Formosa Journal of Computer and Information Science (FJCIS)
Vol. 5, No. 1 2026: 63-82

Analysis of PT PLN (Persero)'s New Installation Waiting List Using the K-Means Clustering Algorithm

Ernawati[1*], Dewi Agushinta R[2]
Universitas Gunadarma
Corresponding Author: Ernawati taslimbatubara98@gmail.com

ARTICLEINFO

Keywords: Clustering, Data Mining, K-Means, New Installation, Waiting List

Received : 16, November
Revised : 18, January
Accepted: 20, March

©2026 Ernawati, Dewi Agushinta R: This is an open-access article distributed under the termsof the Creative Commons Atribusi 4.0 Internasional.

ABSTRACT

This study examines the application of the K-means clustering algorithm to analyze new installation waiting list data obtained from the last three months of 2024. Only entries categorized under new installation requests were selected as the primary dataset. The analysis began by determining the optimal number of clusters: a high volume of new installation waiting lists (C1), a medium volume (C2), and a low volume (C3). Data mining processes were carried out using the RapidMiner tool, producing the following results: 6 UIDs/UIWs were classified into the high cluster (C1), 7 into the medium cluster (C2), and 9 into the low cluster (C3). The clustering performance was subsequently validated using the Davies–Bouldin Index, yielding a final score of 0.486, consistent with the RapidMiner output.

DOI prefik: https://doi.org/10.55927/fjcis.v5i1.16429
ISSN-E: 2830-3040
https://journal.formosapublisher.org/index.php/fjcis

INTRODUCTION

Electricity is an essential component that supports various social, economic, and industrial activities in modern life. The community's dependence on electrical energy continues to increase in line with national economic growth and the expansion of the industrial and public service sectors. In this context, PT PLN (Persero), as a national electricity service provider, is required to provide fast, accurate, and quality services to maintain customer trust and support national development (Darmawan et al., 2022). One important aspect of service is the speed of the new connection process, which is regulated by ESDM Regulation No. 27 of 2017 concerning the New Connection Service Time Limit (HPL) based on service categories. Non-compliance with service quality standards not only reduces service quality but also has a financial impact on the company due to the obligation to provide compensation to customers. With a wide service coverage, encompassing 22 distribution and regional units (UID/UIW) throughout Indonesia, the challenge of managing new connection waiting list data has become increasingly complex and requires a systematic and data-driven analytical approach (Tyas & Purnamasari, 2023).

In addressing these issues, the use of cluster analysis methods has become increasingly relevant as a data mining technique capable of grouping objects based on similar characteristics. K-Means Clustering, a popular non-hierarchical algorithm, is considered effective in partitioning large amounts of data, including in the context of grouping PLN work areas based on the level of new connection waiting lists. Previous studies have demonstrated the effectiveness of this method in grouping electricity distribution and disturbance data; however, studies that specifically analyze the waiting list for new connections at the PLN regional unit level are still limited (Kapita et al., 2022). Therefore, this study was conducted to fill this gap by applying the K-Means algorithm to group regional units based on the number of new connection waiting lists, thereby allowing for more accurate identification of service load patterns. The results of this analysis are expected not only to contribute to the development of science, particularly in the application of data mining in electricity service management, but also to serve as a basis for strategic considerations for PLN in its efforts to improve service quality and reduce the risk of compensation due to delays in the new connection process (Mulyadi et al., 2024).

Customer service in the context of electricity is a fundamental element that determines the quality of interaction between companies and service users. PT PLN (Persero) continues to update its service system, including through the use of the PLN Mobile application for a more effective and efficient new installation application process. Literacy regarding new installation services shows that the process involves administrative stages, installation feasibility checks, and physical connections in the field. A deep understanding of this process is important because the effectiveness of service directly affects customer perception, organizational image, and the financial risks that arise if service standards are not met. Theoretical studies on customer service emphasize that the success of a service organization is largely determined by its ability to meet customer needs quickly, accurately, and precisely (Akromah & Kusumasari, 2023).

Customer service in the electricity sector operates within a complex service chain that integrates administrative, technical, and operational dimensions. Each stage is interconnected and requires careful coordination to ensure that the entire process runs efficiently. A delay in one phase can affect subsequent stages and disrupt overall service performance. This condition highlights the importance of synchronization between internal units. Effective coordination ensures that service delivery remains consistent and aligned with established standards (Hikmah & Rusdianto, 2024).

Service effectiveness is closely related to the organization's ability to manage time-sensitive processes. In the context of new electricity installations, timeliness becomes a critical indicator of service performance. Customers generally expect immediate access to electricity for household or business activities. Even minor delays can lead to dissatisfaction because electricity is a fundamental necessity. Therefore, maintaining punctual service delivery becomes a strategic priority for service providers (Septianingsih, 2022).

The use of digital platforms such as PLN Mobile reflects a shift toward more modern and accessible service systems. Through digitalization, customers can access services more easily and monitor their application status in real time. This transparency reduces uncertainty and increases customer confidence in the service process. From an organizational perspective, digital systems also help streamline workflows (Pasaribu et al., 2021). They provide structured data that can be used for evaluating performance and improving service quality. Regional differences present additional challenges in the implementation of electricity services. Variations in infrastructure, geographical conditions, and resource availability influence the speed of service delivery. Areas with difficult terrain or limited infrastructure may experience longer processing times. These disparities require flexible strategies that can adapt to local conditions. At the same time, the organization must maintain uniform service standards across all regions (Karmanita et al., 2024).

Human resources play a crucial role in determining the success of service delivery. The competence and responsiveness of staff directly affect how customers perceive the quality of service. Employees are responsible for executing procedures accurately and handling customer inquiries effectively. Their ability to solve problems in real situations is also essential (Fernaldy et al., 2025). Strong human resource performance supports the overall effectiveness of the service system. Service performance is also influenced by fluctuations in demand for new electricity installations. An increase in applications can place pressure on existing systems and operational capacity. Without proper planning, this situation may lead to service delays and longer waiting times. Managing demand requires careful allocation of resources and efficient scheduling. This ensures that service quality remains stable even during peak periods (Setiawan & Perdhana, 2019).

