Journal of Research and Development 11[26]1-10: e41126110
Article

Analytical platform for operational performance assessment using Random Forest and K-Means

Plataforma analítica para la evaluación del desempeño operativo utilizando Random Forest y K-Means

Vázquez-De los Santos, Laura Cristina[a], Burciaga-Alarcón, Ricardo*[b], Rodríguez-Silva, Jesús Rolando[c] and Rodríguez-Arzola, Adrián[d]

[a] Universidad Autónoma de Coahuila • 0000-0002-0291-7774 • 615088
[b] Universidad Autónoma de Coahuila • 0000-0001-9711-6824 • 2064690
[c] Universidad Autónoma de Coahuila • 0000-0001-8410-9174 • 492229
[d] Universidad Autónoma de Coahuila • 0009-0002-1112-2372 • 2195267

SECIHTI classification:

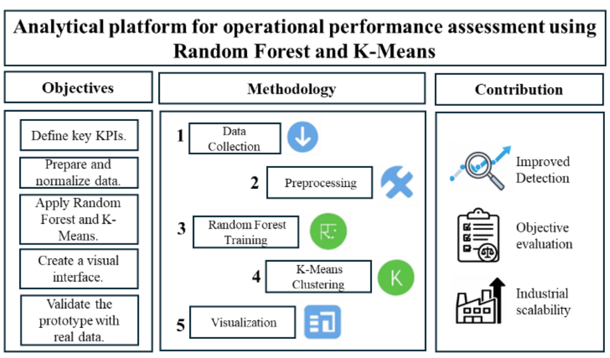
Area: Engineering
Field: Technological Sciences
Discipline: Computer Technology
Subdiscipline: Automated production systems

https://doi.org/10.35429/JRD.2024.11.26.4.1.10
Article History:
Received: January 30, 2025
Accepted: December 05, 2025

* ✉ [rburciaga@uadec.edu.mx]

Abstract

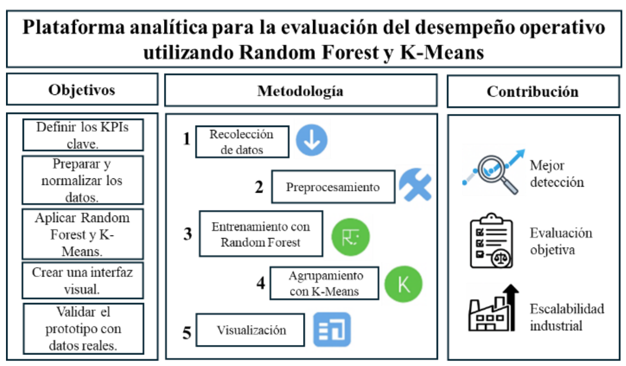
This study presents an analytical platform for assessing operational performance in industrial environments through the use of Random Forest classification and K-Means clustering. Five key performance indicators—scrap rate, cycle time, quality, attendance, and supervisor-based performance rating—were selected to evaluate a dataset of 293 operators from different production areas. After data cleaning and normalization, the Random Forest model achieved an accuracy of 89%, correctly classifying operators into High, Regular, and Critical categories. In parallel, K-Means generated three performance clusters that highlight distinct behavioral patterns among operators. The system was developed with a modular architecture using FastAPI, PostgreSQL, and React.js, and includes interactive dashboards created in Power BI to support decision-making. Practical validation showed benefits such as increased objectivity, transparency, and improved identification of both high and low performers. The results demonstrate that data-driven methods can enhance operational evaluation and provide a foundation for scalable industrial applications.



Random Forest, K-Means Clustering, Operational Performance

Resumen

Este estudio presenta una plataforma analítica para evaluar el desempeño operativo en entornos industriales mediante Random Forest y K-Means. Se seleccionaron cinco indicadores clave: tasa de scrap, tiempo de ciclo, calidad, asistencia y evaluación del supervisor, aplicados a un conjunto de 293 operadores de distintas áreas de producción. Tras la limpieza y normalización de datos, el modelo Random Forest alcanzó una precisión del 89%, clasificando correctamente a los operadores en categorías Alto, Regular y Crítico. Paralelamente, K-Means generó tres grupos que revelan patrones de comportamiento diferenciados. El sistema se desarrolló con arquitectura modular usando FastAPI, PostgreSQL y React.js, e incluye paneles interactivos en Power BI para apoyar la toma de decisiones. La validación práctica mostró mayor objetividad, transparencia e identificación más clara de desempeños altos y bajos. Los resultados evidencian que el análisis basado en datos fortalece la evaluación operativa y permite aplicaciones industriales escalables.



