INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

CLUSROOF: Integrating Descriptive Analytics and K-Means Algorithm for Nishin Metal Corporation

Princess Junail Antonio., Ariel Kyla Estives., Aleck Jovy Marcial., Eleonor Jane Villegas., Gabby C. Vargas., Ezekiel N. Borja

(SY 2025-2026) Arellano University, Pasig Campus

DOI: https://doi.org/10.51583/IJLTEMAS.2025.1410000041

Received: 18 October 2025; Accepted: 24 October 2025; Published: 07 November 2025

Abstract: This project created a web-based system called ClusRoof for Nishin Metal Corporation, a company that offers commercial and industrial roofing services. The system helps the company manage customers and projects while also using the K-Means Algorithm to group customers based on their service requests. It includes an easy-to-use dashboard that shows important information such as project cost, duration, and type of service. The system also helps manage job orders, materials used, and service history. It was developed using Python, PHP, MySQL, JavaScript, and CSS, making it functional, secure, and simple to use. The system was tested and evaluated using the ISO 25010 standard, which checks quality in areas like functionality, reliability, usability, efficiency, and security. Based on feedback from both users and technical evaluators, ClusRoof was found to be effective, user-friendly, and reliable. It successfully met its goals by giving accurate customer grouping, clear data displays, and useful insights to help the company make better decisions. For future improvements, it is recommended to add new features like real-time payment options, automatic email notifications for inquiries, and an inventory module for materials. Continuous updates, security checks, and user support are also advised to keep the system running smoothly. Overall, this project shows that ClusRoof can help improve business processes, and customer service.

Keywords: ClusRoof System, K-Means Algorithm, Web-Based System, Data Analytics, Customer Management, ISO 25010, Project Monitoring, Python, PHP, MySQL, JavaScript, CSS.

I. Introduction

In today's modern world, businesses must find better ways to understand their customers and meet their needs. Nishin Metal Corporation, a company in the roofing industry, aims to improve its website by adding an interactive dashboard that allows the admin to view customer information, company sales, and summary reports. This project, called ClusRoof System, is a web-based platform that uses the K-Means Algorithm, a type of machine learning, to group customers based on their data and behavior.

According to Ioan-Daniel Borlea et al. (2022), the K-Means Algorithm remains widely used in data mining because of its ability to organize large sets of information quickly and efficiently. It helps reduce processing time while identifying hidden patterns within datasets.

Similarly, Das Saumendra and Janmenjoy Nayak (2023) explain that K-Means Algorithm dividing customers into groups with similar characteristics helps companies improve their marketing strategies and understand client preferences more effectively.

By grouping customers into categories, Nishin Metal Corporation can provide better services and manage projects more effectively. The ClusRoof System makes this possible through an easy-to-use web interface that collects, analyzes, and displays information clearly. It helps the company make smarter business decisions and improve customer satisfaction.

Currently, Nishin Metal Corporation does not have a data-driven system to analyze customer information. This makes it difficult to identify customer preferences and service demands, limiting the company's ability to improve its services. The ClusRoof System solves this problem by organizing client data, tracking job orders, and presenting insights through a dashboard. It does not collect sensitive information, making it safe and efficient for company use.

Scope

This project centers on creating a web-based system for Nishin Metal Corporation, a company in the roofing industry. The system combines descriptive analytics with the K-Means Algorithm to help the company better understand and manage its business data. One of its main features is an interactive dashboard that shows key information through simple visual charts and summaries. It presents insights such as total revenue, most requested services, and project statistics like cost, duration, and profit margins.

The system also includes several project management tools that allow the admin to record and organize customer details, job descriptions, materials used, service history, and financial data such as project costs and income. Another important part of the system is the report generation module, which gives a complete overview of the company's performance and operations. It helps management easily monitor progress and check the efficiency of each project.

