor that reduces the daily average to the new target value. This way, the profile maintains peak and valley times, as well as the relative proportion between consumption intervals, making it visually similar to a real day, but with consistently lowers values.

Type 5– Artificially constant consumption. This type of fraud aims to completely mask the real consumption profile and replace it with an artificially stable pattern. The average of the original vector replaced all values in the daily vector. If the averages daily consumption was X (kWh), all 48 consumption points were set equal to X. This generated a perfectly flat consumption profile, without any variations throughout the day, which would in reality be highly unlikely for a typical residential consumer, who typically experiences significant fluctuations throughout the day (e.g., consumption peaks in the morning, afternoon, and evening).

Type 6– Temporal Inversion. This class of fraud exploits the temporal order of consumption data, a characteristic that is essential and predictable in real residential demand profiles. The 48 values of the daily vector were inverted, which means that consumption recorded in the early hours of the day is swapped with consumption in the later hours, and vice versa. For example, early morning consumption (which is usually low) is swapped with midday or evening consumption (which is generally high). This maneuver was designed to exploit potential vulnerabilities in pricing or recording schemes that strictly depend on the time of use. A complete inversion in the profile is highly anomalous and would not represent the actual consumption behavior of a domestic user; therefore, it is a clear indicator of manipulation.

 doi:10.20944/preprints202507.2247.v1

Type 7– Substitution with random minimum values. In this scenario, the fraud aims to drastically reduce the energy bill, but with sufficient complexity to evade more basic detection systems. For each value in the daily vector, it was replaced with a random number between zero and the minimum value observed in the original consumption vector for that day. The result of this manipulation is an artificially low consumption profile, but one that still presents minor random variations, simulating a severely underestimated reading. It could represent fraud where the meter is constantly forced to register values very close to zero, with minor fluctuations, to try to appear more conventional.

To ensure that the anomaly detection model was not biased by the predominance of the majority class (benign days), which is a common situation in real NTL databases; a controlled database balancing strategy was adopted from the outset. This balancing was not performed as a later step, but rather incorporated directly into the fraud characteristic data generation process. To this end, the database preparation followed the following logic:

- For each fraud typology (Types 1 to 7): The original benign database, consisting of 7,000 consumption days, was divided into two equal parts. One half (3,500 rows) was retained to represent the benign days, while the other half (the remaining 3,500 rows) was used to generate the fraudulent samples specific to that typology. After generation, the benign case base used to create the fraudulent cases was discarded, resulting in a final database for each fraud typology composed of 3,500 benign days and 3,500 fraudulent days of the same type.
- For the general set: This set, which represents a balanced combination of all seven fraud types, was created similarly. A portion of the 3,500 benign days was retained, and for the fraudulent days, 500 new instances were generated for each of the seven typologies, totaling 3,500 fraudulent days.

With the database properly balanced and the consumption vectors formatted and supplemented, they were then provided as input to the Fuzzy-ART algorithm. This care ensured that the model had a sufficient and representative number of fraud examples to learn their patterns and perform effective, unbiased clustering.

3.4. Configuration and Operation of the Fuzzy-ART Algorithm

The foundation of this NTL detection methodology lies in the Fuzzy-ART algorithm, an artificial intelligence that "learns" to adaptively group data. This algorithm doesn't require pre-classified fraud examples to begin working, which is a significant advantage, as fraud data is challenging to obtain. The Fuzzy-ART ANN stands out for its ability to find hidden patterns and structures in energy consumption profiles, both standard and fraudulent. Three parameters control the way the Fuzzy-ART ANN operates:

Choice parameter (α): Defines the initial selection of the cluster most similar to a new sample, in this case, influencing the affinity calculation before the vigilance test (ϱ). Higher values tend to favor larger or more general existing clusters, while lower values prioritize exact similarity, which can lead to the creation of more clusters.