Random Forest, Agrupamiento K-Means, Desempeño Operativo

Area: Development of strategic leading-edge technologies and open innovation for social transformation

Citation: Vázquez-De los Santos, Laura Cristina, Burciaga-Alarcón, Ricardo, Rodríguez-Silva, Jesús Rolando and Rodríguez-Arzola, Adrián. [2025]. Analytical platform for operational performance assessment using Random Forest and K-Means. 11[26] 1-10: e41126110.

Journal of Research and Development 11[26]1-10: e41126110
Article

Introduction

In today's highly competitive industrial environment, organizations must sustain elevated levels of efficiency, quality, and operational consistency to ensure their long-term viability. Within this context, the performance of line operators becomes a strategic variable, as their daily actions directly affect productivity, scrap generation, process stability, and overall operational outcomes. Accurate performance assessment is therefore essential not only for measuring results but also for identifying improvement opportunities, developing workforce capabilities, and strengthening organizational culture.

However, despite its importance, many manufacturing plants continue to rely on manual, subjective, and low-frequency evaluation methods that offer limited feedback, hinder timely decision-making, and may generate distrust among personnel.

The emergence of Industry 4.0 has transformed this landscape by introducing digital tools that enable the automation of data capture, processing, and interpretation. Artificial intelligence and machine learning allow organizations to analyze large datasets, recognize behavioral and operational patterns, and generate precise, real-time predictions that support more objective and consistent evaluations. These technologies facilitate the transition from intuition-based assessments toward evidence-based decision systems aligned with continuous improvement principles.

This article presents the development of a functional prototype of an intelligent digital system designed to assess operator performance in industrial settings. The proposed solution integrates key performance indicators (KPIs) with classification algorithms (Random Forest) and clustering techniques (K-Means), enabling both predictive modeling and the segmentation of operational profiles. Additionally, a user-friendly visual interface provides supervisors with accessible, timely insights that enhance fairness, transparency, and strategic decision-making. The results demonstrate the potential of combining data analytics with modern computational techniques to strengthen operational management and promote more efficient, equitable, and data-driven industrial environments.

Conceptual Foundations

Performance Evaluation in the Manufacturing Industry

In the maquiladora environment, performance evaluation is essential for identifying operational strengths and weaknesses, directly impacting overall efficiency. Indicators such as the number of units produced, adherence to standard cycle times, errors per shift, product quality, and defect rates are widely recognized as critical metrics for this purpose [Vega et al., 2022]. These indicators enable continuous monitoring of performance and support decision-making processes focused on operational improvement.

Beyond technical aspects, human factors such as job attachment, motivation, and recognition also exert a significant influence on performance. Their omission may lead to an underestimation of human talent and to a limited interpretation of operational results [Soto et al., 2024]. Likewise, the presence of an effective internal control system contributes to greater productivity and operational discipline, while its absence has been associated with recurrent failures and lower individual performance [Valdez & Salcido, 2021].

Furthermore, in sectors such as the textile industry, logistical conditions, production infrastructure, and supply chain management have been shown to directly affect the quality of operator performance. Factors such as material availability, efficiency of production flows, and workspace organization influence daily consistency and stability in operator output [International Labour Office, 1998; Algawatta & Jayasekara, 2025; Rahman et al., 2024]. These elements highlight the need to consider not only technical variables but also human and contextual factors to achieve more comprehensive and accurate performance assessments in industrial environments.

Lean Manufacturing, Technical Processes, and Digitalization

The Lean approach has contributed significantly to improving efficiency in various maquiladora environments; however, its full potential is often constrained when implementation is only partial, particularly when Lean principles are not fully integrated into technical and production processes [Chávez, 2022].

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

Vázquez-De los Santos, Laura Cristina, Burciaga-Alarcón, Ricardo, Rodríguez-Silva, Jesús Rolando and Rodríguez-Arzola, Adrián. [2025]. Analytical platform for operational performance assessment using Random Forest and K-Means. 11[26] 1-10: e41126110
https://doi.org/10.35429/JRD.2025.11.26.4.1.10

Partial adoption typically results in fragmented workflows, limited standardization, and inconsistencies that prevent organizations from achieving the continuous improvement and process stability that Lean aims to promote.

