Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Electricity Customers Segmentation: A Mechanism to Curb Non-technical Losses in Power

Hadiza Ali Umar,
Department of Computer Science, Bayero University, Kano
haumar.cs@buk.edu.ng

Ibrahim A. Lawal
Department of Information Technology, *Bayero University, Kano*
ialawan.it@buk.edu.ng

Abstract: *Power utilities globally have a huge task of enhancing their services to enable them to compete in emerging markets. However, power utilities in Nigeria cannot meet their obligations due to challenges of revenue deficits that ensue due to collection losses known as Non-Technical losses (NTL). Utilities and government agencies must continue to place a high premium on assessing the intensity of NTLs to boost revenue generation and enhance the dependability of the power value chain. Using machine learning (ML) techniques, researchers can leverage the massive amounts of data produced by power companies to determine customer trends in energy payment and usage. Segmenting electricity customers with similar characteristics into behaviourally comparable groups is one use for this technique. This study groups customers into segments based on tariff plans and bill payments using the k-means, DBSCAN, and expectation maximisation clustering algorithms. The goal is to segment the customer base identify outliers and.*

Keywords: Non-technical loss, Gaussian Mixture Model, Customer Segmentation

1. INTRODUCTION

Power consumption has increased globally by 5%, estimated to reach more than 33,000 Tera Watt hours (TWh), and is consequently seen as the fuel of the future. Due to its rapid expansion, electricity now accounts for more than 20% of all final energy consumption, which is attributed to the rising population, industrialisation, high living standards, and power output. The International Energy Agency (IEA) predicts that global power demand will increase by 2050 (IEA, 2022). This evaluated increase has left utilities worldwide with no choice but to seek solutions to tackling some very challenging power issues. It has been established severally, that reducing collection losses (non-technical losses – NTL) and increasing revenue generation have continued to be two of the most critical of the challenges. Collection and commercial losses stemming from energy utilised that has not been invoiced due to theft, tampering with metering technology, or bill evasion are terms used to describe non-technical losses (NTLs). Since NTLs are usually detected through power distribution, this is the driving force behind the study.

Like in many other countries, Nigeria provides power through a value chain that includes generation, transmission, and distribution. Several establishments staff these services. These establishments are collectively working hard right now to manage a massive debt load of more than $2 billion. It has been stated unequivocally that Nigeria has spent over $38.45 billion since 1999. Sadly, most of emerging countries are following this path which if not arrested will lead to failure.

NTLs have resulted in significant revenue losses for power providers in various countries. Different kinds of electricity customers bear a disproportionate share of these losses. Consumers bear the brunt of these shortfalls as a result of utility tariff rate increases. However, since power has remained uneven and

SLUJST
Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

consumers are still defaulting, such an increase does not provide the expected outcome. The prevalence of this issue is particularly higher in emerging countries like Nigeria with unstable economies. As a result, low-income and disadvantaged people can no longer afford to purchase power at a discount, making it difficult for them to pay their bills. These and several more factors are the main causes of NTLs (Depuru et al., 2011; León et al., 2011). It has become imperative for academics and power companies to identify the primary causes and locations of these losses to mitigate them.

According to an assessment by Glauner, machine learning (ML) approaches—which involve supervised and unsupervised data analysis techniques—have been used in a variety of ways to address these problems. The writers examined recent research in the area of AI (Artificial Intelligence) (Ghori et al., 2021; Glauner et al., 2017). Several of the approaches utilised techniques such as Deep reinforcement learning (Lee et al., 2022), fuzzy clustering (Messinis & Hatziargyriou, 2018), neural networks (Fei et al., 2022), Feature engineering (Yadav & Kumar, 2021), by identifying notable dips in power use, regression analysis was utilised to examine the link between time and monthly consumption (Monedero et al., 2010) among others. Other studies explored the use of ensemble methods (Avila et al., 2018), and some studies employed the SVM (Support vector machines) algorithm only, or in combination with other methods (Haq et al., 2021) (Nagi et al., 2011). In a different study, the researchers divided the locations of their clients into grids of different sizes after analysing the neighborhood's specific qualities. To determine the quantity of NTLs and the number of clients, each of the created grids was carefully examined (Glauner et al., 2016). A related work generated the local estimate of NTL using a generalised additive model and using a Markov chain to determine future variations in the probability (Faria et al., 2016). Massaferro et. al also conducted a similar study engaging customers' local features to detect NTLs with a random search procedure (Massaferro et al., 2018).