A comprehensive understanding of the service process enables organizations to identify areas that need improvement. By examining each stage in detail, inefficiencies can be detected more easily. This analysis helps determine which aspects require optimization, whether administrative or technical. Continuous evaluation supports the development of more effective service strategies. As a result, the organization can provide services that better meet customer needs and expectations (Rohanifar et al., 2026).

In the field of data analysis, data mining provides a robust methodological framework for extracting patterns and meaningful information from large data sets. The Knowledge Discovery in Databases (KDD) process provides a systematic flow from data selection, cleaning, and transformation to interpretation of analysis results (Akromah & Kusumasari, 2023). One data mining method widely used in recent research is clustering, which aims to group objects based on similarity characteristics. In the context of electricity services, this method is relevant for identifying areas with specific patterns such as high waiting lists for new installations that require priority handling. K-Means, as a simple but effective non-hierarchical clustering algorithm, partitions data into several clusters based on the proximity of attribute values to the cluster centroid. Validating clustering results, such as through the Davies Bouldin Index, allows researchers to assess the quality of cluster separation so that the analysis results can be scientifically justified (Nuraeni et al., 2023).

Previous studies have shown that the K-Means algorithm and its variants are effective approaches for various problems in the field of electricity services. Research on electricity distribution, fault data clustering, and analysis of new connection services in various PLN regions proves that this method is capable of providing accurate data pattern mapping and can support strategic decision-making. In addition, research in the academic and health fields that implements K-Means also shows the consistency of this algorithm in producing clusters with clear and practically relevant structures (Apriyani et al., 2023).

LITERATURE REVIEW

This study builds upon previous research in data mining and customer service optimization by applying the K-Means clustering algorithm to analyze waiting list patterns for new electricity installations at PT PLN (Persero), aiming to identify priority groupings, improve service efficiency, and support data-driven decision-making.

METHODOLOGY

This study uses a quantitative approach with data mining methods through the K-Means Clustering algorithm to analyze patterns in the new connection waiting list at PT PLN (Persero) (Dilawati et al., 2024). The research stages began with planning, which included a literature review on new connection services, data mining concepts, and the effectiveness of clustering algorithms in electricity operational analysis, as described in the literature review. The literature study was conducted to strengthen the theoretical basis and identify gaps in previous research so that the research topic could be formulated appropriately. The next stage was the collection and processing of secondary data sourced from PT PLN

(Persero)'s internal database, specifically data on the new connection waiting list for the last quarter of 2024. The data was then cleaned, selected, and transformed using Microsoft Excel software to meet the analytical structure required in the clustering process (Farissa et al., 2021; Nuraeni & Firdaus, 2025).

Once the data was ready for processing, clustering was performed by applying the K-Means algorithm through the RapidMiner application. The K-Means stages refer to standard procedures, namely determining the number of clusters (K), initializing the initial centroid, calculating the distance of objects to the centroid using Euclidean Distance, and grouping data based on characteristic proximity. The process was carried out iteratively until the centroid was stable and did not experience significant changes. This study established three main clusters to identify waiting list levels, namely high (C1), medium (C2), and low (C3). To ensure the quality of the clustering results, this study uses the Davies Bouldin Index (DBI) as an internal validation method. A lower DBI value indicates a more optimal cluster structure, so this approach allows researchers to obtain a methodologically accountable cluster division (Kapita et al., 2022).

The final stage of the research includes interpreting the clustering results to identify work units (UID/UIW) with the highest waiting lists and factors that could potentially affect delays, as referred to in the Service Quality Level (TMP) criteria. This analysis provides a comprehensive overview of the distribution of service loads and forms the basis for developing strategic recommendations for improving the quality of new connection services. The entire research process was carried out from December 2024 to January 2025 at PT PLN Icon Plus, with adequate hardware and software support to ensure the accuracy of the analysis and reproducibility of the results (Hartono & Widiantoro, 2024).

RESEARCH RESULT

Results of K-Means Clustering of the New Connection Waiting List

The application of the K-Means algorithm to the new fitting waiting list data for the last quarter of 2024 produced three clusters with different characteristics, as formulated in the previous research method. The first cluster identified the operational area with the highest waiting list and became the main focus of the study. The results show that the East Java UID is the unit with the largest waiting list accumulation, namely 69,892, followed by the Central Java and DIY UID (57,851), the South Sulawesi UIW (56,870), UIW S2JB (47,516), West Java UID (42,957), and South and Central Kalimantan UIW (41,755). These findings reflect the high concentration of service load in these six regional units, making them relevant as priority areas in the service quality improvement strategy.

The clustering results reveal a clear imbalance in the distribution of service loads across operational regions. This variation indicates that demand for new electricity connections is not evenly distributed, but instead concentrated in specific areas. Regions with higher demand are likely influenced by factors such as population density, urbanization, and economic activity. These external conditions significantly shape the volume of service requests received by each unit. Such findings highlight that service challenges are not solely internal, but are also driven by regional characteristics.

The first cluster represents areas with significantly higher service pressure compared to other groups. This condition suggests that these regions require greater operational capacity to handle the volume of pending requests. Without proper resource allocation, the risk of service delays becomes more pronounced. High accumulation of waiting lists may also reflect limitations in infrastructure readiness or workforce availability. Therefore, these regions demand priority attention in service improvement strategies. The second cluster reflects regions with a moderate level of waiting lists. These areas appear to maintain a relatively balanced relationship between demand and service capacity. However, the stability observed in this cluster does not eliminate the potential for future increases in demand. If not properly managed, these regions may gradually shift toward higher service pressure. Preventive measures are necessary to ensure that service performance remains consistent over time.