In this context, digitalization has emerged as a critical complement to Lean initiatives. Digital tools such as interactive dashboards and real-time visualization systems provide immediate access to operational data, enabling supervisors and managers to continuously monitor performance, detect anomalies, and respond promptly to deviations that may compromise productivity or quality [Vega et al., 2022; Guzmán-Anaya, 2019]. These technologies support more agile and data-driven decision-making, strengthening the alignment between operational practices and Lean objectives.

In the textile sector specifically, researchers have emphasized the pressing need to modernize information systems to comply with international standards and to strengthen traceability across production and supply-chain operations. Improvements in digital infrastructure are essential for enhancing responsiveness, reducing uncertainty, and ensuring that production flows meet the growing demands of global markets [Díaz et al., n.d.; Silva-Castro et al., 2025]. These findings collectively underscore the importance of integrating Lean methodologies with robust digital systems to achieve greater efficiency, transparency, and operational consistency in modern manufacturing environments.

Application of Machine Learning

The application of machine learning has proven highly effective for processing large datasets and enhancing analytical decision-making in industrial environments. Algorithms such as Random Forest have demonstrated strong performance in multivariable classification tasks, particularly in scenarios where numerous operational indicators must be evaluated simultaneously. Their robustness to noise, ability to model nonlinear relationships, and capacity to handle heterogeneous feature sets make them well-suited for complex production systems [Kang et al., 2020; Tobar-Díaz et al., 2023].

These characteristics are especially valuable in maquiladora settings, where operator performance can be influenced by dynamic interactions between time, quality, workload, and equipment conditions.

Similarly, unsupervised learning techniques such as K-Means provide powerful tools for exploring underlying behavioral patterns in operational data. By grouping operators into clusters based on similar measures, K-Means enables the identification of natural performance profiles—ranging from highly consistent operators to those exhibiting irregular or risk-prone patterns—even when no prior labels or categories are available [Sandoval, 2018]. This capacity to discover latent structures is critical in environments seeking to implement preventive and improvement-oriented strategies.

When these analytical models are integrated into interactive graphical interfaces, traditional evaluation practices evolve into automated, transparent, and feedback-driven systems. Interfaces designed to visualize classifications, cluster assignments, and deviations allow supervisors to interpret results more easily, accelerating corrective action and promoting continuous improvement [Tobar-Díaz et al., 2023]. Such systems reduce subjective bias, standardize evaluation criteria, and strengthen the alignment between human resource management and operational performance.

Moreover, advances in artificial intelligence—supported by emerging fields such as quantum computing and nanotechnology—are expanding the predictive and computational capabilities available to industry, pointing toward increasingly sophisticated and autonomous evaluation mechanisms [Díaz-Ramírez, 2021]. These developments reinforce the need for organizations to adopt intelligent, data-driven tools capable of adapting to growing operational complexity. In highly competitive industrial contexts, the integration of supervised and unsupervised learning models becomes a strategic asset, enabling companies to anticipate issues, optimize performance, and maintain resilience amid rapidly evolving global demands.

Recent studies have also explored hybrid machine-learning approaches that combine supervised and unsupervised models to enhance analytical accuracy.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

For example, in the domain of connected healthcare systems, an enhanced Random Forest classifier integrated with K-Means clustering was used to detect and categorize anomalous behaviors in IoMT networks, demonstrating superior performance compared with standalone algorithms [Al-Abadi et al., 2023].

Although applied in a different domain, this work reinforces the effectiveness of combining classification and clustering techniques for improving decision-making in complex operational environments.

Additionally, work in textile manufacturing illustrates how AI-based systems are transforming traditionally manual and error-prone inspection processes. A recent study implemented an enhanced deep convolutional neural network (DCNN) integrated with robotic motion and sensor-based alerts, achieving automatic classification of 13 types of fabric defects with a mean Average Precision of 97.49%.

These results highlight the effectiveness of AI-driven inspection systems and reinforce the broader industrial shift toward objective, data-driven evaluation tools—an approach aligned with the intelligent operator-assessment framework proposed in this study [Hassan et al., 2024].

