ISBN: 978-628-96613-1-6. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2025.1.1.1960

Enhancing Industrial Efficiency: AI-Powered Data Analysis and Visualization with Tkinter

Javier Francisco Hall Sevilla[1]; Alberto Max Carrasco[2]; Héctor Villatoro Flores[3]
[1, 3] Universidad Tecnológica Centroamericana, UNITEC, *kiriallfibel@unitecl.edu, hectorvillatoro6@gmail.com*
[2]School of Advanced Technology, Algonquin College, Canada, *carrasa@algonquincollege.com*

Abstract– This research project developed a system to enhance efficiency in industrial data processing and analysis, addressing the growing demand for improved productivity and evidence-based decision-making. By applying data analysis and visualization techniques, the system transformed CSV-based production reports into clear and interactive visual formats. Built using a Python-based framework, it integrates libraries such as Tkinter, Pandas, smtplib, Scikit-learn, and Matplotlib to support data management, graphical representation, and user interaction. The system also includes email functionality, enabling the automated distribution of analysis reports in image format, thereby supporting collaborative review and communication. The system aligns with IoT and Industry 4.0 principles by automating the handling and sharing of machine-generated data. This reduces manual processing time, facilitates informed decision-making, and makes analytical insights accessible even to non-experts.

Keywords-- Data Analysis, Machine-Learning, Pandas, Matplotlib, Tkinter.

I. Introduction

This study embarks on a pivotal exploration aimed at revolutionizing the way industrial data is managed and analyzed, particularly focusing on the transformation of information shown in CSV data sets into coherent and intelligible visual formats. Utilizing cutting-edge programming libraries for data manipulation, alongside matplotlib for visualization and tk (Toolkit) for interface development, also adding a machine-learning module to make the analysis more flexible, this work is poised at the forefront of industrial innovation. The urgency and relevance of this endeavor are underscored by the contemporary industrial landscape's pressing demand for enhanced data comprehension and optimization of production workflows.

In the heart of this project lies the ambition to forge a tool that transcends mere data analysis program; it seeks to instill a culture of continuous improvement and operational efficiency within the industrial sector. The envisioned system is not only designed to provide exhaustive analytical insights and amalgamated multifaceted reports but also to streamline the communication of these findings through an accessible and user-friendly interface, by using data from an industrial machine that gives reports or industrial processes in the form of CSV documents having data as defects and total production units.

By setting its sights on creating an indispensable asset for industrial data management, this thesis aspires to contribute significantly to the strategic decision-making processes within industry. It is a testament to the project's commitment to elevating data management practices, ensuring that the resultant application emerges as a cornerstone of industrial efficiency and a catalyst for transformative change in the realm of data-driven decision-making.

II. Background and Technical Challenges

In the area of data analysis, growth has been exponential in search of achieving more efficient systems, the need to visualize data is observed as something fundamental for the industrial field [1]. Intelligent machines that provide data are an increasingly present reality, delivering information in large quantities, to acquire this information, different ways are used, such as collecting data in compressed documents, which can be used in a wide variety of programs and manipulated as necessary [2].

In addition to supervising operations, manufacturing companies face the crucial task of controlling processes, a task that places the observation of machinery in a highly relevant position. These systems become critical and indispensable, especially when the quality of the products changes based on the importance or unique value of each piece [3]. Industrial machines, by delivering more production data, allow exponential improvement if the data obtained can be used well to make strategies based on it.

The correlation of existing methods for the identification of resources and necessary data is fundamental in the sequence of any processing instance, given that in said instance the available information may be restricted, this gives rise to what is known as problems in the distribution of resources where you do not have a complete view of the situation [4].

Knowing that the data indicators that can be obtained by industrial machines are delivered in compressed documents, manipulating them or being able to view their information takes a considerable amount of time and in many cases, it is necessary to have the information quickly to avoid production errors [5].

In the area of computing, something essential is to represent information in a way that is easy to understand even for people who do not know how to program or be able to interact with a terminal, so a user interface allows the manipulation of an analysis program much more easy overcoming technical mastery challenges [6].

GUIs enable non-experts to interact with complex analytical tools through intuitive interfaces, making data-driven insights more accessible across production roles.