The main types of reports generated by the system include:

- Project and Inquiry Status Report - This report shows the current stage of each project and the status of customer inquiries. It helps keep track of ongoing projects and pending requests, giving a clear overview of workloads.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

- Sales Trend Report – This report visually shows sales performance, helping the admin see business growth, seasonal changes, and overall profit trends.
- PDF Report Generation for Project Data – The system can create professional-looking PDF reports that summarize all important project details. These reports are well-organized, easy to share, and useful for both internal and external communication.

Overall, the ClusRoof System aims to make company operations more efficient, transparent, and data-driven. It helps both management and staff make smarter decisions, improve customer service, and maintain better communication within the organization.

Limitation

This project is limited to using the existing sales data stored in the company's database. It does not include any real-time data updates or predictive analytics features. The customer grouping function only uses the K-Means Algorithm and does not test or compare other clustering or classification methods.

In addition, the system does not include features such as automated marketing, customer feedback collection, or financial forecasting. These tools are not part of the current system design but may be added in future versions to expand its capabilities.

The testing and evaluation of the ClusRoof System were done only in controlled and academic environments, which means it was not fully deployed in an actual business setting. Because of this, the system's long-term effects on company performance were not measured. Future development may include real business testing and further improvements based on user experience.

Theoretical Framework

This framework explains the main ideas and concepts that support the project and help in understanding how it works.

- A common technique for organizing data is K-means clustering, which divides observations into clusters according to how close they are to the cluster center.
- Based on the distance between their centroids, the iterative K-means clustering algorithm divides a dataset into groups that are similar to one another. The centroid, also known as cluster center, is the median or mean of all the points in the cluster, depending on the data's properties (Kavlakoglu, 2024).

According to Winland (2024) K-means clustering works by using a mathematical measure, typically Euclidean, from the cluster center, the k-means clustering method groups data points into clusters. Reducing the total distance between data points and their designated clusters is the goal. The data points that fall into the same category are those that are closest to a centroid. A higher k value, or the number of clusters, signifies smaller clusters with greater detail, while a lower k value results in larger clusters with less detail.

Conceptual Framework

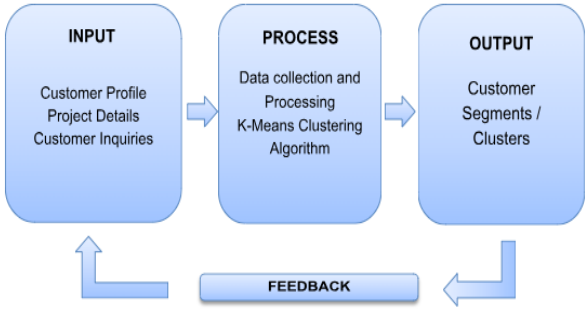

Figure 1: Input-Process-Output Mode

Input

Customers provide basic details like their name, location, and preferred roofing services. Sensitive information such as birthdays or bank details is not collected for safety. The main goal is to know how often customers buy and what products they prefer.

Process

The system uses the K-Means Clustering Algorithm to group customers based on their service requests, such as commercial or industrial roofing. This helps identify patterns and organize customers into meaningful groups.

Output

ClusRoof displays the results on a clear dashboard showing customer groups, service needs, and buying habits. This helps Nishin Metal Corporation plan better and assign the right teams for projects.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

Feedback

Administrators can adjust the number of clusters or features to improve accuracy. Users may also add insights or corrections, helping the system become more accurate and useful over time.

Significance of the Study

This study highlights the usefulness of using the K-Means Clustering Algorithm to improve how a company understands and serves its customers. It helps create better marketing strategies, strengthens customer relationships, and allows the company to offer more personalized services based on client needs.

This study could help the following individuals and groups:

- To Faculty and Students: This study gives teachers real examples of how K-Means Clustering can be applied in the roofing business industry. It can be used for classroom discussions, case studies, or future research about customer segmentation and business analytics. The study provides learning material about how machine learning can be used in real businesses. The students can use it as a reference for creating their own projects or theses related to consumer data and analytics.
- To the Administration: For company leaders, this research offers ways to solve existing problems in managing customer data. It helps improve marketing plans, system performance, and overall company services by applying the K-Means algorithm effectively.
- To IT Professionals: This study can provide insights into how the algorithm is integrated, specifically the K-means algorithm in business industry. This will serve as reference for them to improve business efficiency using data analytics and data-driven technology.
- To Future Researchers: This study can serve as a model or reference for others who want to explore clustering algorithms in different industries. It gives a guide on how data grouping can be applied to improve customer analysis and business decision-making.