Integration of an Intelligent Evaluation System

The proposed technological architecture is built upon modern tools that enable efficient data management, analytical processing, and user-friendly visualization. FastAPI (Python) serves as the backend framework, providing a high-performance environment for handling requests, executing analytical models, and managing system logic. PostgreSQL is used as the main database engine, offering reliability, scalability, and robust support for structured operational records.

On the client side, React.js allows the development of dynamic and responsive interfaces capable of displaying performance indicators clearly and interactively. During the initial design stages, visualization modules were prototyped using Power BI to generate rapid dashboards and intuitive data exploration views [Guzmán-Anaya, 2019; Tobar-Díaz et al., 2023].

This technological structure supports the continuous ingestion, processing, and interpretation of operator-performance data, enabling the platform to provide timely analytical insights and to integrate machine-learning components into the evaluation workflow. Its modular and extensible design also makes it adaptable to a wide range of production contexts. In sectors such as apparel and textile manufacturing, previous studies have demonstrated the relevance of monitoring variables such as operation times, stitching quality, and process-control elements to reduce production failures, improve order fulfillment, and strengthen the consistency of manufacturing flows [Silva-Castro et al., 2025]. These findings reinforce the importance of intelligent evaluation systems capable of aligning digital infrastructure with the specific requirements of each production environment.

Methodology

This study is framed within an applied technological research approach, focused on the development and experimental validation of a functional prototype for evaluating operator performance. The methodology follows a quantitative orientation and a non-probabilistic experimental design. Its main objective was to develop and validate an intelligent evaluation system grounded in objective indicators and machine-learning algorithms, structured into five core phases.

A. Data Collection and Structuring

Five priority KPIs were identified for assessing operational performance: average scrap (number of defective units per operator), cycle time (average duration required to produce a single unit), quality (ratio of accepted to total produced units), attendance (percentage of days worked), and a global operator evaluation (categorized as Critical, Regular, or High based on supervisor criteria). These indicators were automatically captured from control sheets and operational databases, ensuring consistency and minimizing manual intervention. The dataset incorporated a total of 293 operators from different production areas, each anonymized using identification codes (e.g., OP297, OP613) to guarantee confidentiality while enabling full traceability of performance patterns across the evaluation period.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

This structured dataset provided a reliable foundation for the subsequent analytical phases and for the implementation of the machine-learning models integrated into the intelligent evaluation platform.

B. Data Preprocessing

The collected information was subjected to a comprehensive preprocessing stage that included data cleaning, statistical normalization, and manual verification to ensure internal consistency and reliability. The following techniques were applied during this phase:

- Data normalization, used to standardize the scale of numerical indicators and prevent disproportionate weighting in subsequent analytical models.

- One-hot encoding, implemented to convert categorical variables into binary vectors, allowing supervised algorithms to interpret them correctly.

- Training–testing split, performed to generate independent datasets for model training and performance evaluation, ensuring objective measurement of generalization capacity.

All preprocessing procedures were carried out using Python, primarily through the Pandas library for data manipulation and Scikit-learn for machine-learning transformations. This step ensured that the dataset was clean, structured, and computationally compatible with the supervised and unsupervised models used in later phases, forming a robust foundation for accurate analytical modeling.

C. Modeling with Machine Learning Algorithms

For the system's modeling stage, a complete workflow was implemented in Python integrating the selected algorithms through the Scikit-learn library. The Random Forest model was trained using historical operator records that had been previously labeled by supervisors. The normalized KPIs served as input features, ensuring that all variables contributed proportionally to the model's decision structure.

The dataset was partitioned into training and testing subsets, and cross-validation techniques were applied to guarantee consistency, reduce overfitting, and assess the model's generalization capability.

In parallel, the K-Means algorithm was applied using the same standardized indicators, which were normalized through z-score transformation to ensure comparable feature scales. The value of $k = 3$ was defined to generate three operator clusters, corresponding to distinct performance profiles. The k-means++ initialization method was used to improve centroid selection and promote stable convergence, reducing the likelihood of suboptimal clustering results.

Both models were encapsulated into modular Python scripts and integrated into the system's backend, allowing them to be executed automatically whenever new operator data are received. This modular architecture ensures scalability, maintainability, and seamless communication between the analytical components and the visualization interface, enabling real-time performance analysis within the intelligent evaluation platform.

D. Visual Interface Design

An interactive dashboard was developed as a functional prototype using Power BI (Fig. 2), with the purpose of facilitating the visual interpretation of results for supervisors and talent-management personnel. The interface was designed to provide a clear, accessible, and intuitive overview of operator performance, allowing users to explore the analytical outputs generated by the system.