The third cluster includes regions with relatively low waiting list values, indicating that service capacity is generally sufficient to meet demand. This condition suggests a more efficient operational environment or lower service pressure. Regions within this cluster may serve as a reference point for best practices in service management. Studying their operational strategies could provide insights into improving performance in higher-load regions. These findings reinforce the importance of comparative analysis across clusters. The variation in the number of members within each cluster provides additional insight into the overall distribution pattern. A larger number of regions in a cluster indicates a broader representation of similar service conditions. This distribution helps identify which category dominates the national service landscape. It also supports the development of differentiated policies based on cluster characteristics. As a result, decision-making can be more targeted and effective.

The application of the K-Means algorithm enables the transformation of complex data into structured and interpretable information. By grouping regions with similar characteristics, the analysis simplifies the identification of underlying patterns. This approach enhances the ability to evaluate service performance across multiple regions simultaneously. It also supports data-driven decision-making in operational planning. Consequently, resource allocation can be optimized based on actual needs. These findings also open opportunities for implementing region-specific service strategies. Each cluster requires a different approach depending on its level of service pressure. High-load regions need acceleration and capacity expansion, while moderate regions require stabilization strategies. Low-load regions can focus on maintaining efficiency and consistency. This segmentation approach strengthens the overall effectiveness of service management within the organization.

Table 1. Results of the New Connection Waiting List Cluster for the Last Quarter of 2024

Group	Number of Members	UID/UIW Name	Number
1	6	UIW S2JB	47516
		UIW South and Central Kalimantan	41755
		UIW South Sulawesi and West Sulawesi	56870
		UID East Java	69892
		UID Central Java and Yogyakarta	57851
		UID West Java	42,957
2	7	UIW North Sumatra	23677
		UID Lampung	19912
		UIW West Kalimantan	20728
		UIW Central Sulawesi	28146
		UIW Papua and West Papua	19166
		East Nusa Tenggara UIW	25869
		UID Bali	27127
3	9	UIW Aceh	7382
		UIW West Sumatra	7202
		UIW Bangka Belitung	2580
		UIW Riau and Riau Islands	17053
		East Kalimantan UIW	11188
		UIW Maluku and North Maluku	11549
		UIW West Nusa Tenggara	14192
		UID Greater Jakarta	16848
		UID Banten	10863

Based on Table I, the clustering results are divided into three groups with distinct quantitative characteristics. Cluster 1 consists of six regions with the highest waiting list values, indicating a concentration of service demand in these areas. Regions such as East Java and Central Java & Yogyakarta show significantly higher numbers compared to others, reflecting strong demand pressure likely driven by economic and population factors.

Cluster 2 includes seven regions with moderate waiting list values. The distribution within this group appears relatively balanced, with no extreme values compared to cluster 1. Regions such as Central Sulawesi and Bali demonstrate stable levels of demand, suggesting that their service capacity is still relatively aligned with incoming requests. This cluster represents a transitional category between high and low service loads.

Cluster 3 contains nine regions and represents the largest group, but with the lowest waiting list values. This indicates that most regions fall into the low-demand category. Areas such as Bangka Belitung and West Sumatra show very small numbers, which may reflect either efficient service systems or lower levels of demand. These regions demonstrate relatively stable service conditions.

Overall, the table highlights an uneven distribution of service loads across regions. Cluster 1 plays a dominant role in contributing to the national waiting list volume and requires immediate attention. Cluster 2 serves as a buffer group with potential for growth in demand. Cluster 3 reflects stable conditions and can be used as a benchmark for efficient service management practices.

Cluster Quality Validation Using the Davies Bouldin Index

The validity of the clusters was verified using the Davies Bouldin Index (DBI), in accordance with the evaluation approach in data mining described in the literature review. The DBI calculation result of 0.486 indicates that the clustering quality is in the good category, because values closer to zero indicate more optimal clusters. This value indicates that the units in the cluster have a high degree of similarity (cohesive), and the boundaries between clusters are quite clear. Thus, the K-Means model used has been proven to produce a stable cluster structure that is methodologically accountable. The validation process using the Davies Bouldin Index provides an important layer of assurance regarding the reliability of the clustering results. This evaluation focuses on measuring how well each cluster is internally compact while remaining distinct from other clusters. A strong validation outcome indicates that the grouping structure is not formed randomly, but reflects meaningful patterns in the data. This strengthens confidence in the analytical approach used. It also ensures that the clustering output can be interpreted with a higher level of credibility.

A low DBI value reflects a balance between intra-cluster cohesion and inter-cluster separation. This balance is essential in clustering analysis because it determines how clearly different groups can be distinguished. When clusters are well-separated, the risk of overlapping characteristics between groups becomes minimal. This allows for more precise interpretation of each cluster's identity. As a result, decision-making based on these clusters becomes more accurate. The quality of clustering directly influences the usefulness of the analysis for practical applications. When clusters are well-formed, they provide a clearer representation of real-world conditions. This makes it easier to identify patterns, trends, and anomalies within the data. In the context of operational management, such clarity supports more effective planning and control. The validation result therefore plays a key role in ensuring that analytical findings are actionable.

Another important aspect of cluster validation is its contribution to methodological robustness. A validated model demonstrates that the analytical process follows sound scientific principles. This strengthens the legitimacy of the study and enhances its acceptance in academic and professional contexts. It also indicates that the model can be replicated or adapted for similar datasets. Such robustness is essential for sustaining long-term analytical applications.

The stability of the cluster structure suggests that the grouping remains consistent under the given data conditions. Stable clusters are less sensitive to minor variations in data, which improves the reliability of the analysis. This characteristic is particularly important when dealing with operational data that may change over time. A stable model can accommodate updates without significantly altering the overall structure. This makes it suitable for continuous monitoring and evaluation. The validation outcome also supports the integration of clustering results into strategic decision-making processes. Reliable clusters can be used as a basis for segmentation, prioritization, and resource allocation. This allows organizations to tailor their strategies according to the specific characteristics of each group. The ability to differentiate between clusters enhances the effectiveness of targeted interventions. It also improves the efficiency of operational planning.