II. Review of Related Literature

In today's digital era, customers are becoming smarter and more unpredictable. They often change their choices depending on their needs, which makes it important for companies to truly understand their customers. One effective way to do this is through customer segmentation, a method that groups people with similar needs or behaviours.

According to Jiaqi Zhao (2024), customer segmentation helps businesses strengthen relationships with customers by studying organized data to understand their buying habits and preferences.

Meanwhile, Fabregas et al. (2017), K-Means is a simple yet powerful clustering technique because of its speed, accuracy, and low memory use. It works well when the dataset has clear separations between groups. The algorithm's easy implementation and adaptability make it one of the most commonly used clustering methods in research and business applications.

Synthesis

The reviewed studies highlight that the K-Means clustering algorithm is a key tool for unsupervised learning and customer segmentation because of its speed, simplicity, and efficiency (Fabregas et al., 2017; Bagundang & Rael, 2021). However, some studies note its weaknesses, such as the need to set the number of clusters in advance. Improved versions like U-K-Means were developed to fix this issue (Sinaga & Yang, 2020; Ahmed et al., 2020).

Foreign research, such as Tabianan et al. (2022), shows how K-Means helps businesses group customers by buying behavior to improve marketing. Local studies like Paguio et al. (2015) and Tan et al. (2021) also prove its usefulness in classifying data based on certain features, such as roof types or customer traits.

While K-Means is widely used in many fields, there is little research applying it in the roofing or metal industry. This study addresses that gap through ClusRoof, a web-based system for Nishin Metal Corporation that uses K-Means to group customers, find sales patterns, and improve marketing strategies.

III. Methodology of the Study

The techniques that directed the course of this investigation are covered in this chapter. Research design, data collection, system development life cycle, database design, development tools, evaluation method and tool(s), study participants, and statistical tools utilized are all covered in the parts that follow.

The approach we used is applied research. It is focused on solving a practical problem faced by Nishin Metal Corporation. The company wants to improve the way they manage and accept client requests through the admin dashboard.

This study used Client Inquiry data, The data included information like what type of roofing services and location requested of the customers. The initial steps involved organizing the data and taking out any incorrect or incomplete records.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

Once completed, descriptive analytics was applied to organize and summarize the data. This helped in identifying basic patterns, like which services had most requests.

Data collection refers to the process of acquiring information from various sources, along with both primary and secondary data. Primary data is now gathered through many first-hand methods, such as with surveys, and with interviews, also observations, while secondary data comes from existing sources like with books, also journals, and online databases.

This study used the primary data using the survey questionnaire. The respondents must interact with the system prior to data collection. Doing so ensures relevant understandings will be obtained. Surveys can be conducted through Google Forms, whereby participants will provide important responses which can serve as key data sources that allow analysis.

The Software Development Life Cycle (SDLC) is a structured and efficient process that development teams follow to create high-quality web-based systems. It helps reduce costs and development time by providing clear steps for planning, designing, building, testing, and maintaining software. The main goal of SDLC is to minimize risks and ensure that the final product meets user needs and expectations.