Depending on the data that is being processed, we can delve into the various critical concepts integral to the project, from the graphical interface and document management to data visualization in graphs and functions like merging multiple documents into one.

This allows for the sharing of information generated by the program via email. Currently, data visualization plays a vital role in developing evaluation metrics for strategies and time series. Key performance indicators, for instance, are essential for quantitatively assessing current situations and the quality of target objects.

However, crafting metrics, particularly for data with dimensional complexities and temporal characteristics, poses challenges even for domain experts. This innovative approach has proven effective through examples and user experiments based on hypotheses derived from these examples, emphasizing the significance of data visualization in extracting meaningful metrics from complex data sets [7].

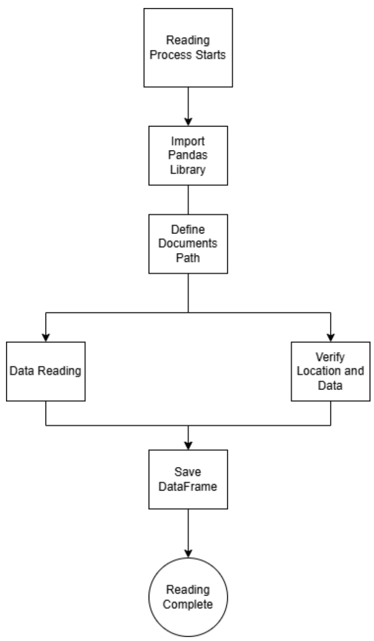

Fig. 1 Flowchart of steps for the reading of Data.

The recent emphasis on optimizing productive efficiency in the industry highlights the importance of a structured methodology that spans from initial data collection to validation, underlining the opportunity for continuous learning and improvement for businesses. This structured methodology is visually summarized in the data processing flowchart shown in Fig. 1.

The proposed interface aims to overcome these challenges by extending the concept of semantic interaction to address the temporal features of data, enabling analysts to manipulate visualized objects in an efficient and user-friendly GUI. On the other hand, data mining, linked to process analysis, has become an indispensable tool for decision-making across various fields, notably in manufacturing.

The wealth of data generated by production lines offers fertile ground for applying data mining techniques, enabling detailed analyses, and promoting greater efficiency across the entire production line [8]. This highlights data mining's value in improving efficiency by leveraging the extensive information gathered across production lines.

However, a significant challenge in this area is understanding the information provided by production machines. This process, which includes data acquisition, comprehension, and preparation, can account for over 60% of the total time dedicated to data analysis according to recent research [9]. As data collection is crucial, it becomes a time-consuming and resource-intensive task, complicating the effective interpretation of information.

A. *Data Analysis Interface*

In the realm of computing, data analysis emerges as a critical component, particularly within the context of data mining, which stems from the necessity to visualize information based on processes. Decision-making across various domains heavily relies on production data, underscoring the importance of analyzing data, especially through time series, to tackle numerous challenges within these fields [10].

Various methodologies exist for data collection, yet Graphical User Interfaces (GUIs) stand out for their unique ability to assess the current state of gathered data. This approach enables a more comprehensive and effective perspective, which is crucial for competently addressing and resolving challenges in the productive environment [11].

Modern infrastructure is characterized by its complexity, founded on the integration of cyber-physical systems. This synergy between material processes and the intercommunication of technological elements facilitates data collection that enriches the understanding of the operational environment [12].

The convergence of data mining with physical systems in contemporary computing signifies a substantial leap in productive efficiency and understanding, which are key to making informed decisions based on data.

B. *Data Management and Analysis*

Recently, there has been a significant focus on enhancing productive efficiency within the industrial sector, characterized by the necessity to implement a series of sophisticated controls for a variety of processes. This focus has generated a demand for analyzing extensive volumes of data to meet the requirements of minimizing and managing the energy consumption of machinery [13].

Improving productive efficiency is achieved through optimization and meticulous management of processes, thereby providing a learning opportunity for organizations focused on specific goals. An instance of this is research aimed at reducing the environmental impact through data analysis of industrial equipment, focusing on non-ferrous metals [14].

It's essential to segment the process into distinct phases for proper data interpretation, starting from the initial stage of data collection and description, followed by the selection and extraction of relevant data, and concluding with the validation of the data analysis process [15].