Furthermore, the validation process highlights the importance of combining analytical techniques with evaluation metrics. Clustering alone does not guarantee meaningful results without proper validation. By incorporating evaluation measures such as DBI, the analysis becomes more comprehensive. This approach ensures that the results are not only technically correct but also practically relevant. It reinforces the role of data-driven methods in supporting organizational performance.

Figure 1. Performance Vector Results.

The performance vector results presented in Figure 1 summarize the evaluation metrics used to assess clustering quality. The displayed values indicate that the model produces a relatively low error level and maintains consistency across clusters. This suggests that the algorithm performs efficiently in grouping data with similar characteristics. The presence of clear metric outputs also indicates that the evaluation process is systematically conducted (Ifaldiansyah & Hertati, 2023; Nurwanda et al., 2024).

The figure reflects that the clustering process has been successfully executed without significant deviations. The recorded performance values support the conclusion that the model operates within acceptable analytical standards. Each metric contributes to understanding how well the clusters are formed. The consistency of these values reinforces the reliability of the clustering outcome.

In addition, the performance vector provides a concise overview of the model's analytical strength. It allows researchers to quickly assess whether the clustering meets the expected quality criteria. The structured format of the output also facilitates comparison with other models or future analyses. This makes it a useful tool for ongoing evaluation and improvement.

Overall, the figure confirms that the clustering model achieves a satisfactory level of performance. The evaluation metrics align with the interpretation of a well-structured clustering result. This strengthens the overall conclusion that the applied method is appropriate for the dataset. The performance vector therefore serves as a supporting element in validating the analytical findings.

Analysis of TMP Criteria in Priority Clusters

In addition to cluster mapping, this study also evaluated the achievement of Service Quality Targets (TMP) as operational indicators. The results of the analysis show that the TMP category with the largest proportion is the 5-day completion time without expansion, accounting for 55% of total new installation requests. This category describes relatively simple requests that do not require additional network development. The high achievement in this category indicates that PT PLN (Persero) has implemented basic service procedures effectively, particularly in areas included in the priority cluster.

The evaluation of TMP criteria provides a deeper understanding of operational performance beyond cluster classification. It allows the analysis to move from identifying patterns to assessing how well service standards are implemented in practice. By examining completion time categories, the study captures the efficiency of service execution under different technical conditions. This approach helps distinguish between routine and more complex service scenarios. As a result, operational insights become more nuanced and context-specific. The dominance of the 5-day completion category without expansion reflects the efficiency of handling standard service requests. These requests typically involve minimal technical complexity and rely on existing infrastructure. The ability to process such requests within the specified timeframe indicates that internal workflows are functioning effectively. It also suggests that coordination between administrative and technical units is well-established. This condition supports consistent service delivery in routine cases.

The high proportion of this category also highlights the importance of maintaining operational stability in core service processes. Routine installations form the majority of service demand, making their efficient handling critical for overall performance. Any disruption in this segment could significantly affect customer satisfaction levels. Therefore, sustaining performance in this category is essential for preserving service reliability. It also serves as a baseline for evaluating improvements in other service areas. In priority clusters, the effectiveness of TMP achievement becomes even more significant. These regions face higher service pressure, which increases the risk of delays. The ability to maintain strong performance in standard service categories under such conditions reflects operational resilience. It indicates that despite higher demand, the system can still deliver consistent results. This strengthens confidence in the organization's capacity to manage workload variations.

However, the focus on simpler service categories also implies the need to examine more complex cases. Requests requiring network expansion or additional technical work may present different challenges. These cases often involve longer processing times and greater resource allocation. Understanding their performance is important for achieving a balanced evaluation of service quality. This ensures that improvements are not limited to routine operations but extend to more demanding scenarios.

The distribution of TMP categories can also be used to identify potential inefficiencies in service handling. A concentration in one category may indicate strong performance in that segment, but it may also reveal gaps in others. This perspective encourages a more comprehensive evaluation of operational performance. By analyzing the full spectrum of TMP criteria, organizations can better understand their strengths and limitations. This supports more informed decision-making.

The integration of TMP analysis with cluster mapping enhances the overall analytical framework. It connects spatial distribution patterns with performance indicators, creating a more holistic view of service conditions. This combined approach enables more targeted strategies for service improvement. It also allows for better prioritization of resources based on both demand and performance metrics. Such integration strengthens the effectiveness of data-driven management practices.

Identification of Service Challenges in Clusters

Although the 55% percentage reflects fairly good service performance, 45% of applications could not be completed within the 5-day time limit. This percentage indicates obstacles related to network expansion, geographical conditions, or additional licensing requirements as outlined in the TMP criteria. These results suggest that the service load in priority cluster areas is not only high in terms of the quantity of requests, but also has significant technical complexity. Therefore, analysis of these results is an important basis for formulating strategic steps to improve services.

The proportion of applications exceeding the 5-day completion target reflects the presence of structural and operational constraints within the service system. These constraints are not isolated incidents but are embedded in the broader workflow of installation processes. When a significant portion of requests requires extended handling time, it indicates that existing procedures must accommodate a wider range of technical conditions. This situation highlights the importance of differentiating service strategies based on complexity levels. A uniform approach may not be sufficient to address diverse operational challenges.

The unmet time target also signals the need to examine the readiness of supporting infrastructure. In areas where network expansion is required, the availability of materials, equipment, and technical personnel becomes a critical factor. Limitations in any of these components can delay the execution process. This condition suggests that service performance is closely tied to logistical efficiency. Strengthening supply chain coordination can therefore play a key role in improving completion times (Apriyani et al., 2023; Farissa et al., 2021).