In order to manage power-related challenges, machine learning has been impacted by customer segmentation (Yuan et al., 2018) such as electricity tariff (Nafkha et al., 2018), micro/mini-grids (Williams et al., 2018) and demand side management (Cominola et al., 2018). These represent ideas intended to improve the operations and activities of utilities providing customers with consistent power. By using customer segmentation to target customer groups with comparable payment and usage patterns, utilities may optimise the benefits of these activities. As a result, it has been established that clustering is highly effective in customer segmentation (Kansal et al., 2019). Due to its rapid convergence and ease of use, K-means, an unsupervised learning algorithm, has shown to be highly effective in customer segmentation. The K-means algorithm was utilised to partition the electricity load profiles of customers to get insight into their consumption patterns (Viola, 2016). Electricity customers were characterized and clustered through an electricity demand signature (Camero et al., 2018). Customers were divided into five categories, and K-means was utilised to determine their purchasing habits. The outcomes were then compared with those of alternative clustering techniques (Kansal et al., 2019). The density information of the clusters, however, cannot be measured by the K-means method. As a result, considering the characteristics of the data set, we sought for a better clustering technique, such as DBSCAN, to manage the density and find outliers (Wang & Govindarasu, 2019). Moreover, Forero et al. (2012) introduced clustering techniques based on the Gaussian mixture model (GMM) that are sensitive to data and inconsistent with outliers because of their functional dependence on the Euclidean distance measure (Forero et al., 2012). To further demonstrate the existence of outliers in the power utility dataset, the Expectation Maximisation (EM) technique was chosen.

Through an examination of consumers' tariff classes and payment patterns, this research investigates the potential of clustering approaches to evaluate the intensity of NTLs in Nigeria's electricity industry. With the use of the K-means, DBSCAN, and Expectation Maximisation algorithms, the data will be analysed to identify outliers and segment customers. The evaluation or identification of NTL is an instance of anomaly or fraud detection, which is in line with the research objective.

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

The data was acquired from a Nigerian distribution company for inclusion in the study. The introduction, methods, findings, and conclusion make up the order in which the document is structured.

2. BACKGROUND REVIEW

2.1 The Nigerian Tariff structures

Tariff rates reflect the expenses of the entire power value chain, from the generating plant, transmission, distribution, metering, billing, and ultimately to the customer. Power utilities will have more operational stability if income exceeds costs. There were five (5) tariff classes in the Nigerian power sector according to electricity usage; Residential (R), Commercial (C), Industrial (D), Special (A), and Street lights (S). These customer classes were different because of the operational and infrastructure expenses related to the supply and distribution of electricity to the various customer groups. Customers in the industrial category which are mostly manufacturing facilities, had the highest tariff rates among all the distribution companies, while customers in the Residential class (especially the R1 subclass) who are households utilising home utilities and had energy demands under five kilovolt-ampere (kVA), had the lowest tariff rates. However, in 2020, the electricity regulatory released a new tariff plan known as "Service Reflective Tariffs (SRT)" which forms the benchmark for determining future tariff rates for electricity consumers in Nigeria. Each of the classes consists of subclasses which are derived according to the customer's consumption and hours of electricity supplied as per demand (maximum or non-maximum demand). The concentration of NTL by tariff classes is more appropriate in segmenting customers in the Nigerian power industry. The divisions of customers are portrayed in Figure 1. The residential class constitutes customers where NTLs have the most influence. This may be explained by the facts surrounding NTLs that are perpetrated in residents.

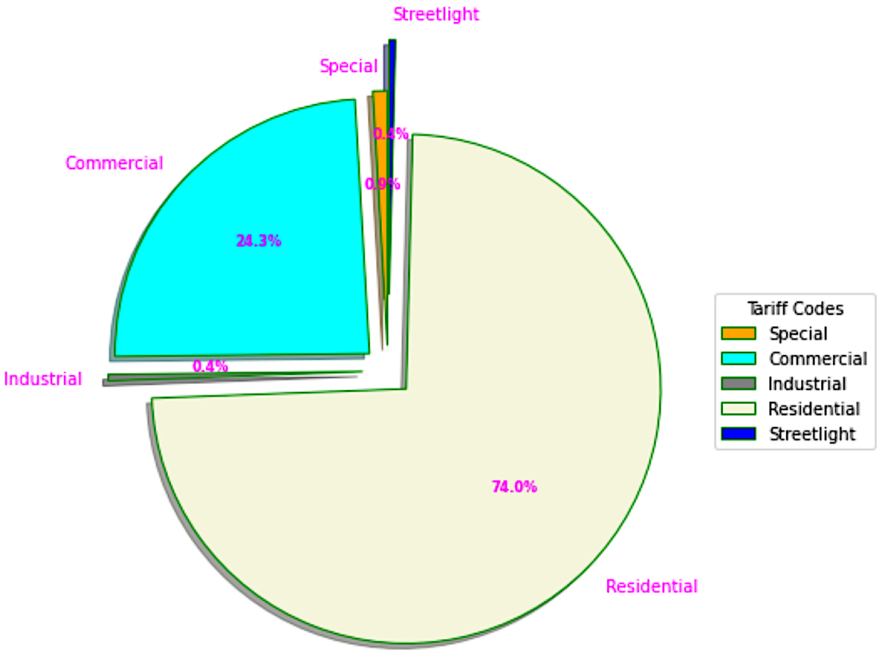

Figure 1. Consumers Tariff Class
(Source: Data from Distribution Company in Nigeria)