Its main features include the visualization of key performance indicators for each operator, graphical distributions of evaluation outcomes, historical trends showing the evolution of performance over time, and the segmentation of operators based on cluster assignments. The dashboard also incorporates dynamic tables, filters, and color-coded indicators (similar to traffic-light systems) to highlight critical conditions and facilitate rapid decision-making. These elements collectively enhance usability and support a more transparent, data-driven evaluation process.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

Journal of Research and Development 11[26]1-10: e41126110
Article

Box 1

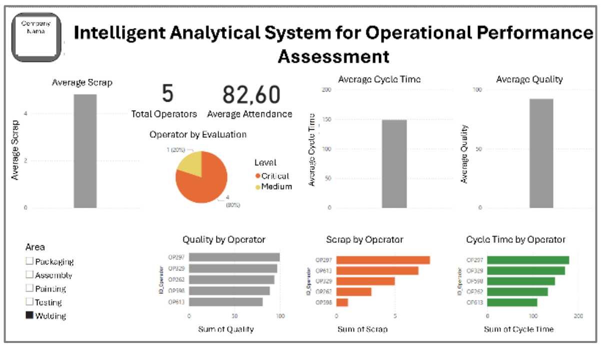


Figure 1
Dashboard of the intelligent system for operator performance evaluation
Source: Own Elaboration

The system was validated in a controlled testing environment using real operational data, replicating the complete evaluation workflow under conditions similar to those of a real production setting. Throughout this process, supervisory personnel highlighted several practical advantages. Among the most frequently mentioned were the increased objectivity and transparency of the evaluations, the reduction in the time required to generate performance reports, and the system's ability to reveal both high-performing operators and previously unnoticed low-performance cases. Supervisors also reported an improvement in the perceived fairness of the evaluation process, noting that the analytical criteria provided clearer and more consistent justification for classification outcomes. Although this validation was experimental and limited to a testing environment, the results demonstrate strong potential for future deployment in real industrial contexts. The system's capacity to integrate diverse performance indicators, automate analytical workflows, and generate interpretable visual outputs positions it as a valuable tool for organizations seeking to standardize evaluations and support data-driven decision-making.

System Architecture

The system was designed following a modular and scalable architectural approach based on modern web technologies, as illustrated in figure 2. This structure allows each component—backend services, database management, machine-learning models, and visualization interfaces—to operate independently while maintaining seamless communication across the platform.

Such modularity greatly simplifies maintenance, facilitates the integration of new analytical models or data sources, and enhances the system's adaptability to different operational areas within an organization.

This architectural flexibility ensures that the platform can evolve in response to new evaluation requirements, additional KPIs, or changes in production dynamics, making it a robust and future-oriented solution for intelligent performance evaluation.

Box 2

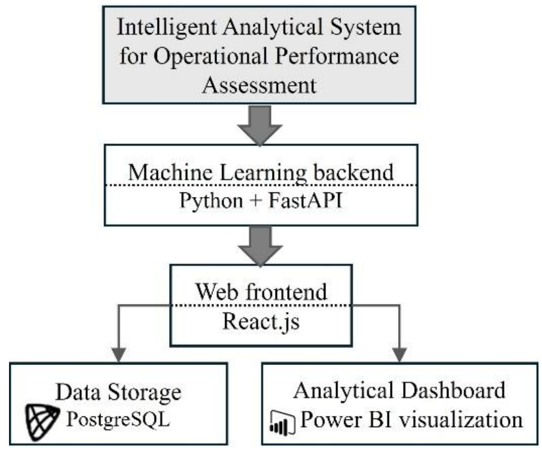


Figure 2
General system architecture with AI components and visualization
Source: Own Elaboration

Results and Discussion

As part of the development of the functional prototype, the machine-learning models (Random Forest and K-Means) were integrated and executed using a real operational dataset. These models constituted the core analytical components of the system, and their performance was evaluated alongside the data-processing workflow and the visualization interface. This joint evaluation made it possible to observe not only the predictive accuracy of each model but also their practical contribution to the decision-support capabilities of the intelligent evaluation platform.