Geographical conditions further contribute to variations in service delivery performance. Regions with difficult terrain or limited accessibility may require additional time for installation activities. These physical constraints are often beyond immediate organizational control, yet they must be anticipated in operational planning. Adapting service models to local conditions becomes essential in such contexts. This ensures that performance expectations remain realistic while still striving for efficiency.

Administrative and regulatory aspects also influence the duration of service completion. Additional licensing requirements or procedural approvals can extend processing time beyond standard targets. These steps are necessary to ensure compliance and safety, but they introduce complexity into the workflow. Efficient coordination with external stakeholders is therefore crucial. Streamlining administrative processes can help reduce unnecessary delays without compromising regulatory standards.

The presence of technically complex requests in priority clusters suggests that high demand is accompanied by higher operational difficulty. This combination creates a dual challenge for service providers. Not only must they handle a large volume of requests, but they must also manage cases that require more intensive resources. This condition calls for a more segmented approach to service management. Differentiating between simple and complex cases allows for more efficient allocation of effort (Hartono & Widiantoro, 2024).

Understanding the distribution of complex cases can support more targeted performance improvements. By identifying where and why delays occur, organizations can develop specific interventions for each type of challenge. This may include increasing technical capacity, improving planning accuracy, or enhancing coordination mechanisms. Such targeted strategies are more effective than general improvements applied uniformly. They allow for a more precise response to operational realities.

The analysis of these conditions provides a foundation for long-term service optimization. It encourages the development of adaptive strategies that respond to both demand volume and technical complexity. This perspective shifts the focus from merely achieving targets to understanding the underlying factors that affect performance. By addressing these factors systematically, organizations can improve both efficiency and reliability. This approach supports sustainable improvements in overall service quality.

Implications of Clustering Results Mapping on Operational Planning

Overall, the clustering results, DBI validation, and TMP analysis confirm that the six priority UIDs/UIWs require special attention in new connection service planning. Mapping these priority areas provides an objective basis for PT PLN (Persero) to allocate resources in a more targeted manner, increase technical capacity, and optimize business processes so that completion times are closer to the ideal standard. These findings also ensure that the recommendations in the conclusion can be formulated based on data rather than mere assumptions.

The integration of clustering outcomes into operational planning introduces a more structured approach to managing service performance across regions. By identifying priority areas through data-driven methods, planning activities can be aligned with actual service demands rather than relying on generalized assumptions. This alignment improves the accuracy of operational decisions and reduces the likelihood of misallocation of resources. It also enables organizations to anticipate challenges more effectively. Such an approach strengthens the overall responsiveness of the service system. The identification of priority units allows for more precise distribution of human and technical resources. Regions with higher service pressure can be supported with additional workforce, equipment, and operational support. This targeted allocation helps balance workload disparities between regions. It also minimizes the risk of service bottlenecks in high-demand areas. Efficient resource distribution ultimately contributes to more consistent service delivery outcomes.

Operational planning also benefits from the ability to prioritize interventions based on cluster characteristics. Each priority area can be assigned specific improvement strategies that address its unique conditions. This ensures that planning is not uniform but adaptive to varying operational contexts. Tailored strategies increase the likelihood of achieving performance targets. They also enhance the effectiveness of implementation efforts across different regions. The incorporation of analytical results into planning processes supports better coordination between organizational units. When planning is guided by clear data patterns, communication between departments becomes more focused and goal-oriented. This reduces ambiguity in decision-making and clarifies operational priorities. It also facilitates the synchronization of administrative and technical activities. As a result, service processes can be executed more efficiently.

Another important implication is the improvement of performance monitoring mechanisms. With clearly defined priority areas, performance indicators can be tracked more systematically. This enables organizations to evaluate whether implemented strategies produce the expected outcomes. Continuous monitoring also allows for timely adjustments when necessary. Such a feedback system strengthens the adaptability of operational planning. The mapping results further support the development of long-term strategic planning. By understanding patterns of service demand and complexity, organizations can design capacity-building initiatives that address future needs. This includes investments in infrastructure, technology, and human resource development. Forward-looking planning reduces the risk of recurring service issues. It also ensures that the organization remains prepared for increasing demand.

Finally, the use of empirical evidence in formulating recommendations enhances the credibility of strategic decisions. Decisions based on validated data are more likely to gain acceptance from stakeholders. This approach reduces reliance on subjective judgment and strengthens accountability. It also ensures that planning outcomes are grounded in measurable realities. Such evidence-based decision-making contributes to sustainable improvements in service performance.

DISCUSSION

Relevance of Cluster Findings to Service Issues

The results of this study directly address the issue raised in the introduction, namely, the identification of areas with high waiting lists that have an impact on the delivery of public services. Clustering shows that high waiting lists, particularly in East Java UID and other areas in cluster 1, have the potential to influence public perception of the effectiveness of PLN's service delivery. This is in line with literature that emphasizes that service queues are a key indicator in evaluating the quality of public services. The identification of areas with high waiting lists provides a clearer picture of how service performance varies across regions. These variations reflect differences in operational capacity, demand intensity, and local conditions. When certain areas consistently experience higher waiting lists, it indicates a gap between service demand and available resources. This gap becomes a critical point of concern in public service delivery. Addressing it requires a focused and context-sensitive approach.

High waiting lists also have broader implications beyond operational performance. They influence how the public evaluates the reliability and responsiveness of the service provider. When delays become frequent, users may develop perceptions of inefficiency or lack of preparedness. These perceptions can spread quickly, especially in regions with high service demand. Maintaining public trust therefore depends on the ability to manage waiting times effectively.