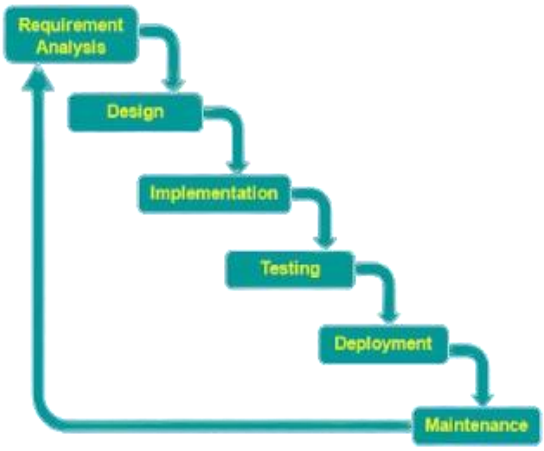

Figure 2: SDLC Model

The development of ClusRoof system was built using the Incremental Model, A software development approach where the system is incrementally designed, implemented, and tested until the complete system is completed. For ClusRoof, core modules involving customer data input along with descriptive analytics visualization and also K-Means clustering were developed incrementally. Each phase plans, implements, tests, also integrates. It does so by way of previously completed components in detail. Database design can be defined as a set of procedures or collection of tasks involving various steps taken to implement a database. And a well-structured database design helps you get the right information when you need it.

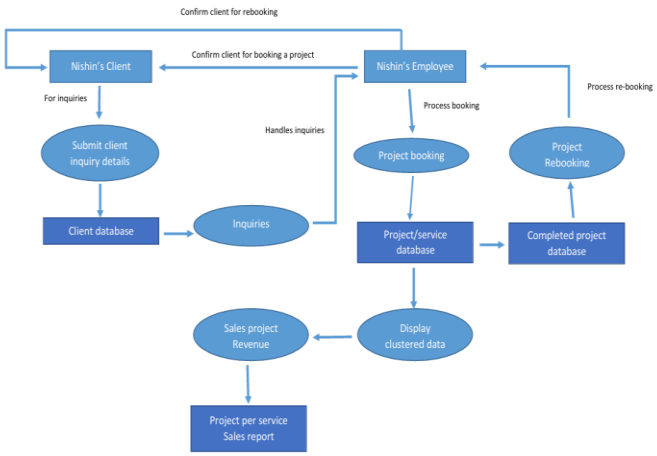

Figure 3: Data Flow Diagram

The system's operation is shown in this diagram. When a client provides their information, including personal and service-related facts, the procedure begins. Based on the data, the system then groups or segments the client's request using the K-Means Clustering algorithm. Conversely, a system administrator or authorized employee logs in.

After they have successfully logged in, the system administrator will process the booking which will be entered in the project/service database. It will then display the clustered data, sales project revenue, and project per service sales report. The completed project

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

database is based on the project/service database that was inputted by the admin. And, the project rebooking opens a new way for the client to be able inquire and request for another project they need in the future.

Respondents of the Study

The respondents of this study are divided into two groups: (a) the user group, which consists of staff from Nishin Metal Corporation, and (b) the technical group, composed of selected IT professionals. The total respondents are 50. There are 30 individuals for the user group and 20 respondents for the technical group. This method is chosen because it is simple, fast, and practical. It allowed us to gather information from people who are easily available and willing to take part. Instead of randomly selecting participants, the proponents of the study approached those who are already present or easy to reach. The surveys are conducted through Google Forms to make data collection more convenient and accessible.

Development And Evaluation Procedure

The system is developed using Python, PHP, and JavaScript as the main programming languages, with MySQL as the database for storing and managing data. The coding is done in Visual Studio Code, which is the main tool used for writing and editing the code. For hardware, the computer should have at least an Intel Core i5-7Y54 processor and 4GB of RAM to run smoothly. It must use a 64-bit version of Windows 10 Pro as the operating system. To access the system online, users can use popular web browsers like Google Chrome or Microsoft Edge.

Data Analysis Plan

Two evaluation techniques were used in the study. The first approach evaluated the application's usability and user satisfaction. The second approach assessed the application's usefulness and technical performance. Only one evaluation form is used, specifically following the ISO 25010 standards for assessment, even though two evaluation methodologies are used.