C. Data Visualization with Graphs

Assessments indicate that the manufacturing sectors provide an extensive field for the application of data mining due to the vast amount of information generated along production lines.

This facilitates the performance of detailed analyses for each production sequence, expanding the possibilities of gaining valuable insights [16].

There are various autonomous methodologies for data handling, including OLAP (Online Analytical Processing), which segments procedures into distinct stages, fitting a wide range of uses. Conventionally, OLAP consolidates several data processing systems under a single framework [17].

It has been observed that efficiency in collecting data from production lines can significantly improve by applying temporal models to analyze production speeds and error frequency on a specific line, utilizing information collected at different times from the same machine [18]. The manufacturing domain emerges as a primary area for data mining, where methods like OLAP analysis and temporal models enrich the process's understanding and efficiency through detailed data analysis collected along production lines.

D. Data Processing and Unification

In the manufacturing sphere, beyond routine operations, meticulous monitoring of systems and machinery is established as a priority, crucial for ensuring consistency in product quality.

Tracking data across dates ensures consistent product quality and highlights each component's role in the process [3]. In the industrial context, database configuration is based on data directly from machinery, presenting challenges in both its implementation and management. Transforming these data into more accessible formats is essential for accurately replicating real operational conditions, facilitating their analysis, and understanding [14].

Coordinating existing methods to identify essential resources and information becomes a sequential requirement in any process stage, facing the challenge of having restricted information at certain times, known as the difficulty in resource allocation, where foresight is limited [15]. The intersection of operations, precise monitoring, and database management in the manufacturing industry underscores the need to continuously adapt and refine processes to face the challenges inherent in handling complex and diverse information, ensuring efficiency and effectiveness in all production cycle phases.

E. Reports Automation and Results Communication

Real-time reporting was enabled through email functionality, allowing users to send analysis graphs as PNG files via a preconfigured email account. The system uses SMTPLib protocols for internet connectivity and authentication, streamlining remote access to production data. Also having the importance of IoT and the way of accessing information without the need for being in the actual place is one functionality that helps for more efficiency and for faster checking of the data of production in an industry. The program to have the ability to connect with the internet had to use many protocols that are mainly rooted in the SMTPLib library. From creating a google account to giving its credentials to the program so it can access and send messages through that account.

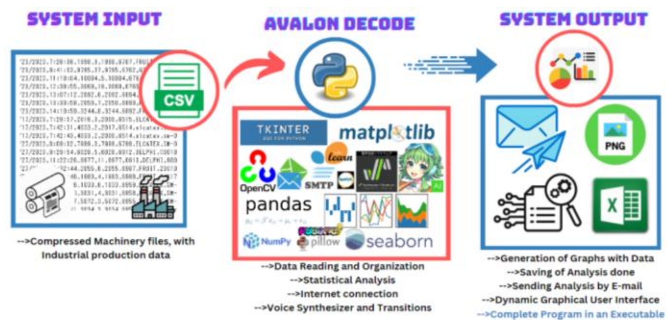

Fig. 2 Conceptual Diagram

The conceptual diagram in Fig. 2 illustrates the system's data analysis workflow from input to output. First, the system receives compressed CSV files containing industrial production data. Then, in the Avalon Decode phase, the data is processed using Python and various libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Tkinter, enabling data organization, statistical analysis, and visualization.

It also incorporates internet connectivity, voice synthesis, and email functions for sharing reports. Finally, in the output phase, the system generates PNG graphs, allows report storage, sends analysis via email, and provides an interactive graphical interface. All these features are integrated into an executable program, making it user-friendly without requiring advanced programming knowledge [16]. Industrial production systems face challenges due to higher quality standards and the integration of numerous sensors for data collection.

III. Limitations and Solutions

The system resolved multiple initial limitations through targeted improvements. Manual statistical processing was automated using mathematical libraries, reducing errors and improving efficiency. A simplified graphical interface was introduced to make the tool accessible for non-technical users, while email integration enabled efficient report sharing.

Navigational complexity was addressed through a button-based system, and the dependence on internet connectivity was minimized by adding an offline mode. The ability to export graphical results as PNG files improved data accessibility and archiving.