The concentration of waiting lists in specific regions highlights the need for differentiated service strategies. Not all areas face the same level of demand or complexity, and uniform solutions may not produce optimal results. Regions with higher waiting lists require more intensive intervention and resource allocation. This targeted approach helps reduce disparities in service performance. It also ensures that improvements are directed where they are most needed. Service queues can also be viewed as a reflection of system capacity under pressure. When demand exceeds processing capability, waiting lists naturally increase. This condition provides valuable insight into the limits of current operational systems. Understanding these limits is essential for designing improvements that are both realistic and sustainable. It allows organizations to strengthen their systems in a structured manner (Anggraeny et al., 2025; Fernaldy et al., 2025).

The findings emphasize the importance of integrating analytical results into service evaluation frameworks. By linking waiting list data with performance assessment, organizations can better understand the root causes of service challenges. This approach supports more accurate diagnosis and more effective solutions. It also reinforces the role of data in guiding improvements. Ultimately, this contributes to enhancing the overall quality of public service delivery.

Quality of Clustering and Methodological Accuracy

The DBI value of 0.486 obtained shows that the K-Means method has a good level of accuracy in mapping PLN operational data. This reinforces the methodological findings in the literature review, which state that K-Means is effective when applied to large-scale and heterogeneous data. The high cluster validity ensures that the six priority UIDs/UIWs identified truly have similar characteristics, namely, high waiting lists that are not found in other clusters (Karmanita et al., 2024). The obtained validation outcome indicates that the clustering structure is sufficiently robust to represent real conditions within the dataset. This level of accuracy suggests that the grouping is not influenced by random variation, but rather reflects consistent patterns embedded in the data. Such consistency is essential when the results are used to support operational decisions. It ensures that identified patterns can be relied upon for further analysis. This strengthens the overall analytical framework of the study.

The ability of the method to handle heterogeneous data is particularly important in this context. Operational data from multiple regions typically contain variations in scale, distribution, and characteristics. An effective clustering model must be able to accommodate these differences without compromising accuracy. The validation result confirms that the model performs well under such conditions. This indicates that the analytical approach is suitable for complex, real-world datasets. The identification of priority units based on similar characteristics highlights the strength of the clustering process. Regions grouped together share comparable levels of service pressure, which justifies their classification within the same cluster. This similarity provides a strong basis for comparative analysis. It also enables the development of uniform strategies for regions facing similar challenges. Such grouping enhances the efficiency of decision-making processes.

A reliable clustering outcome also reduces uncertainty in interpreting the results. When clusters are well-defined, the distinction between groups becomes clearer. This clarity supports more precise identification of problem areas. It also minimizes the risk of misclassification, which could lead to ineffective interventions. Accurate grouping is therefore essential for ensuring that strategies are appropriately targeted. The validation result further supports the integration of clustering analysis into operational and strategic planning. With a dependable model, the findings can be confidently used to guide resource allocation and performance improvement efforts. This integration enhances the practical relevance of the analysis. It also demonstrates the value of combining computational methods with real-world applications. Such an approach contributes to more effective and data-driven organizational management.

Service Performance Analysis Based on TMP

The discussion on TMP shows that PLN has responded quickly to simple new connection requests. However, the 55% completion rate also reveals that nearly half of the requests still require more time. Referring to public service theory, this condition can be caused by external factors (e.g., licensing and geographical conditions) and internal factors (e.g., technical capacity and resource limitations). Thus, improvements should not only target operational processes, but also inter-agency coordination and technical strategy optimization (Mulyadi et al., 2024). The presence of a significant portion of requests requiring extended completion time indicates that service performance is influenced by multiple interacting factors. These factors operate simultaneously and shape the overall efficiency of the service system. The complexity of handling different types of requests requires a flexible operational approach. A single standardized procedure may not be sufficient to accommodate varying service conditions. This situation highlights the need for adaptive service management.

External constraints play a crucial role in shaping service timelines. Conditions such as geographical accessibility and regulatory procedures introduce elements that are beyond immediate operational control. These factors can create unavoidable delays despite efficient internal processes. Addressing such challenges requires coordination beyond the organization itself. Strengthening collaboration with relevant stakeholders becomes essential to minimize these external barriers. Internal capacity also determines how effectively the organization can respond to service demands. The availability of skilled personnel, technical equipment, and operational infrastructure directly affects service speed. When these resources are limited, processing time tends to increase. This condition emphasizes the importance of capacity planning and resource optimization. Enhancing internal capabilities can significantly improve service performance.

The interaction between internal and external factors creates a layered challenge in service delivery. Improvements in one area may not produce optimal results without corresponding adjustments in others. For example, faster internal processes may still be constrained by regulatory delays. A comprehensive approach is therefore required to address all contributing elements simultaneously. This ensures that improvements are balanced and effective. A strategic perspective is necessary to manage these complexities in a sustainable manner. Organizations need to develop integrated solutions that combine operational efficiency with external coordination. This includes refining technical strategies while also improving communication with external entities. Such an approach supports more consistent service outcomes. It also enhances the organization’s ability to adapt to varying service conditions over time.

Operational Impact and Priority Strategies

The six UIDs/UIWs in the priority cluster have major strategic implications. The high waiting list burden in these areas should be used as a reference for resource allocation, such as adding technical personnel, increasing material availability, and accelerating survey coordination and connections. If this strategy is implemented, PLN can reduce the potential for compensation or sanctions due to non-fulfillment of TMP, while increasing public trust in the national electricity service provider (Tyas & Purnamasari, 2023). The designation of these six units as priority areas highlights the need for a more focused and responsive operational framework. Concentrating efforts in these regions allows the organization to address the most critical service pressures in a timely manner. This targeted approach ensures that interventions are not dispersed across all regions, but directed where they are most impactful. It also supports a more efficient use of available resources. Such prioritization strengthens the overall effectiveness of service management.