In this study, the system is evaluated to satisfy the ISO 25010 criteria such as:

- Functional Sustainability - The degree to which a product or system meets all stated and implied needs when used under the specified conditions.
- Reliability - The degree to which a system, product, or component can carry out particular tasks under given circumstances for a particular period of time.
- Efficiency - This refers to the way the system works when a certain number of resources are being used.
- Usability - The system can be used by specific users to achieve goals with effectiveness, efficiency, and satisfaction within particular contexts.
- Security - The system secures data and information to allow users, services, or other systems to access it according to their authorization levels and varieties.

The study applied the following statistical tools. These are the samples of statistical tools normally used in quantitative research:

- To measure the effectiveness of the ClusRoof platform, a five-point Likert scale was utilized, where respondents selected from five possible choices for each survey question. The weighted mean formula was used to compute the average rating of the system's evaluation based on its performance, usability, and effectiveness.
- Frequency distribution is the detailed breakdown of responses for each question according to the study's categories or the predefined survey options. This method provides insight into how responses are spread, helping to better understand participants' feedback related to ClusRoof: Integrating Descriptive Analytics and K-Means Algorithm for Customer Segmentation in a Roofing Industry for Nishin Metal Corporation.
- A Likert scale is a quantitative analysis data collection instrument used in research and evaluations to gauge people's views, beliefs, and attitudes. This scale asks respondents a series of questions or comments. The responses are given numerical values so that the data can be analyzed quantitatively. The options on a standard Likert scale would be "strongly disagree" and "strongly agree." This study uses a 4-point Likert scale, with 1 representing "strongly disagree" and 4 representing "strongly agree."

The System

The output of the study is a web-based app called "ClusRoof System". This is a web-based system designed for Nishin Metal Corporation, a company in the roofing industry. The system combines customer and project management with descriptive analytics using the K-Means algorithm. It provides an interactive dashboard that shows important insights such as services, project costs, and duration of the project. It also has tools for handling job orders, including customer details, materials used, and service history.

The system is built using web development tools such as Python, MySQL, JavaScript, PHP, and CSS. These tools allow the system to be user-friendly, functional, and efficient. It is evaluated based on ISO 25010 standard, which checks important qualities like

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

functional sustainability, reliability, usability, efficiency, and security. The system will be tested to make sure it works properly, is safe, and easy to use.

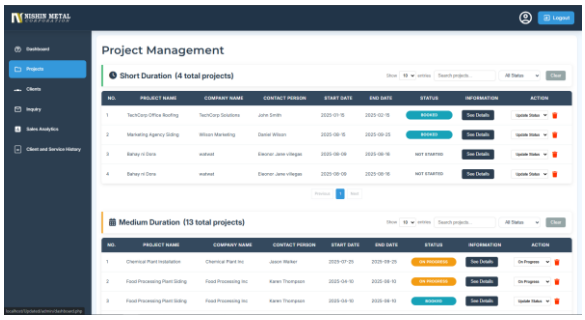

Figure 4: Project Page

This page is where the algorithm is applied that is based on the nearest centroid. It typically displays the clustered data based on the duration of the project and then clusters it after re-computing the nearest centroid into mean.

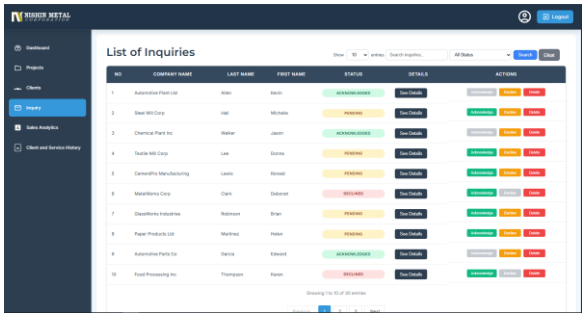

Figure 5: Inquiry Page

This page displays the list of inquiries from the client side. This will manage the booking inquiry whether if its acknowledged or declined.