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

2.2 Clustering

Objects that have similar qualities or characteristics are grouped to form clusters. Clustering is the process utilised to form clusters. Learning the characteristics of each object is an extremely tough effort; alternatively, it would be simple to group similar things together and have a shared set of attributes that the group adheres to. It is an unsupervised machine-learning technique that captures and separates structured and unstructured data into meaningful groups called clusters based on their similarity (Han & Kamber, 2006). Subsequently, clusters with similar features or properties are formed. The primary goal of clustering is capturing the inherent significance in the structure of the grouped data. Numerous techniques, including K-means, mean shift clustering, DBSCAN, expectation maximisation, and agglomerative hierarchical clustering, are employed in the process of clustering.

The clustering algorithm K-means is based on partitions. By initialising the coordinates of K centroids, it splits the data. By assigning data points to the closest centroid and utilising the Euclidean distance metric to find the centroids, a cluster is created. Until the position of the centroid does not change, the process is performed again. Finding the ideal cluster size k, is essential for K-means clustering. The techniques employed to determine the ideal quantity of k consist of direct techniques like elbow, average silhouette, and statistical testing techniques (Gap Statistic). Finding clusters that will minimise the overall Within-cluster Sum of Squares (WSS) is the primary objective of partitioning techniques like k-means clustering. It is usually desirable to have a lower overall WSS value. The compactness of the clustering is measured by the WSS. The inertia property is utilised to determine the (WSS) distances of the samples that are closest to the cluster centroid once k-means is initialised for each k value. The WSS distance tends to zero as the k numbers rise. All samples will form their clusters, and the WSS distances will equal zero, assuming that k is set to the maximum value of n. If the plot resembles an arm, the ideal value of k is found.

2.2.2 DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

In machine learning issues, the density-based clustering technique known as DBSCAN is used to separate high-density clusters from low-density clusters. It has shown to be incredibly effective in identifying regions of the data with high observation density as opposed to regions with low observation density. Data may be arranged into clusters with different forms using DBSCAN (Jang & Jiang, 2018). The main objective is to locate neighbours on an n-dimensional sphere with a radius ϵ based on their densities. Different classifications of points specified by DBSCAN are the Core point (which forms a cluster with points reachable from it), Border point (lies in a cluster and is reachable by other points in the cluster), and Outlier (does not lie in any cluster, is not reachable, and is not related to other points by density).

2.2.3 Expectation Maximisation

When there are unobserved, hidden, or incomplete variables in a model, the Expectation-Maximization (EM) process determines the greatest probability estimate for those variables. It approximates the greatest likelihood function iteratively. It is regarded as a k-means algorithm extension. The Expectation/Estimation Step (E Step) and the Maximisation Step (M Step) are the two fundamental phases of the EM algorithm.

3. MATERIALS AND METHODS

3.1 Research Framework

This section explains the steps involved in carrying out research. The research framework is depicted in Figure 2. It presents the processes and methods employed in the study.

Information was taken from a power distribution utility in Nigeria. Preprocessing was done on the CSV data, and pertinent characteristics were chosen for analysis. After the data was scaled using the standard scale (z-score), it was centered around 0 with a standard deviation of 1.