Enhancing technical capacity in these areas requires careful planning and execution. The addition of personnel must be aligned with the specific needs and workload characteristics of each region. Simply increasing the number of workers without considering skill compatibility may not yield optimal results. Training and specialization therefore become essential components of capacity development. This ensures that the workforce can handle both routine and complex service requirements effectively. Material availability is another critical factor that directly influences service completion time. Delays in the supply of equipment or installation components can disrupt operational schedules. Ensuring a steady and timely flow of materials helps maintain continuity in service processes. This requires improved inventory management and coordination with supply chains. Efficient material distribution supports faster and more reliable service execution.

Coordination processes, particularly in surveys and field connections, also play a significant role in improving service outcomes. Effective communication between administrative units and field teams ensures that each stage of the process is executed without unnecessary delays. Streamlining these coordination mechanisms reduces redundancy and enhances workflow efficiency. It also minimizes the risk of miscommunication that could affect service timelines. Strong coordination contributes to smoother operational performance. Implementing these strategies can create a positive cycle of improvement in service delivery. As performance improves, the likelihood of delays and associated risks decreases. This leads to greater customer satisfaction and reinforces confidence in the organization's capabilities. Sustained improvements in priority areas can also influence overall service perception at a national level. Over time, this contributes to a more reliable and trusted service system.

The Urgency of Continuous Improvement

The findings of this study emphasize the need for PT PLN (Persero) to adopt a data-driven approach to quality improvement. Although service performance for the 05-day category shows a positive trend, PLN still needs to strengthen monitoring, implement a quality control system, and identify the root causes in the 45% of cases that have not met the time standard. In addition, simplifying bureaucracy, digitizing field processes, and monitoring potential fraud are important elements in ensuring that new connections are faster, more transparent, and free of obstacles, in line with the recommendations in the research conclusion (Umagapi et al., 2023). From the method applied, namely the K-Means algorithm, the cluster results obtained from the UID/UIW with the highest number of new waiting lists were UID East Java with a total of 69,892, UID Central Java and DIY with 57,851, UIW SULSELRABAR with 56,870, UIW S2JB with 47,516, UID West Java with a total of 42,957, and UIW South and Central Kalimantan with a total of 41,755. The results of this analysis should be used as input for PT. PLN (Persero) to prioritize these UID/UIWs in improving electricity connection services that are easier, more transparent, less complicated, and free from any form of fraud.

The result of the test calculation using the Davies Bouldin Index showed that the cluster members have a fairly good similarity. The DBI is considered good if the value is close to 0. Then, the largest percentage of TMP criteria is 05 days, with the reason that the TMP criteria are without extension, meaning that the maximum completion time given is 5 (five) calendar days. No additional time or extension (expansion) is permitted, so the entire process must be completed within the specified time frame. If there is no "expansion" or fulfillment of requirements within those 5 days, there will be negative consequences related to PLN services, such as compensation or sanctions. The DBI outcome indicates that the clustering structure is sufficiently coherent to support meaningful interpretation of the grouped data. This level of similarity among cluster members suggests that each group represents a distinct operational condition rather than a mixed or ambiguous category. A well-formed structure allows the analysis to capture underlying patterns with greater clarity. It also reduces the risk of overlap between groups that could weaken the interpretation. Such a condition strengthens the analytical reliability of the clustering results.

The dominance of the 5-day completion criterion reflects the organization’s ability to operate within strict procedural boundaries. This category requires all processes to be completed within a fixed timeframe, which demands high coordination and efficiency. The absence of flexibility in time allocation places greater emphasis on operational precision. Every stage, from administrative verification to field execution, must function without delay. This requirement highlights the importance of time management within the service system. The fixed completion window also introduces a performance benchmark that must be consistently maintained. When the process is completed within the defined period, it reflects a well-functioning operational flow. However, any deviation from this timeframe can immediately be identified as a performance gap. This makes the criterion a clear and measurable indicator of service effectiveness. It also enables easier monitoring and evaluation of operational outcomes.

The potential consequences associated with unmet timelines create additional pressure on service performance. These consequences are not limited to operational inefficiencies but extend to organizational accountability. The presence of sanctions or compensation mechanisms reinforces the importance of meeting established standards. It also encourages stricter adherence to procedures and timelines. This accountability framework plays a role in maintaining service discipline. The combination of validated clustering and strict TMP criteria provides a comprehensive perspective on service performance. The clustering identifies patterns and priority areas, while the TMP framework defines the expected level of service execution. Together, they create a structured basis for evaluating operational effectiveness. This integrated view supports more informed decision-making. It also enables the organization to align analytical findings with practical service targets.

CONCLUSIONS AND RECOMMENDATIONS

PT. PLN (Persero) has been effective in handling new installations in the UID/UIW area within Cluster 1, as previously mentioned; however, further improvement is needed as the percentage stands at 55%.

ADVANCED RESEARCH

This research can be expanded further to include not only new connection transactions but also other types of transactions, such as power changes and reconnection, using the latest methods or approaches to obtain more accurate results. Combining cluster evaluation tests, such as internal cluster dispersion (icdrate) or the elbow method, to obtain more optimal values.

REFERENCES

Akromah, S., & Kusumasari, I. R. (2023). Penerapan Aplikasi PLN Mobile dalam Meningkatkan Pelayanan Pelanggan PLN ULP Bojonegoro Kota. Jurnal Masyarakat Mengabdi Nusantara, 2(2), 105–116. https://doi.org/10.58374/jmmn.v2i2.154.

Anggraeny, F., Eviyanti, A., & Sumarno. (2025). Penerapan K-Means dengan Evaluasi Davies-Bouldin Index untuk Pengelompokan Kelas Unggulan SMP Wijaya Sukodono. SMATIKA Jurnal: STIKI Informatika Jurnal, 15(2), 372–382. https://doi.org/10.32664/SMATIKA.V15I02.1689.

Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. Hello World Jurnal Ilmu Komputer, 2(1), 20–33. https://doi.org/10.56211/helloworld.v2i1.230.