Vigilance parameter (ϱ): This parameter can be thought of as a "similarity criterion," resonance, or tuning. It ranges from 0 to 1 and defines how similar a new consumption reading must be to an existing cluster to be added to it. A high value (close to 1) means the algorithm is very demanding: it will only accept a new reading in a cluster if it is rigorously similar to the patterns that the cluster already represents. This results in many small, particular clusters, each capturing a very distinct type of behavior. A low value (close to 0) makes the algorithm more flexible, as it will accept new readings even if they are significantly different from the cluster, resulting in fewer, but larger and more comprehensive clusters. Choosing an appropriate ϱ is crucial for

 doi:10.20944/preprints202507.2247.v1

Fuzzy-ART to efficiently distinguish between fraudulent and standard consumption patterns without creating excessive data fragmentation.

Learning parameter (β): This parameter, β, is thought of as the model's "adaptation speed." It also ranges from 0 to 1 and determines how quickly existing groups adapt to new patterns they encounter:

A value close to 1.0 causes the model to learn quickly. Groups adapt intensely to each new reading they absorb.

A lower value results in slower learning and more stable groups. Groups are less influenced by a single reading, becoming more representative of the average of all readings in that group. In certain situations, a lower β is preferable to prevent the model from being overly influenced by atypical or "noisy" readings. The logic behind the Fuzzy-ART ANN involves a process of continuous comparison and adjustment.

- A value close to 1.0 allows the model to learn quickly. Groups adapt intensely to each new reading they absorb.
- A lower value results in slower learning and more stable groups. Groups are less influenced by a single reading, becoming more representative of the average of all readings in that group. In certain situations, a lower β is preferable to prevent the model from being overly influenced by atypical or "noisy" readings. The logic behind the Fuzzy-ART ANN involves a process of continuous comparison and adjustment.

When a new consumption reading (already prepared and standardized) is presented, the algorithm processes according the Fuzzy-ART algorithm presented on Appendix A.

Appendix A presents the Fuzzy-ART algorithm in detail. This architecture allows the Fuzzy-ART ANN to continuously learn and detect abnormal patterns without needing to be "taught" about all types of fraud. It has been extremely valuable for dealing with fraud, which is constantly evolving and requires a dynamic and autonomous detection system. To ensure that the Fuzzy-ART ANN performs at its maximum efficiency in detecting NTLs, it is essential to find the "best configuration" for its parameters ϱ and β, given that the parameter α has been fixed. It has been achieved through a process known as hyperparameter search. In this study, an exhaustive search was performed, testing a wide range of values for ϱ (from 0.60 to 0.99) and β (from 0.1 to 1.0), with α set to 0.005. Combining all the possibilities resulted in 400 different configurations for the model. To make this process efficient, parallel computing resources were utilized, meaning that instead of testing the 400 configurations one by one in sequence, the system was able to test many of them simultaneously, using multiple processing cores. This approach drastically accelerated the time required to find the optimal combination of ϱ and β for the aforementioned α value. After testing all 400 combinations, the one that performed best in identifying fraud was selected. This optimized Fuzzy-ART ANN configuration was then used to evaluate the final performance of our NTL detection system, providing a clear picture of its effectiveness.

3.5. Model Performance Evaluation

Evaluating the performance of an anomaly detection model, such as the Fuzzy-ART ANN applied to NTL detection, requires the use of a comprehensive set of metrics. Although the Fuzzy-ART ANN is an unsupervised clustering algorithm, its effectiveness in fraud detection can be quantified a posteriori by analyzing cluster composition, in terms of the agreement between benign and fraudulent samples. This inference endows the model with its intrinsic ability to differentiate suspicious profiles based solely on the morphology of daily consumption vectors, without requiring prior knowledge of the labels during training. To quantify and compare the performance of the fraud detection model, the following metrics were used. These metrics are qualifying classification scores calculated on the test set using the predictions generated by the model with the best hyperparameters found:

 doi:10.20944/preprints202507.2247.v1

Accuracy: This parameter measures the total proportion of correct predictions (benign samples correctly classified as benign or fraudulent samples correctly classified as dishonest). This parameter is calculated as follows: Accuracy (Equation (2.31)), where: TP (True Positives) are the fraud cases correctly identified, TN (True Negatives) are the normal days correctly identified as usual, FP (False Positives) are the normal days incorrectly classified as fraud, and FN (False Negatives) are the fraud cases that were not detected. Although it is an intuitive and widely used metric, in imbalanced datasets (where the fraudulent class is a minority), high accuracy can be misleading if the model classifies most samples as the majority class. However, in this experiment, the dataset was intentionally balanced to minimize this problem during training and evaluation. The following metrics have been widely used and accepted as indicators of the quality of results and experiments. They are, therefore, compiled within the context of confusion matrix theory (DUDA & STORK, 2012).