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

$$\text{z-score}(x) = \frac{x - mean(X)}{stdev(X)} \quad (1)$$

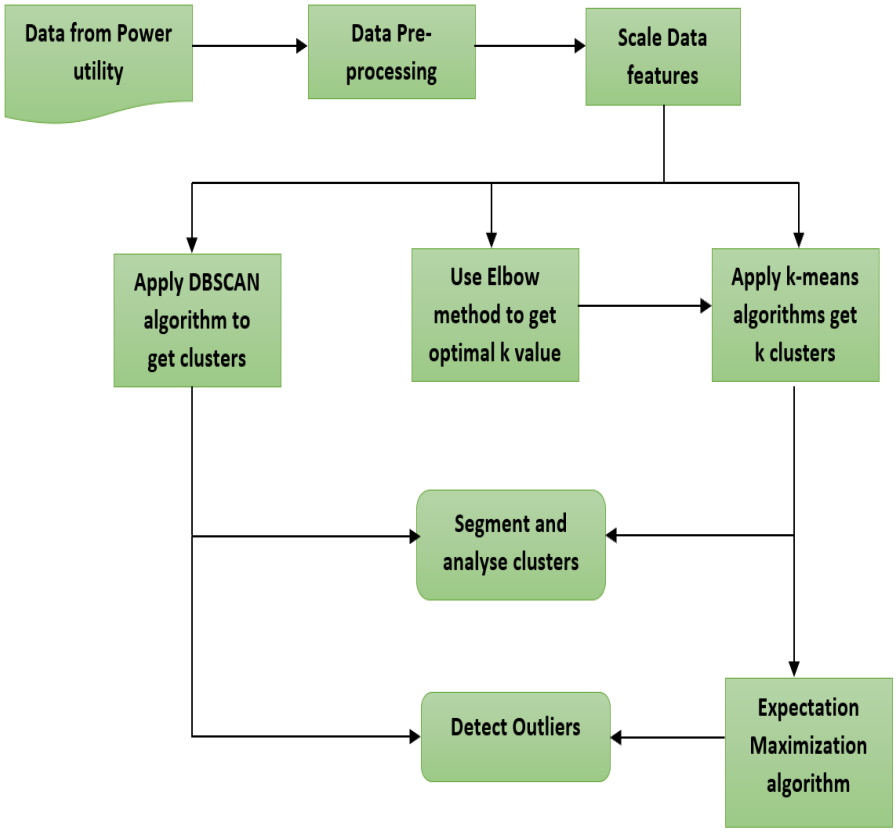

Figure 2. Research Framework

In a given feature set xi ϵ X, x represents the entry, mean (X) is the mean of the feature set, and stdev (X) denotes the standard deviation of X. The elbow method is then applied to calculate the value of k for the dataset. The result obtained from the elbow method will influence the efficiency of the k-means clustering. Furthermore, the DBSCAN and Expectation Maximization clustering algorithms will be used to segment power consumers in a bid to show the effects of NTLs and to demonstrate the efficiency of the algorithm in showing the density of the
clusters and outliers.

4. EXPERIMENTAL RESULTS

In order to segment customers, this article explores the data as mentioned in the preceding part to look at the payment habits of power users. Using k-means, DBSCAN and Expectation Maximization clustering algorithms, we categorise the consumers based on their bill payment and the resulting clusters were analysed.

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

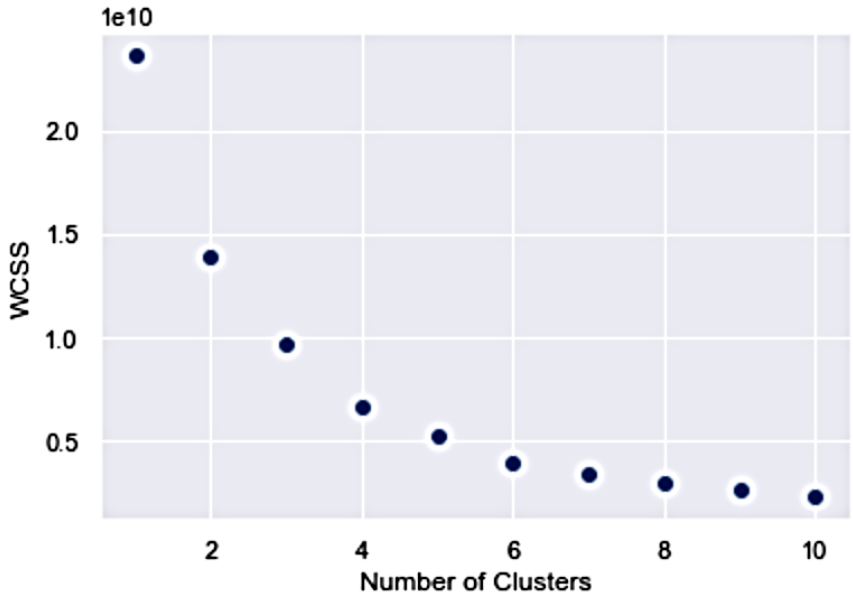

Figure 3. *Elbow method for* values of k

It is evident from Figure 3 that the number of clusters equal to 6 is where the curve's point of inflection is. Therefore, k = 6 is the ideal number from our dataset. The score inevitably drops as the number of clusters rises. The WCSS rate is expected to drop as soon as the ideal number of clusters (k) is attained.