These enhancements significantly improved usability and data management. The rise in industrial monitoring has increased data volume, emphasizing the need for efficient automated communication processes [17]. Automatic report generation and result sharing via graphs and email improve clarity and speed in data-driven decisions [18]. Industrial IoT enables remote access to machines and analytics, increasing operational flexibility [19], while the importance of accurate data collection and management remains central to effective decision-making [20].

IV. Methodology

This chapter focuses on and describes in detail the methodology implemented to develop and execute the data analysis application designed with the objective of processing, analyzing, and viewing CSV reports related to various operational areas. The methodology applied for the development of this tool ranges from the design phase of the graphical user interface (GUI) to the implementation of advanced data analysis functionalities, including data visualization through graphs, comparison of reports, and the unification of multiple reports in a single document. In addition, it has functionality for sending emails that allows you to share the analysis results efficiently.

A. Techniques and Instruments Applied.

The program efficiently integrates the graphical interface that is created through Tkinter with the data analysis capabilities of the Pandas and NumPy libraries, allowing effective manipulation and analysis of the data sets.

Through the Matplotlib library, the application facilitates the graphical representation of the analyses, offering clear and understandable graphs that summarize the results obtained. In addition, communication functionality is incorporated through the SMTPLib protocol, allowing analysis results to be shared directly by email. This synergy of tools not only promotes deep and detailed data analysis but also provides a seamless and complete user experience, from data import to visualization and sharing. The machine-learning module used was scikit-learn which granted the ability to give more feedback based on the data of each file that was being analyzed, from single file analysis to a whole repertoire analysis based on the data being recollected by the machine.

B. Study Methodology

The following section addresses the steps to develop the program of data analysis for csv documents. The methodology used to develop the program was based on an iterative approach, inspired by the incremental model.

This process began with the precise definition of requirements, followed by the design of the user interface, the implementation of analysis and visualization capabilities, culminating in a phase of testing and adjustments based on user feedback, thus ensuring that the application not only satisfied the initial needs, but also adapted to future requirements and new functionalities suggested by users, maintaining the relevance and effectiveness of the tool over time.

The process begins with the collection of structured data in CSV format, which is imported and cleaned by the program to eliminate unwanted entries. Subsequently, they are processed using Pandas for tasks such as time conversion and categorical variable coding, unifying multiple reports for comparative analysis and organizing them in DataFrames.

Matplotlib is used to generate visualizations that facilitate the identification of trends and patterns. Finally, the program allows the generation and export of reports and graphs, with options to send them by email, thus optimizing the distribution and access to the analyzed results.

C. Data Processing

Once the data is imported, the processing proceeds using the Pandas library. This process may include converting time formats to seconds, as well as entering missing point values, encoding categorical variables, and aligning DataFrame columns with those required for subsequent analysis. The tool also allows the unification of multiple reports into a single data set for comparative analysis. The data that will be worked with will be production data, production totals, defects, time, and clients who ordered this data, organizing them in DataFrames and then visualizing them graphically.

D. Data Visualization and Analysis

With the data already processed, the capabilities of the Matplotlib library integrated into the program are used to generate visualizations such as bar and line graphs. These visualizations serve both for initial exploration of the data and for identifying trends and patterns. Among the generated graphs, the ease of visualizing the required data is taken into account, so different graphs are used at different points, such as to compare defects with totals of single reports, a foot graph is used, to visualize time with totals and defects, line graphs are used, both histogram graphs and density graphs are used for reports that have more information, as are bar graphs.

E. Evaluation Metrics

For evaluation metrics, the base tool is the surveys, both surveys for experts and the survey for general users, each of these tools provides a more extended view of how the program is received and its impact and usefulness in different user perspectives according to as displayed according to each of the groups that you want to have as validation metric points. The expert surveys will themselves be experts in data analysis, it is designed to collect detailed and specialized feedback. Professionals will evaluate the accuracy and efficiency with which the program handles large data sets, including compressed documents that are common in production environments. They will be asked to assess the effectiveness of data import functions, analytical capabilities of the system, and accuracy of visualizations with the help of integration of libraries such as pandas and matplotlib. This critical feedback is vital to improve the usability of the system for complex analysis and handling large volumes of information.