Sensitivity (Recall): This parameter, known in the literature as the True Positive Rate or Recall, infers the proportion of fraudulent samples that were correctly identified as fraudulent by the aforementioned metric relative to the total number of actual fraudulent samples present in the test set. It is calculated using the following equation: Sensitivity = TP/(TP+FN). It is a crucial metric for assessing fraud detection problems, where identifying as much fraud as possible is a high priority, even if it results in an increase in false positives. It should be noted that a high sensitivity value is desirable to prevent real fraud from going undetected, thereby resulting in continued financial losses for both the utility and consumers who comply with their contractual obligations.

Specificity: Infers the proportion of benign samples that were correctly identified as benign by the model, compared to the total number of real benign samples. It is calculated as: Specificity = TN/(TN+FP). This metric is essential to complement sensitivity, as it indicates the model's ability to avoid false positives. False positives can lead to unnecessary inspections, increased operational costs, and ultimately, inspector fatigue, which compromise efficiency and confidence in the system. It is worth noting that high specificity indicates that the model is unlikely to classify a standard sample as fraudulent.

Matthews Correlation Coefficient (MCC): This is an inference metric for evaluating the quality of binary classifiers, being more robust than accuracy and F1-score, especially in imbalanced datasets. The MCC ranges from -1 to +1, where +1 indicates a perfect prediction, 0 indicates a random prediction (no better than chance), and -1 indicates an inverse prediction (completely incorrect prediction). This coefficient takes into account all four components of the confusion matrix (TP, TN, FP, FN). It is a more balanced and reliable metric for evaluating the overall performance of the classifier in cases of imbalanced class problems, offering a more honest assessment of how well the model distinguishes between the two classes.

4. Results Obtained

A comparative analysis between input scenarios consisting of 48 data points related to daily consumption, designated type "C," and consumption, also daily, considering the binary temporary attributes "C+D," reveals that the addition of binary date-related variables often did not yield consistent performance gains for NTL detection. The database used was the one mentioned above, which was obtained from smart meter trials conducted by the Commission for Energy Regulation (CER) of Ireland. It consists of the daily consumption of benign consumers spread over 48 half-hour records throughout an entire year. These records were modified, as mentioned above, to generate the seven types of fraudulent consumers. The metrics for the overall set (which, as mentioned, represents the balanced combination of the seven fraud types) indicate that accuracy and specificity decreased with the inclusion of these attributes in "C+B" (from 0.8100 to 0.7529 and from 0.7819 to 0.6895, respectively). At the same time, sensitivity remained relatively stable, and the overall MCC decreased significantly (from 0.6210 to 0.5098). The most significant difference is observed in Type 2, where the inclusion of binary features leads to a substantial reduction in specificity (from 0.9133 to 0.5724) and MCC (from 0.7563 to 0.3571). It is suggested that, for certain types of fraud, the addition of explicit

 doi:10.20944/preprints202507.2247.v1

temporal information may, in some cases, impair the model's ability to distinguish between them, potentially introducing unnecessary noise or complexity into the clustering and classification task.

Detailed results of model performance for different fraud types and input scenarios are presented in Table 2.

Table 2. Performance metrics by fraud type considering scenarios with consumption only (C) and with consumption plus temporal attributes with binary coding (C+B).