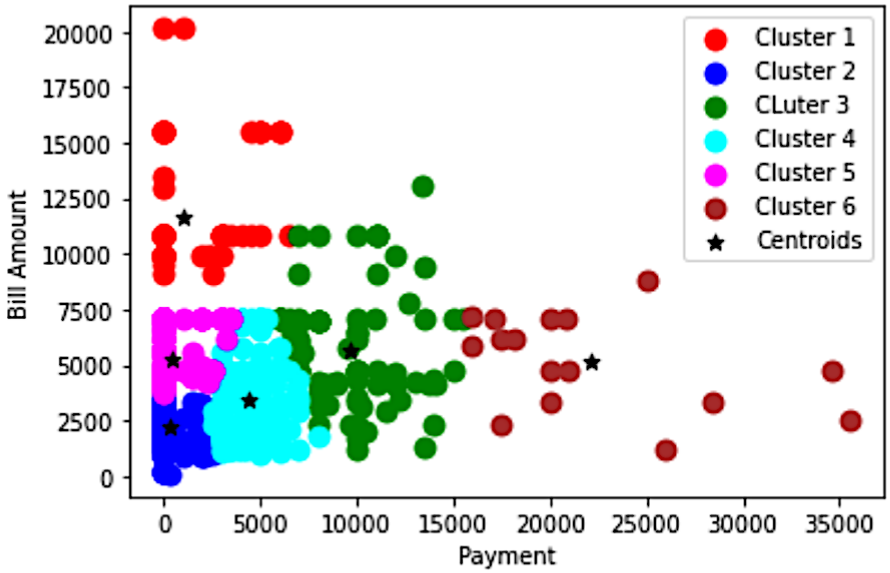

Figure 4. Payment Clusters

Figure 4 depicts 6 payment clusters with their centroids obtained from running the K-means algorithm. The cluster analyses are as given:

Cluster 1 – Consists of customers who have paid part of their bills to the utility provider. However, some of the customers have not even paid at all. Residential and Commercial tariff schemes are in effect for these consumers.

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Cluster 2 – This group of customers has the lowest bill amount. due to the large concentration of no payments. However, the least payment is received from them and has the greatest concentration of zero payments (a great collection loss -NTL) with the centroid at zero. It shows that the consumers have not been paying their monthly bills or made very few payments to the distribution facility. These consumers belong to the Residential category customers and have a low demand (consumption is low).

Cluster 3 **–** This cluster consists of customers who have an accumulated debt profile due to not paying for the power they have used. It can be observed that the consumers in this category are largely on the Special tariff plan and the others are on the Commercial and industrial plans. It clearly shows that they have a high consumption profile (maximum demand MD) and evidently their inability to reimburse for the power utilised has resulted in skyrocketing bills.

Cluster 4 – consists of those very few consumers who have made some payments since zero payment is not recorded against the group. A few are in the Residential category and the remaining belong to the commercial category.

Cluster 5 **–** appears to be the cluster with substantial zero payments as well. Though some of them have low consumption demands, their obligations have not been met. All the consumers in this category belong to the Residential Tariff plan.

Cluster 6 **–** shows a large concentration of customers in the commercial, industrial, and special categories. A portion of the customers in this group have either made partial or full bill payments. The NTL menace is not as extreme as seen in the other clusters.

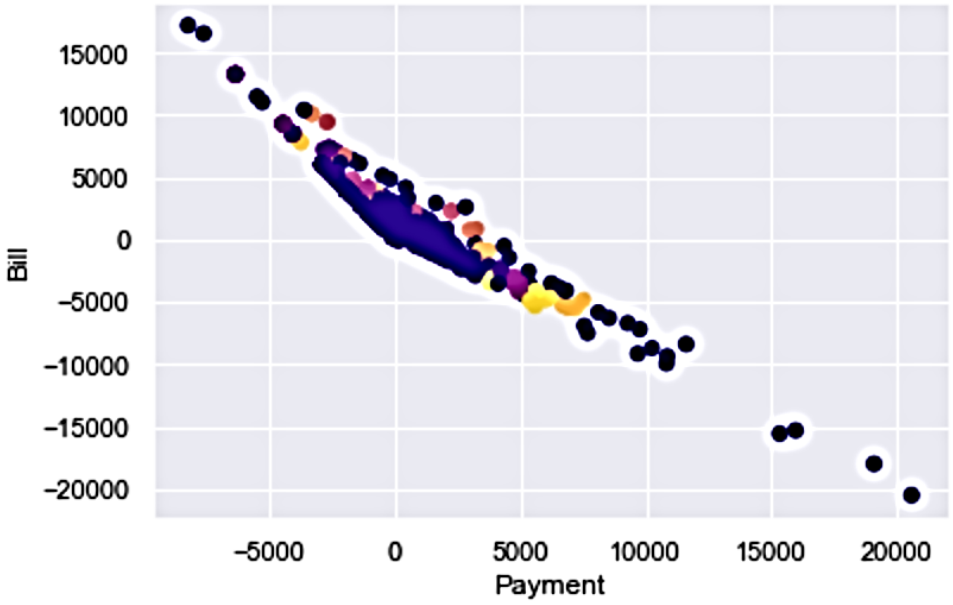

Figure 5. DBSCAN clusters

Figure 5 shows that there are outliers in the data because the points on the outside margins of the dataset are not identified. This suggests a high concentration of losses due to serious debt scenarios. The high-density spots are grouped, suggesting that severe instances of power non-payment are present in all customer groups.