Darmawan, S. D., Agushinta R, D., & Murniyati, M. (2022). Penerapan Algoritma K-Means Clustering pada Data Keluhan Pelanggan PT. PLN (Persero). Journal of Information System, Applied, Management, Accounting and Research, 6(2), 327–340. https://doi.org/10.52362/JISAMAR.V6I2.761.

Dilawati, H., Widianto, H., & Kuswiadji, A. (2024). Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means Clustering di Rumah Sakit Widodo Ngawi. BIOS : Jurnal Teknologi Informasi Dan Rekayasa Komputer, 5(2), 139–147. https://doi.org/10.37148/bios.v5i2.134.

Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokkan Data Obat dengan Silhouette Coefficient di Puskesmas Karangsambung. Journal of Applied Informatics and Computing, 5(2), 109–116. https://doi.org/10.30871/jaic.v5i1.3237.

Fernaldy, F. A., Arifiyanti, A. A., & Kartika, D. S. Y. (2025). Klasterisasi Tracer Study Alumni Universitas XYZ Menggunakan Algoritma K-Means. Jurnal Informatika Dan Teknik Elektro Terapan, 13(1). https://doi.org/10.23960/jitet.v13i1.5581.

Hartono, P. C., & Widiantoro, A. D. (2024). Analisis Prediksi Harga Saham Unilever Menggunakan Regresi Linier dengan RapidMiner. Journal of Computer and Information Systems Ampera, 5(3), 174–190. https://doi.org/10.51519/JOURNALCISA.V5I3.481.

Hikmah, D. U., & Rusdianto, R. Y. (2024). Prosedur Pelayanan Pelanggan terhadap Permohonan Pasang Baru, Perubahan Daya dan Migrasi Pada PT PLN (Persero) Unit Layanan Pelanggan Bojonegoro. Economics And Business Management Journal (EBMJ), 1(02), 66–72. https://doi.org/10.66084/EBMJ.V1I02.229.

Ifaldiansyah, M. R., & Hertati, D. (2023). Efektivitas Program Pemasangan Listrik Baru Menggunakan Aplikasi PLN Mobile di Wilayah Kerja Unit Pelaksana Pelayanan Pelanggan (UP3) Surabaya Selatan. Jurnal Ilmu Sosial Dan Ilmu Politik (JISIP), 12(3), 347–354. https://doi.org/10.33366/jisip.v12i3.2684.

Kapita, S. N., Mubarak, A., Abdullah, S. Do, & Fhadli, M. (2022). Penerapan Algoritma Clustering Khonen-Som dengan Validasi Davies Bouldin Index pada Pengelompokan Potensi Udang di Indonesia. IJIS - Indonesian Journal On Information System, 7(2), 134. https://doi.org/10.36549/ijis.v7i2.218.

Karmanita, D., Defit, S., & Sumijan. (2024). Metode K-Means dalam Mengukur Tingkat Pemahaman Materi Mata Kuliah Inti dan Penilaian Mahasiswa di Prodi Informatika. Jurnal KomtekInfo, 132–238. https://doi.org/10.35134/komtekinfo.v11i3.551.

Mulyadi, S., Insani, F., Agustian, S., & Afriyanti, L. (2024). Pengelompokan Data Pendistribusian Listrik Menggunakan Algoritma Mini Batch K-Means Clustering. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 4(3), 1051–1062. https://doi.org/10.57152/malcom.v4i3.1425.

Nuraeni, F., & Firdaus, M. H. M. (2025). Model Klasterisasi Data Pertanian Menggunakan Algoritma K-Means. Jurnal Algoritma, 22(2), 1–12. https://doi.org/10.33364/algoritma/v.22-2.2054.

Nuraeni, F., Susilawati, H., & Agustin, Y. H. (2023). Perbandingan Implementasi Algoritma K-Means++ dan Fuzzy C-Means pada Segmentasi Citra Wajah. JuTI "Jurnal Teknologi Informasi," 1(2), 47. https://doi.org/10.26798/juti.v1i2.722.

Nurwanda, M. D., Liviasari, A. N., & Pambudi, D. D. (2024). Penerapan K-Means untuk Meningkatkan Strategi Pemasaran Melalui Segmentasi Rumah Tangga yang Efektif. Jurnal Penelitian Ilmu Komputer, 2(3), 128–136.

Pasaribu, D. F., Damanik, I. S., Irawan, E., Suhada, & Tambunan, H. S. (2021). Memanfaatkan Algoritma K-Means dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat. BIOS : Jurnal Teknologi Informasi Dan Rekayasa Komputer, 2(1), 11–20. https://doi.org/10.37148/bios.v2i1.17.

Rohanifar, Y., Levine, R., Sultana, S., & Ahmed, S. I. (2026). Data Migration in HCI: The Politics of Invisible Borders, Informal Networks, and Immigrants' Data. Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems, 1–15. https://doi.org/10.1145/3772318.3790605.

Septianingsih, A. (2022). Analisis K-Means Clustering pada Pemetaan Provinsi Indonesia Berdasarkan Indikator Rumah Layak Huni. Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika, 3(1), 224–241. https://doi.org/10.46306/lb.v3i1.116.

Setiawan, D., & Perdhana, M. S. (2019). Analisis Keterlambatan Penyambungan Baru Listrik Menggunakan Business Model Canvas (Studi Kasus di PLN UID Jateng & DIY). Universitas Diponegoro.

Tyas, T. M. M., & Purnamasari, A. I. (2023). Penerapan Algoritma K-means dalam Mengelompokkan Demam Berdarah Dengue Berdasarkan Kabupaten. Blend Sains Jurnal Teknik, 1(4), 277–283. https://doi.org/10.56211/blendsains.v1i4.231.

Umagapi, I. T., Umaternate, B., Hazriani, H., & Yuyun, Y. (2023). Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index pada Pengelompokan Data Prestasi Siswa. Prosiding Sisfotek, 7(1), 303–308.