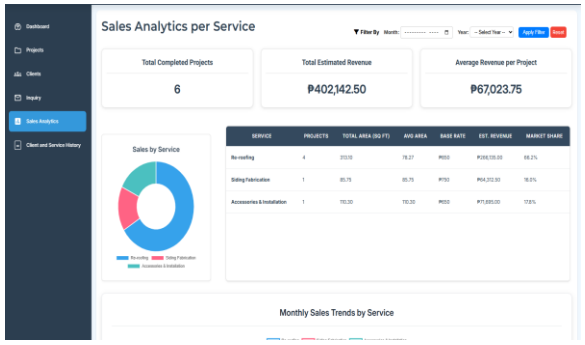

Figure 6: Sales Trend Report

This page is applied with descriptive analytics where it displays the past data that are made. It monitors on what will be the estimated revenue, sales, and etc.

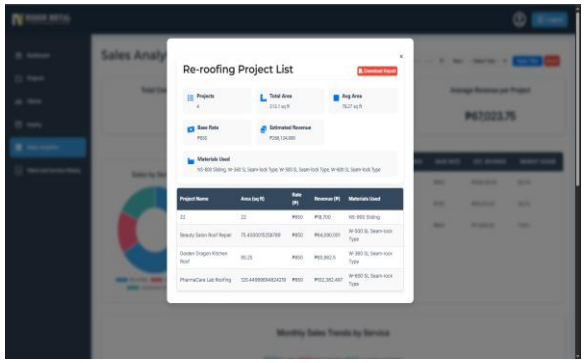

Figure 7: PDF Report Generation for Project Data

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

The modal displays the summarization of data in each project. Along with this is an option of pdf generation, including the list of projects.

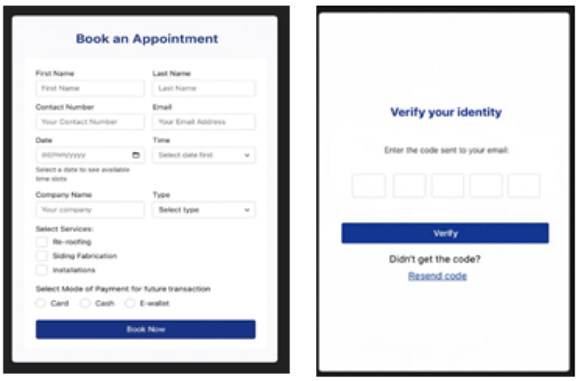

Figure 8: Payment Transaction

When a client schedules an appointment, a payment transaction is included, as you can see on the book an appointment page. A code delivered to the customer's email serves as the primary means of confirming the client's identity on a subsequent page once the clients select the payment method.

Assessment: Summary Of Respondents on the System

Respondents (groupings)	Size (n)	Percentage
Users	30	60.0%
Technical	20	40.0%
Total (n)	50	100.0%

Table 1: Distribution of Respondents

Table 1 shows the total number and percentage of respondents who participated in the evaluation activities. There are 30 User respondents or 60% of the total, and 20 Technical respondents or 40% of the total who joined the evaluation of the application.

The evaluations of the technical and user responders are included in the following tables, along with a summary and comparison of the two groups' evaluations.

Criteria (ISO25010)	Respondents (50)			
	Users (30)		Technical (20)	
	WM	VI	WM	VI
1. Functional Sustainability	3.3	SA	3.5	SA
2. Reliability	3.5	SA	3.8	SA
3. Efficiency	3.6	SA	3.8	SA
4. Usability	3.6	SA	3.7	SA
5. Security	3.3	SA	3.9	SA
Overall Average Mean	3.5	SA	3.7	SA

Table 2: Summary & Comparison of Evaluations of Respondents

Table 2 shows the summary and comparison of evaluations of respondents based on the criteria of ISO25010. Overall, both user & technical respondents Strongly Agreed on the acceptability and usability based on the criteria of ISO 25010.

Ethical Considerations

The study guarantees the confidentiality and integrity of the data collected from participants. Information provided by respondents is secure, and no personality identifiable information is shared without permission. Respondents are free to leave the study at any moment without facing any repercussions, in accordance with the principles of voluntary participation. Strict adherence to data security protocols guards against misuse and illegal access to information. Lastly, in order to preserves the study's credibility, all results are presented truthfully and accurately, free from bias or manipulation.