Type	Accuracy		Sensitivity		Specificity		MCC		Created Clusters	
-	C	C+B	C	C+B	C	C+B	C	C+B	C	C+B
1	0.8767	0.8743	0.9476	0.9629	0.8057	0.7857	0.7610	0.7606	59	107
2	0.8771	0.6748	0.8410	0.7771	0.9133	0.5724	0.7563	0.3571	12	169
3	0.8790	0.8829	0.9019	0.9143	0.8562	0.8514	0.7589	0.7672	172	132
4	0.8733	0.8610	0.9257	0.9295	0.8210	0.7924	0.7508	0.7288	64	86
5	0.9495	0.9119	0.9914	0.9048	0.9076	0.9190	0.9022	0.8239	879	176
6	0.8824	0.7586	0.8248	0.7495	0.9400	0.7676	0.7699	0.5172	1126	720
7	0.9533	0.9357	0.9781	0.9895	0.9286	0.8819	0.9078	0.8765	213	124
All	0.8100	0.7529	0.8381	0.8162	0.7819	0.6895	0.6210	0.5098	155	352

5. Conclusions

This research focused on investigating, understanding, and developing methodologies for non-technical electrical losses, with a particular emphasis on individual consumers. As a rule, by law, these losses are passed on (in whole or in part) to the energy bills of customers who pay their bills. Covering amounts of energy they did not consume. Therefore, it is imperative to identify fraudsters and hold them accountable for these errors. Benefiting society and democratically reducing tariffs. Thus, these systems were proposed to "discover" the behavior of electricity customers (residential, commercial, and industrial) linked to an electricity distribution company. Identifying fraudsters as accurately as possible is also part of this objective. This system was developed using techniques from the machine learning context, specifically. Neural networks from the ART family, designed by Stephen Grossberg (cognitive scientist, psychologist, computational theorist, neuroscientist, mathematician, biomedical engineer, and neuromorphic technologist) and Gail Carpenter (cognitive scientist, neuroscientist, and mathematician). This partnership resulted in a series of proposed ANNs aimed at solving a variety of real-world problems. The main contributions are: ART1, ART2, ART2-A, ART3, ARTMAP, ART Predictive, Fuzzy-ART, Fuzzy-ARTMAP, Gaussian-ART, Gaussian-ARTMAP, Fusion-ART, Topo-ART, Hypersphere-ART, Hypersphere-ARTMAP, and LAPART. There may be a few other modules that are little-known. It should be noted that the acronym ART refers to the unsupervised training modality, while "MAP" stands for supervised ANN. This abundance of alternatives motivated this research, as well as the high-quality alternatives offered, enabling even greater contributions toward improving the quality of solutions. The study of this duo is based on understanding the human mind, as highlighted, in particular, by Grossberg's publications [2,30], while also taking into account many other previous studies. Therefore, to infer the quality of solutions, databases are needed for training. In this case, a database provided by an Irish electricity distribution company was used. Later, the research began using Fuzzy-ART ANN, primarily due to its highly valuable scientific properties: stability, rapid response, and, above all, plasticity. Plasticity is almost an exception among the various proposals in the specialized literature. Plasticity is the ability to undergo incremental training (continual, continuous, eternal learning); ideally, it is an innate attribute. Once this was done, the Fuzzy-ART ANN was implemented, taking into account the specific characteristics of the research problem at hand—the study of non-technical losses in electric system distribution. The results are encouraging, as illustrated in Table 2. These results were inferred using a modern metric based on the theory of confusion matrices. A comparative study with the literature was not conducted, as the purpose of this research focused on a topic that, unless I'm mistaken, is distinct from the literature. In this ANN, a large number of alternatives need to be

 doi:10.20944/preprints202507.2247.v1

explored and developed to improve results and meet various demands, based on more effective parameter tuning and experimentation.

Author Contributions: S.F.C.: Conceptualization, Writing—Original Draft, Writing—Review and Editing, Investigation. R.J.D.S.: Conceptualization,Writing—Original Draft,Writing—Review and Editing, Investigation. C.R.M.: Writing—Review and Editing, Funding Acquisition, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study were collected from real customers of the Commission for Energy Regulation (CER) of Ireland.

Acknowledgments: The authors are grateful for the financial support of Brazilian National Council for Scientific and Technological Development (CNPq) under grant 302896/2022-8 and Coordination for the Improvement of Higher Education Personnel (CAPES), grant UNESP/PROPG 37/2023.

Conflicts of Interest: The authors declare that they have no competing interest.

Appendix A. Fuzzy-Art Neural Networks

The algorithm for the training phase is presented below. The ART neural network is composed of three layers: F0 (input layer), F1 (comparison layer), and F2 (recognition layer that stores categories (clusters)). The algorithm for this neural network consists of the following steps [2]. When used to perform a specific task, a step related to normalizing the input data is usually included at the beginning of the processing. However, according to the proposal of this research, the processing of the input and output vectors implicitly satisfies the sine qua non condition of fuzzy logic: that is, all components of the ANN's input-output pair lie in the interval [0,1].