The K-means method is extended by the EM algorithm, a kind of Gaussian mixture model. It applies the expectation-maximization process. Up until it converges, it selects estimates for the shapes and locations of the clusters. Figures 6 and 7 demonstrate that each cluster is connected to a smooth Gaussian model

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

rather than a hard-edged sphere. Multiple random initializations are used by the method to construct the clusters, just like in the k-means expectation–maximization algorithm. The cluster components indicated by numbers 5 and 6 were allocated as the number of clusters created, indicated by n. In all scenarios, the effects displayed the outliers.

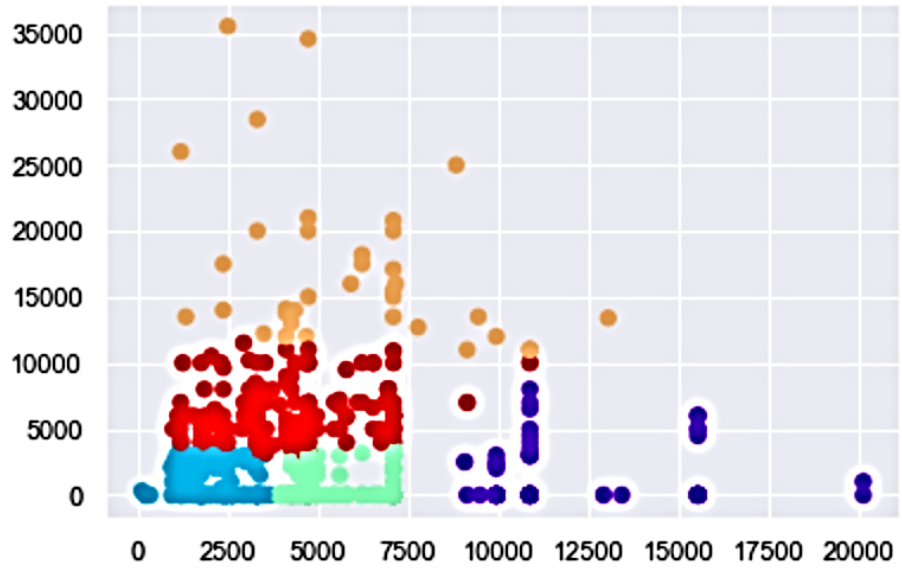

Figure 6. EM cluster assignment, n=5

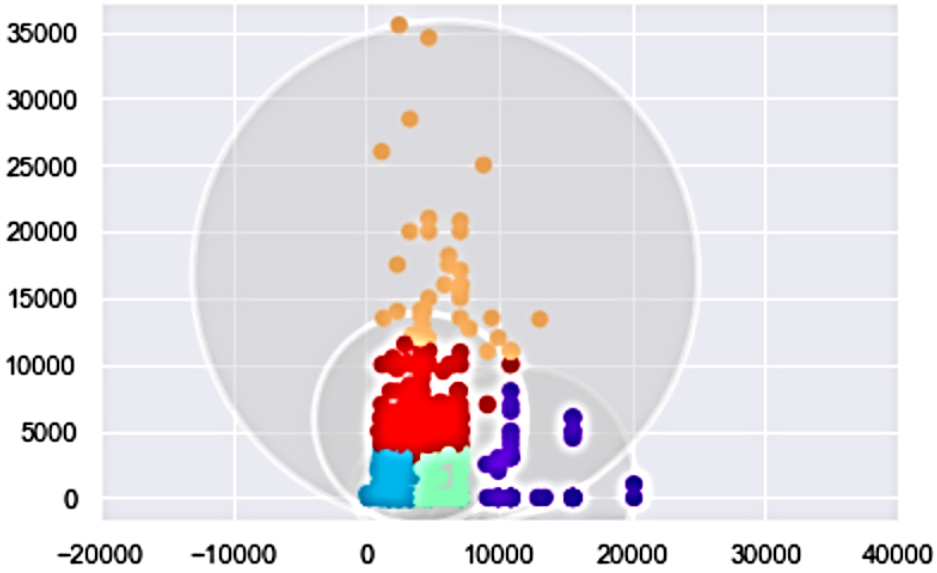

Figure 7. EM cluster assignment, n=5

In Figure 8, we can see that the uncertainty of the cluster assignment is reflected by the dots at the borders of the clusters. Following the transformation of the data, a cluster assignment was made.