Summary

The ClusRoof is a web-based system made for Nishin Metal Corporation, a company that offers commercial and industrial roofing services. The system uses the K-Means Algorithm to group customers based on the services they request. It has a simple dashboard that helps the admin organize customer information, keep track of the project or services that they request. The system is user-

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VI, June 2025

friendly and easy to navigate based on user and technical groups. Both user and technical group evaluated ClusRoof using the ISO 25010 standards. This study is significant because it helps the company understand its customers better and improve services.

IV. Conclusion

The result of the evaluation reveals that both end-users and technical users strongly agreed in the effectiveness and reliability of the ClusRoof system. The platform successfully met its objectives by providing accurate customer segmentation, user-friendly data visualization, and actionable insights that enhance decision-making for Nishin Metal Corporation. The system demonstrates its potential as a valuable tool for improving business strategies, customer engagement, and overall operational efficiency.

Recommendation

Future researchers are encouraged to expand the ClusRoof System by improving the integration of data analytics and the K-Means algorithm, considering the limitations of this study. It is recommended to add useful feature like a material inventory.

Furthermore, putting in place technical support and long-term maintenance procedures will help ensure the system's stability and efficiency after its initial launch. This may include regular security updates, user assistance, and automatic upgrades.

References

1. Tabianan, K., Velu, S., & Ravi, V. (2022b). K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data. *Sustainability*, *14*(12), 7243. https://doi.org/10.3390/su14127243
2. Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K., & Hanzo, L. (2016b). Machine Learning Paradigms for Next-Generation Wireless Networks. *IEEE Wireless Communications*, *24*(2), 98–105. https://doi.org/10.1109/mwc.2016.1500356wc
3. Ahmed, M., Seraj, R., & Islam, S. M. S. (2020b). The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. *Electronics*, *9*(8), 1295. https://doi.org/10.3390/electronics9081295
4. Tan, L. R. E., Acosta, T. J. S., Gumaro, J. J. C., Agar, J. C., Tingatinga, E. a. J., Plamenco, D. a. D., Ereno, M. N. C., Musico, J. K. B., Pacer, J. S., Baniqued, J. R. D., Hernandez, J. Y., & Villalba, I. B. O. (2021). Investigation of the effects of the classification of building stock geometries determined using clustering techniques on the vulnerability of galvanized iron roof covers against severe wind loading. IOP Conference Series Materials Science and Engineering, 1150(1), 012024. https://doi.org/10.1088/1757-899x/1150/1/012024
5. Bagundang, E., & Rael, C. (2021). Clustering commercial and residential electricity consumption using K-Means algorithm. International Journal of Scientific and Technology Research, 10(8), 93–96. https://www.ijstr.org/final-print/aug2021/Clustering-Commercial-And-Residential-Electricity-Consumption-Using-K-means-Algorithm.pdf
6. Fabregas, A. C., Gerardo, B. D., & Tanguilig, B. T., III. (2017). Enhanced initial centroids for K-Means algorithm. International Journal of Information Technology and Computer Science, 9(1), 26–33. https://doi.org/10.5815/ijitcs.2017.01.04
7. Saumendra Das and Janmenjoy Nayak (2022).Customer Segmentation via Data Mining Techniques State-of-the-Art Review https://www.researchgate.net/publication/360413389_Customer_Segmentation_via_Data_Mining_Techniques_State-of-the-Art_Review
8. Ioan-Daniel Borlea, Radu-Emil Precup, Alexandra-Bianca Borlea (2022). Improvement of K-means Cluster Quality by Post Processing Resulted Clusters https://www.sciencedirect.com/science/article/pii/S1877050922000096
9. Manish Suyal and Sanjay Sharma (2024). A Review on Analysis of K-Means Clustering Machine Learning Algorithm based on Unsupervised Learning https://www.researchgate.net/publication/379878557_A_Review_on_Analysis_of_K-Means_Clustering_Machine_Learning_Algorithm_based_on_Unsupervised_Learning