Step 1: Input Data

The input data is denoted by the vector a = [a1 a2 . . . aM] of dimension M. This vector is normalized to avoid the proliferation of categories. Thus:

$$\underline{\mathbf{a}} = \mathbf{a} \,/\, \|\mathbf{a}\| \tag{A.1}$$

where:

$\underline{\mathbf{a}}$: normalized input vector;

$$\|\mathbf{a}\| = \sum_{i=1}^{M} |a_i| \tag{A.2}$$

$\|.\|$: 1-norm of a vector.

Once Step 1 is completed, to simplify the notation, proceed as follows:

$$\mathbf{a} \leftarrow \underline{\mathbf{a}} \tag{A.3}$$

Step 2: Input Vector Encoding

Complement encoding is performed to preserve the amplitude of the information, that is, the norm (norm1) has the same size for the entire set of vectors in the training and diagnostics:

$$a_i^c = 1 - a_i \tag{A.4}$$

where:

$\mathbf{a}^c$: normalized complementary input vector

$$= [a_1^c \; a_2^c \; a_3^c \ldots a_M^c].$$

Therefore, the input vector will be a 2M-dimensional vector, denoted by:

$\mathbf{I} = [\mathbf{a} \; \mathbf{a}^c]$

$$= [a_1 \; a_2 \; a_3 \; \ldots \; a_M \; a_1^c \; a_2^c \; a_3^c \ldots a_M^c] \tag{A.5}$$

 doi:10.20944/preprints202507.2247.v1

$\|\mathbf{I}\| = M$.

Therefore, all vectors with normalization and complemented encoding will have the same length M.

Step 3: Activity Vector

The activity vector of F2 is symbolized by y = [y1 y2 . . . yL], where L is the number of categories created in F2. Thus, we have:

$$y_j = \begin{cases} 1, & \text{If node } j \text{ of F2 is active;} \\ 0, & \text{otherwise.} \end{cases} \quad (A.6)$$

Step 4: Network Parameters

The parameters used in the processing of the ART-Fuzzy network are:

4. Choice Parameter: $\alpha > 0$;
5. Training rate: $\beta \in [0,1]$;
6. Vigilance Parameter: $\rho \in [0,1]$.

Step 5: Weight Initiation

Initially all weights have a value equal to 1, that is:

$$w_{j1}(0) = \ldots = w_{j2M}(0) = 1 \quad (A.7)$$

indicating that there is no active category.

Step 6: Choosing the Category

Given the input vector **I** in F1, for each node j in F2, the choice function Tj is determined by:

$$T_j = \|\mathbf{I} \wedge \mathbf{W}_j\| / (\alpha + \|\mathbf{W}_j\|) \quad (A.8)$$

Being:

$\wedge$: Fuzzy operator AND defined by:

$$(I \wedge \mathbf{W})_i = \min (I_i, w_i),\ i = 1, 2, \ldots, 2M. \quad (A.9)$$

The category is chosen as being the active node (neuron) J (in fact it is a candidate for the condition of active neuron), that is:

$$\Gamma = \arg \{\max (Tj)\} \quad (A.10)$$

j = 1, 2, ..., L.

Using equation (A.10), if there is more than one active category, the category chosen will be the one with the lowest index. The index J obtained from equation (A.10) is only a candidate for indicating the winning neuron, viz., Γ is a temporary winning neural index. If it is confirmed as the winning neuron, it is labeled as J. Final confirmation will occur after passing the vigilance test (A.8) onwards.

Step 7: Resonance or Reset

Resonance occurs if the vigilance criterion (A.10) is satisfied:

$$\{\|\mathbf{I} \wedge \mathbf{W}_J\| / \|\mathbf{I}\|\} \geq \rho \quad (A.11)$$

If the criterion defined by relation (A.11) is not satisfied, the reset device must be activated. During the reset, node J of F2 is excluded from the search process given by (10), that is, TJ = 0. Then, a new category is chosen using equation (A.10) for the resonance process. This procedure will be performed until the network finds a category that satisfies inequality (A.11).