Sule Lamido University Journal of Science & Technology Vol. 9 No. 1&2 [May, 2024], pp. 197-207
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

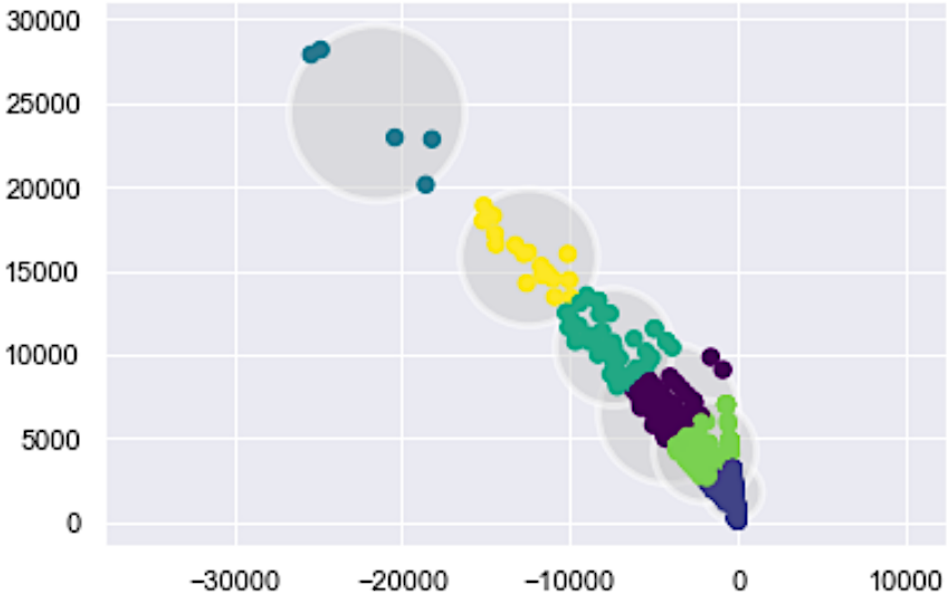

Figure 8. Stretched Five-Component GMM

In Figure 8, a six-component GMM for the initial data in Figure 6 was used to fit the stretched dataset. This allowed for a full covariance for the model to fit the oblong and stretched-out clusters. This clearly shows the outliers, making the EM algorithm a good tool for detecting unusual observations in the dataset.

5. Conclusion

The robust capabilities of clustering approaches and their behaviours concerning the power distribution dataset as they are used in customer segmentation are demonstrated in this research. The data, illustrates the degree of bill payment default among Nigerian power consumers. K-Means provided a clear visual representation of the consumer segments by displaying six clusters along with their centroids. However, we were unable to create cluster morphologies that would have allowed for accurate outlier detection. Conversely, DBSCAN showed dense clusters that represented noise and outliers (extreme NTL instances) in the data. We saw a thick cluster that demonstrated the severity of Nigerian power users' nonpayment of electricity. Clusters were clearer with EM. The clusters that emerged provided information on how many NTLs are present in Nigeria's electricity industry and where they are most prevalent. The results of the study have demonstrated the applicability of the clustering algorithms DBSCAN and Expectation Maximisation for determining the extent of losses sustained in the power industry. This research is intended to assist electricity companies in handling defaulters with caution and determination. In order to provide microgrids to deserving areas, customer segmentation of power customers might be investigated. The utilities must also find the consumers who will help them transition to a more financially stable regime.

REFERENCES

Avila, N. F., Figueroa, G., & Chu, C. C. (2018). NTL Detection in electric distribution systems using the maximal overlap discrete wavelet-packet transform and random undersampling boosting. *IEEE Transactions on Power Systems*, *33*(6), 7171–7180. https://doi.org/10.1109/TPWRS.2018.2853162

Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Camero, A., Luque, G., Bravo, Y., & Alba, E. (2018). Customer segmentation based on the electricity demand signature: The andalusian case. *Energies*, *11*(7), 1–14. https://doi.org/10.3390/en11071788

Cominola, A., Spang, E. S., Giuliani, M., Castelletti, A., Lund, J. R., & Loge, F. J. (2018). Segmentation analysis of residential water-electricity demand for customized demand-side management programs. *Journal of Cleaner Production*, *172*, 1607–1619. https://doi.org/10.1016/j.jclepro.2017.10.203

Depuru, S. S. S. R., Wang, L., & Devabhaktuni, V. (2011). Electricity theft: Overview, issues, prevention and a smart meter based approach to control theft. *Energy Policy*, *39*(2), 1007–1015. https://doi.org/10.1016/j.enpol.2010.11.037

Faria, L. T., Melo, J. D., & Padilha-Feltrin, A. (2016). Spatial-Temporal Estimation for Nontechnical Losses. *IEEE Transactions on Power Delivery*, *31*(1), 362–369. https://doi.org/10.1109/TPWRD.2015.2469135

Fei, K., Li, Q., & Zhu, C. (2022). Non-technical losses detection using missing values' pattern and neural architecture search. *International Journal of Electrical Power & Energy Systems*, *134*, 107410. https://doi.org/https://doi.org/10.1016/j.ijepes.2021.107410

Forero, P. A., Kekatos, V., & Giannakis, G. B. (2012). Robust Clustering Using Outlier-sparsity Regularization. *IEEE Transactions on Signal Processing*, *60*(234914), 4163–4177. https://doi.org/10.1109/TSP.2012.2196696

Ghori, K. M. U., Awais, M., Khattak, A. S., Imran, M., Abbasi, R. A., & Szathmary, L. (2021). A review on latest trends in non-technical loss detection. *CEUR Workshop Proceedings*, *2874*, 131–139.