A. Results of the Random Forest Model

The supervised classification model (Random Forest) achieved an accuracy of 89%, correctly assigning operators to three performance categories: High, Regular, and Critical, as shown in Table I.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

Vázquez-De los Santos, Laura Cristina, Burciaga-Alarcón, Ricardo, Rodríguez-Silva, Jesús Rolando and Rodríguez-Arzola, Adrián. [2025]. Analytical platform for operational performance assessment using Random Forest and K-Means. 11[26] 1-10: e41126110
https://doi.org/10.35429/JRD.2025.11.26.4.1.10

This level of accuracy indicates that the model was able to capture relevant relationships among the key performance indicators (KPIs) and reproduce the supervisor-assigned labels with notable consistency.

The model's output also contributed to a more objective interpretation of operator performance by reducing subjective bias and providing transparent criteria for each classification. In the context of an industrial environment—where evaluation processes are often manual or partially standardized—this represents a significant improvement. The integration of Random Forest within the system's architecture allowed the classifications to be visualized in real time through the dashboard, enabling supervisors to identify patterns, detect emerging risks, and make more informed decisions regarding personnel management and process improvement.

Box 3

Table 1
Prediction Results

Operator ID	Scrap (%)	Time (s)	Quality (%)	Attendance (%)	RF Prediction
OP297	1.2	12.3	98.7	100	High
OP302	4.5	15.0	91.2	92	Medium
OP613	8.7	18.5	84.3	85	Critical
OP417	2.8	141	95.5	96	Medium
OP509	0.9	11.8	99.1	100	High

Source: Own Elaboration

Figure 3 presents the confusion matrix generated by the Random Forest model when classifying operator performance. Since all instances fall along the main diagonal, the results indicate complete agreement between the model's predictions and the actual evaluations conducted by supervisors. This perfect alignment demonstrates that, under the testing conditions, the model was able to accurately capture the patterns encoded in the training data and reproduce the supervisory criteria without misclassifications.

Such behavior reflects a notably high level of predictive reliability, suggesting that the selected KPIs carry sufficient discriminative power to differentiate between the performance categories.

Moreover, this outcome reinforces the model's potential as a robust decision-support tool, capable of contributing to more objective, consistent, and transparent evaluation processes in operational environments. The results also highlight the feasibility of integrating supervised learning techniques into daily management practices, particularly in contexts where rapid and data-driven assessments are essential.

Box 4

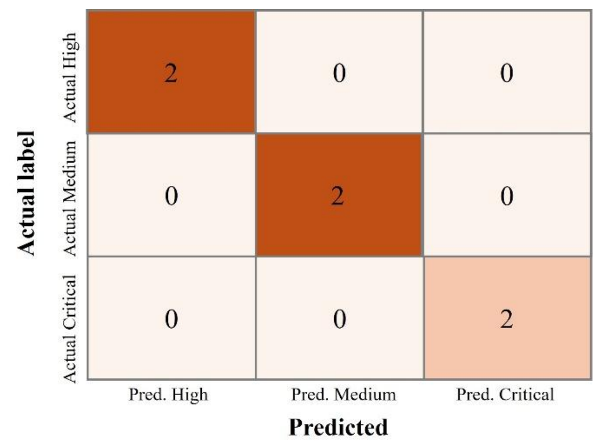


Figure 3
Confusion Matrix of the Random Forest Model
Source: Own Elaboration

B. Results of the K-Means Model

Figure 4 presents the comparison of average KPI performance values across the three clusters generated by the K-Means algorithm. Cluster 0 corresponds to the group of high-performing operators, characterized by consistently superior values in all indicators, including lower scrap, shorter cycle times, higher quality, and stronger attendance records. Cluster 1 represents an intermediate performance group, showing acceptable but not optimal results, which suggests the presence of operators with stable yet improvable behaviors. In contrast, Cluster 2 groups those operators with the highest scrap levels, reduced quality, and lower attendance, clearly identifying them as the critical segment requiring closer supervision and targeted corrective actions.