Step 8: Weight Update (Training)

After input vector **I** have reached resonance, the training process continues, modifying the weight vector given by:

$$\mathbf{W}_J^{new} = \beta (\mathbf{I} \wedge \mathbf{W}_J^{old}) + (1 - \beta) \mathbf{W}_J^{old} \quad (A.12)$$

where:

 doi:10.20944/preprints202507.2247.v1

J: winning active category.
$\mathbf{W}_J^{new}$: updated weight vector;
$\mathbf{W}_J^{old}$: weight vector referring to the previous update.
If $\beta = 1$, there is rapid training.

References

1. Carpenter, G.A. and Grossberg, S. "A self-organizing neural network for supervised learning, recognition and prediction", IEEE Communications Magazine, Vol. 30, No. 9, pp. 38–49, 1992.
2. Grossberg, S. "Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world", Neural Net. Vol. 37 2013, pp. 1-47. Doi: 10.1016/j.neunet.2012.09.017.
3. Carpenter, G.A., & Grossberg, S. (2016). "Adaptive resonance theory", Springer.Doi: 10.1007/978-1-4899-7502-7_6-1.
4. Jené-Vinuesa, M.; Aragüés-Peñalba, M.; Sumper, A. "Comprehensive data-driven framework for detecting and classifying non-technical distribution losses", IEEE Access, Vol. 1, 2024. Doi: 10.1109/access.2024.3426940.
5. Haykin, S. "Neural networks and learning machines", 3. ed. Upper Saddle River: Prentice-Hall, 2008.
6. Brazilian Senate Agency. "CTFC Approves limits on the inclusion of non-technical losses in electricity bills, Senate News, Nov-2021. (In Portuguese).
7. Marchiori, S. C.; da Silveira, M.C; Lotufo, A.D.P, Minussi, C.R. and Lopes, M.L.M. "Neural network based on adaptive resonance theory with continuous training for multi-configuration transient stability analysis of electric power systems", Applied Soft Computing, Vol. 11, No. 1, Jan-2011, pp. 706-715. Doi. 10.1016/j.asoc.2009.12.032.
8. Zadeh, L. A. "Fuzzy sets", Information and Control, 1965, Vol. 8, No. 3, p. 338-353. Doi: 10.1016/S0019-9958(65)90241-X.
9. Messinis, G. M.; Hatziargyriou, N. D. "Review of non-technical loss detection methods", Electric Power Systems Research, 2018, Vol. 158, 2018, pp. 250–266. Doi: 10.1016/j.epsr.2018.01.005.
10. Faria, L. T.; Melo, J. D.; Padilha-Feltrin, A. "Spatial-temporal estimation for nontechnical losses", IEEE Transactions on Power Delivery, 2016, Vol. 31, No. 1, pp. 362–369.
11. Pearson, K. "On lines and planes of closest fit to systems of points in space", Philosophical Magazine, 1901, Vol. **2, No.** 6, pp. 559–572. Doi: 10.1080/14786440109462720
12. Daubechies, I. "Ten lectures on wavelets. Philadelphia: Society for Industrial and Applied Mathematics, 1992, 377 p.
13. Axelsson, S. "The base-rate fallacy and the difficulty of intrusion detection", ACM Transactions on Information and System Security (TISSEC), 2000, Vol. 3, No. 12, p. 186–205.
14. Duda, R. O. and Stork, D.G. "Pattern classification", 2001, 2nd ed., New York: Wiley.
15. Cortes, C. and Vapnik, V. "Support-vector networks", Machine Learning, 1995, Vol. 20, No. 3, pp. 273–297. Doi:10.1007/BF00994018.
16. Silveira, V.G.; Silva-Santos, A.; Lopes, M.L.M.; da Silva, J.F.R. e Faria, L.T. "Detection of non-technical losses via ARTMAP-Fuzzy neural network in electrical energy distribution systems", XXIV Brazilian Congress of Automation, 2022, pp. 1-8. (in Portuguese).
17. Quinde, S.; Rengifo, J.; Vaca-Urbano, F. "Non-technical loss detection using data mining algorithms", IEEE PES Innovative Smart Grid Technologies Conference, Sep. 2021. Doi: 10.1109/ISGTLATINAMERICA52371.2021.9543024.
18. Wang, Z.; Li, G.; Wang, X.; Chen, C.; Huan, L. "Analysis of 10kV non-technical loss detection with data-driven approaches", IEEE Innovative Smart Grid Technologies - Asia, 2019, pp. 4154–4158. Doi: 10.1109/ISGT-ASIA.2019.8881733.
19. Badawi, S. A.; Takruri, M.; Al-Bashayreh, M. G.; Salameh, K.; Humam, J.; Assaf, S.; Aziz, M. R.; Albadawi, A.; Guessoum, D. E.; Elbadawi, I. A.; Al-Hattab, M. "A novel two-stage method to detect non-technical losses in smart grids", IET Smart Cities, 2024, pp. 96-111. Doi: 10.1049/smc2.12078.
20. Breiman, L. "Arcing the edge", Technical Report 486. Statistics Department, University of California, Berkeley, 1997, pp.1-14.