Glauner, P., Meira, J. A., Dolberg, L., State, R., Bettinger, F., & Rangoni, Y. (2016). Neighborhood features help detecting non-technical losses in big data sets. *Proceedings - 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, BDCAT 2016*, 253–261. https://doi.org/10.1145/3006299.3006310

Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques* (J. Gray, Ed.; Second edi). Morgan Kaufmann Publishers.

Haq, E. U., Huang, J., Xu, H., Li, K., & Ahmad, F. (2021). A hybrid approach based on deep learning and support vector machine for the detection of electricity theft in power grids. *Energy Reports*, *7*, 349–356. https://doi.org/10.1016/j.egyr.2021.08.038

Jang, J., & Jiang, H. (2018). *DBSCAN++: Towards fast and scalable density clustering*. http://arxiv.org/abs/1810.13105

Kansal, T., Bahuguna, S., Singh, V., & Choudhury, T. (2019). Customer Segmentation using K-means Clustering. *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 135–139. https://doi.org/10.1109/ctems.2018.8769171

Lee, J., Sun, Y. G., Sim, I., Kim, S. H., Kim, D. I., & Kim, J. Y. (2022). Non-Technical Loss Detection Using Deep Reinforcement Learning for Feature Cost Efficiency and Imbalanced Dataset. *IEEE Access*, *10*, 27084–27095. https://doi.org/10.1109/ACCESS.2022.3156948

León, C., Biscarri, F., Monedero, I., Guerrero, J. I., Biscarri, J., & Millán, R. (2011). Integrated expert system applied to the analysis of non-technical losses in power utilities. *Expert Systems with Applications*, *38*(8), 10274–10285. https://doi.org/10.1016/j.eswa.2011.02.062

Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Massaferro, P., Marichal, H., Martino, M. Di, Santomauro, F., Kosut, J. P., & Fernandez, A. (2018). Improving electricity non technical losses detection including neighborhood information. *IEEE Power and Energy Society General Meeting*, *2018-Augus*(i). https://doi.org/10.1109/PESGM.2018.8586146

Messinis, G. M., & Hatziargyriou, N. D. (2018). Unsupervised classification for non-technical loss detection. *20th Power Systems Computation Conference, PSCC 2018*, 1–7. https://doi.org/10.23919/PSCC.2018.8442797

Monedero, I., Biscarri, F., León, C., Guerrero, J. I., Biscarri, J., & Millán, R. (2010). Using regression analysis to identify patterns of non-technical losses on power utilities. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6276 LNAI*(PART 1), 410–419. https://doi.org/10.1007/978-3-642-15387-7_45

T. (2018). Do customers choose proper tariff? empirical analysis based on polish data using unsupervised techniques. *Energies*, *11*(3). https://doi.org/10.3390/en11030514

Nagi, J., Yap, K. S., Tiong, S. K., Ahmed, S. K., & Nagi, F. (2011). Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system. *IEEE Transactions on Power Delivery*, *26*(2), 1284–1285. https://doi.org/10.1109/TPWRD.2010.2055670

Viola, L. G. (2016). *Segmentation of household load-profiles with K- means clustering algorithm Segmentation of household load-profiles with K-means clustering algorithm*. *July*, 0–9.

Wang, P., & Govindarasu, M. (2019). Anomaly Detection for Power System Generation Control based on Hierarchical DBSCAN. *2018 North American Power Symposium, NAPS 2018*, *September*. https://doi.org/10.1109/NAPS.2018.8600616

Williams, N. J., Jaramillo, P., Campbell, K., Musanga, B., & Lyons-Galante, I. (2018). Electricity Consumption and Load Profile Segmentation Analysis for Rural Micro Grid Customers in Tanzania. *2018 IEEE PES/IAS PowerAfrica, PowerAfrica 2018*, 360–365. https://doi.org/10.1109/PowerAfrica.2018.8521099

Yadav, R., & Kumar, Y. (2021). Detection of non-technical losses in electric distribution network by applying machine learning and feature engineering. *Journal Europeen Des Systemes Automatises*, *54*(3), 487–493. https://doi.org/10.18280/JESA.540312

Yuan, Y., Dehghanpour, K., Bu, F., & Wang, Z. (2018). *A Data-Driven Approach for Estimating Customer Contribution for System Peak Shaving 18*(1) 1–8.