This segmentation provides a structured view of performance variability within the workforce and facilitates the identification of operational patterns that may not be evident through traditional evaluation methods. The detailed values associated with each cluster are presented in Table 2.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

Box 5

Table 2
Cluster Summary

Cluster	Scrap (%)	Time (s)	Quality (%)	Attendance (%)
0	7.28	12.20	90.00	91.36
1	2.65	14.03	94.91	89.57
2	6.88	17.55	90.40	90.62

Source: own elaboration

Box 6

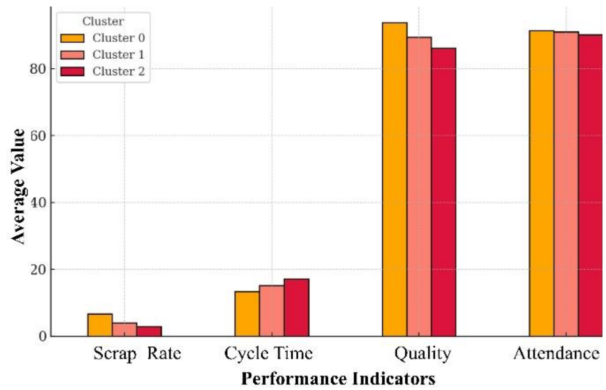


Figure 4
Comparison of average KPIs by cluster (K-Means)
Source: Own Elaboration

Consistent with the findings reported by [Kang et al., 2020], the implementation of machine-learning algorithms in production lines not only enables the automation of complex processes but also facilitates the discovery of meaningful patterns within operational data.

The three clusters identified by the K-Means algorithm—representing high, medium, and low performance profiles—allow organizations to adopt differentiated strategies for training, supervision, and personnel recognition. This aligns with the observations of [Tobar-Díaz et al., 2023], who highlight the usefulness of machine-learning models in multivariate environments. Likewise, the digitalization strategy presented in this study responds to the recommendations of [Díaz-Ramírez, 2021], who emphasizes that the convergence of machine learning, visualization, and automation will serve as a foundational pillar of future industrial systems. This technological integration also promotes meritocracy and transparency in personnel management, aligning with the continuous-improvement principles of Lean Manufacturing described by [Chávez, 2022].

The results demonstrate the value of applying artificial intelligence to performance evaluation in industrial environments. The strong agreement between the Random Forest model and historical supervisory assessments, combined with the clustering insights provided by K-Means, shows that traditional and subjective evaluation methods can be replaced with more objective, consistent, and continuous digital tools.

However, previous comparative studies show that machine-learning performance in industrial settings can vary considerably. [Kharitonov et al., 2022] evaluated ten algorithms for anomaly detection in manufacturing and reported strong differences in accuracy depending on data characteristics and noise levels. While models such as AutoEncoder and Feature Bagging showed limited applicability, KNN unexpectedly performed best. In contrast, the Random Forest and K-Means combination in this study delivered stable and interpretable results, suggesting that the proposed framework is well-suited to heterogeneous operator-performance data.

Conclusions

The system developed represents an initial step toward the intelligent automation of operational performance evaluation in industrial environments. Although this phase was validated only as a functional prototype, the results demonstrate its technical feasibility and its potential to replace traditional methods based on subjective assessments. The proposed architecture enables the solution to scale toward real-world implementation, with real-time integration of artificial intelligence models, continuous data processing, and automatic updating of key performance indicators.

Furthermore, the combination of supervised and unsupervised models proved effective for classifying and segmenting operators based on objective evidence, opening the possibility of applying differentiated strategies for training, task assignment, and continuous improvement. The integration of interactive dashboards strengthens the transparency of the process and facilitates the interpretation of results by supervisors and talent-management areas.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

This work lays the foundation for future research aimed at incorporating new predictive models, expanding the number of indicators analyzed, integrating IoT sensors or MES systems, and evaluating the system's impact on real productivity once deployed in an operational environment. The findings confirm that the adoption of intelligent tools is not only possible but necessary to enhance industrial competitiveness within the framework of digital transformation.

Declarations

Conflict of interest

The authors declare no interest conflict. They have no known competing financial interests or personal relationships that could have appeared to influence the article reported in this article.

Author contribution

Vázquez de los Santos, Laura Cristina: Supervised the methodological design and scientific rigor of the study, validated the theoretical framework, and conducted the critical review.

Burciaga Alarcón, Ricardo: Led the technical development of the system, including the conceptualization, architectural design, implementation of machine-learning models, and preparation of the initial manuscript draft.

Rodríguez Silva, Jesús Rolando: Contributed to data collection and preparation, KPI structuring, and preliminary performance analysis.

Rodríguez Arzola, Adrián: Supported backend and database integration, performed technical validation of the prototype, and assisted in assessing its operational feasibility.