 doi:10.20944/preprints202507.2247.v1

21. Moreno, D. A.; Holguin, M.; Holguín, G. A.; Hernandez, B. “An industry 4.0 based data analytics framework for the detection of non-technical losses in a smart grid”. 2023 IEEE 6th Colombian Conference on Automatic Control (CCAC), 2023, pp. 1–6. Doi: 10.1109/ccac58200.2023.10333569.
22. Esmael, A. A.; da Silva, H.H.; Ji, T.; Torres, R.S. “Non-technical loss detection in power grid using information retrieval approaches: A comparative study. IEEE Access, Vol. 9, pp. 40635–40648, 2021. Doi: 10.1109/ACCESS.2021.3064858.
23. LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; Jackel, L. D. "Backpropagation applied to handwritten zip code recognition", neural computation, 1989, Vol. 1**,** No. 4, pp. 541–551. Doi: 10.1162/neco.1989.1.4.541.
24. Medeiros, M. H.; Sanz-Bobi, M. A.; Domingo, J. M.; Picchi, D. “Network oriented approaches using smart metering data for non-technical losses detection”, IEEE PowerTech Conference, 2021. Doi: 10.1109/POWERTECH46648.2021.9495019.
25. Raggi, L. M. R.; Trindade, F. C. L.; Cunha, V. C.; Freitas, W. “Non-technical loss identification by using data analytics and customer smart meters”. IEEE Transactions on Power Delivery, 2020, Vol. 35, No. 6, pp. 2700–2710, Doi: 10.1109/TPWRD.2020.2974132.
26. Bezerra, U. H.; Soares, T. M.; Nunes, M. V. A.; Tostes, M. E.L.; Vieira, J.P.A.; Agamez, P.; Viana, P. R. A. “Non-technical losses estimation in distribution feeders using the energy consumption bill and the load flow Power Summation Method”, IEEE International Energy Conference, 2016, pp. 1–6. Doi: 10.1109/ENERGYCON.2016.7513947
27. Ferreira, T. S. D.; Trindade, F. C. L.; Vieira, J. C. M. “Load flow-based method for nontechnical electrical loss detection and location in distribution systems using smart meters”. IEEE Transactions on Power Systems, Vol. 35, No. 5, pp. 3671–3681, Sept. 2020. Doi: 10.1109/TPWRS.2020.2981826.
28. Pengwah, A. B.; Razzaghi, R.; Andrew, L. L. H. “Model-less non-technical loss detection using smart meter data”, IEEE Transactions on Power Delivery, Vol. 38, No. 5, Oct. 2023. Doi: 10.1109/TPWRD.2023.3280551.
29. Yeckle, J.; Tang, B. “Detection of Electricity Theft in Customer Consumption Using Outlier Detection Algorithms”. 2018 1st International Conference on Data Intelligence and Security (ICDIS), South Padre Island, TX, USA, 2018, pp. 135–140, Doi: 10.1109/ICDIS.2018.00029.
30. Grossberg, S. “Conscious mind, resonant brain: how each brain makes a mind”, Oxford University Press, Jul-2021, 768 p.