Availability of data and materials

All data used for this research were derived from our own data analysis, no information from third parties was used.

Funding

The research did not receive any funding.

Acknowledgements

We thank the Autonomous University of Coahuila for the facilities provided to carry out this research.

Abbreviations

ANN	Artificial Neural Network
KPI	Key Performance Indicator
RF	Random Forest
s	seconds
%	percentage

References

Antecedents

Algawatta, N., & Jayasekara, R. (2025). Optimizing sewing line performance: The impact of layout on productivity in the garment industry. *Journal of Managerial Sciences*, 19(3), 43–44 (y ss.).

International Labour Office. (1998). Improving working conditions and productivity in the garment industry (Instructional material). *International Labour Organization*.

Rahman, A. et al. (2024). A comparative evaluation of line balancing methods to enhance garment sewing line productivity. *In Proceedings of the IEOM Society Conference*.

Soto Díaz, C., Tarango Hernández, L. E., Zorrilla Briones, F., Terrazas Mata, L. E., & Alvarado Tarango, L. (2024). Elementos de apego laboral en personal operativo en una empresa maquiladora. *Revista NEYART*, 2(3).

Valdez, Y. M., & Salcido, L. F. E. (2021). El control interno como factor determinante de la eficacia en los empleados de empresa maquiladora. *Revista de Investigación Académica sin Frontera*, 36.

Vega, Y., Romero López, R., Rosel Solís, M. J., Chávez Moreno, E. A., & Morales Contreras, O. A. (2022). Indicadores críticos en el desempeño de actividades de cambios rápidos en la industria maquiladora de Baja California. *Revista de Ciencias Tecnológicas (RECIT)*, 5(1), 165–175.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902

Basics

Guzmán-Anaya, L. (2019). Knowledge transfer in the automotive industry: The case of JICA's project for automotive supply chain development in Mexico. *Mirada con Ciencia y Praxis*, 8(23).

Sandoval, L. J. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *REDICCES*. Repositorio Digital de la Ciencia y Cultura de El Salvador.

Supports

Al-Abadi, A. A., Mohamed, M. A., & Fakhfakh, A. (2023). Enhanced Random Forest classifier with K-Means clustering (ERF-KMC) for secure IoMT networks. *Computers*, 12(12), 262.

Díaz Ramos, M. A., Acosta Soberano, M. de los Á., & Pérez Cervantes, G. (Eds.). (n.d.). Proyecciones en innovación empresarial. *Grupo de Ediciones y Publicaciones Xalapa S.A. de C.V.*

Hassan, S. A., Beliatis, M. J., Radziwon, A., Menciassi, A., & Oddo, C. M. (2024). Textile Fabric Defect Detection Using Enhanced Deep Convolutional Neural Network with Safe Human–Robot Collaborative Interaction. *Electronics*, 13(21), 4314.

Silva-Castro, D., Campoblanco-Carhuachin, R., Leon-Chavarri, C., Cardenas-Medina, J., & Borges-Ribeiro, R. (2025). A production management method to reduce non-fulfillment of orders based on lean tools and change management: Case of a Peruvian apparel company. *International Journal of Engineering Trends and Technology*, 73(4), 191–209.

Differences

Kharitonov, A., Nahhas, A., Pohl, M., & Turowski, K. (2022). Comparative analysis of machine learning models for anomaly detection in manufacturing. *Procedia Computer Science*, 200, 1288–1297.

Discussions

Chávez Pineda, J. A. (2022). Adopción parcial e integral de las prácticas del sistema técnico de Lean en la industria maquiladora de manufactura en México. *RECAI*. Revista de Estudios de Contaduría, Administración e Informática, 11(30).

Díaz-Ramírez, J. (2021). Aprendizaje automático y aprendizaje profundo. *Ingeniare*. Revista Chilena de Ingeniería, 29(2), 180–181.

Kang, Z., Catal, C., & Tekinerdogan, B. (2020). Machine learning applications in production lines: A systematic literature review. *Computers & Industrial Engineering*, 149, 106773.

Tobar-Díaz, R., Gao, Y., Mas, J. F., & Cambrón-Sandoval, V. H. (2023). Clasificación de uso y cobertura del suelo a través de algoritmos de aprendizaje automático: *Revisión bibliográfica. Revista Teledetección*, 62, 1–19.

ISSN: 2444-4987
RENIECYT-SECIHTI: 1702902