The following validation tool is focused on the general public in order to obtain the answer of how it is received by the non-expert public and not accustomed to data analysis systems, the point is to see if the graphical interface is intuitive enough so that even non-experts can understand the program and navigate through it without any inconvenience, so the survey for general users is focused on the general audience, who may not be familiar with technical data analysis, but who works with reading production documents, the survey seeks to evaluate the overall usability of the program.

It will focus on the user experience when interacting with the interface, the clarity of the instructions and its intuitiveness, from data loading to report generation. The simplicity of data visualization and the effectiveness of communicating results will be key points in this survey. Feedback on the user experience and GUI will be collected through a survey that asks users to rate the overall experience, aesthetics, and navigation within the app. They will be asked about initial setup, understanding graphs, and ease of access to features such as comparing data and sending reports by email. The results of this survey will be instrumental in making iterative improvements to the interface design, ensuring that the application is not only powerful in terms of functionality, but also accessible and enjoyable to use for all users.

V. RESULTS

In this section, the results of the data analysis show the progressive development of the software through eight increments, highlighting graphical reading and display of CSV files, exporting, and emailing of graphics, and an intuitive user interface that improves with each phase.

The software enhancements included sophisticated capabilities. Furthermore, the user interface was continually improved, becoming more intuitive and user-friendly with each phase. The incremental approach allowed for step-by-step integration and optimization of new features, ensuring a robust and comprehensive data analysis tool by the end of the development process.

A. Results Obtained

The results follow an incremental methodology implemented in eight progressive development stages. Each stage introduced foundational and specific functionalities, enabling gradual system enhancement and user-focused refinement. During each increase there is a base point, which must be carried out progressively. More points are added for its total functionality. Initially, a solid foundation for data handling was established by implementing functions for loading and processing CSV files, setting the stage for more complex analyses (see Fig. 1). Visualization tools like Matplotlib significantly enhanced data presentation and interpretation, while features such as email reporting and the ability to save graphs as PNG files increased interactivity and utility. Advanced algorithms for anomaly detection and comparative analysis enriched the data analysis, providing broader insights. Continuous user feedback through surveys indicated high satisfaction with the software's clarity, navigation, and overall functionality, culminating in a versatile and user-friendly program.

B. First Increment

The first increment focused on setting up the environment and importing necessary libraries to process CSV files. A file selection system ensured valid file formats, while libraries like pandas and numpy enabled basic data handling. This initial step ensured that the data was properly organized for later analysis and visualization.

A file selection system has been implemented that checks the compatibility of data with program requirements before uploading. In addition, important libraries such as pandas and numpy were incorporated for data management, and tkinter for the user interface.

Aesthetic and interactivity elements were added to the interface through libraries such as cv2 and pygame, which allow the inclusion of multimedia files such as videos and sounds, enriching the user experience and facilitating navigation.

The implementation of a button to insert CSV documents, using libraries such as tkinter and pandas, facilitated data loading and established the basis of the system. This paved the way for advanced analysis and visualization of the uploaded data.

C. Second Increment

This stage introduced data visualization and initial analysis by category. The system verified and displayed uploaded CSV files within the interface, allowing dynamic projections without requiring separate pop-up windows.

D. Third Increment

This increment added comparative analysis by enabling side-by-side visualization of two different CSV files, directly integrated into the GUI to improve navigation and interpretation.

E. Fourth Increment

This phase includes data preprocessing and model adjustments, enabling the creation of multiple graph screens with different buttons. This enhancement optimizes dynamic data visualization without distorting the graphical interface, ensuring smooth data display.

F. Fifth Increment

This phase focuses on implementing graph generation using CSV data and saving these graphs as PNG files. This allows users to export and save analysis results with a button that generates and stores the graph, enhancing flexibility and interaction with the data analysis.

G. Sixth Increment

This phase emphasizes communication and reporting, enabling internet connectivity for sending analyzed graphs as PNG files via email. A button triggers a pop-up for entering the recipient's email and attaching the PNG file, facilitating direct report sending through a pre-configured email account.

H. Seventh Increment

This phase improves user interface and experience, adding flexibility to existing functions. Enhancements include a welcome sound using pygame, voice synthesis for notifications via Synthesizer V, and a main menu with access to all functions. Instruction screens explain the program's functionalities upon startup.

I. Last Increment

In the final phase of the development of the data analysis software, dynamic analysis functionality was added, allowing users to select specific machines and analyze their data through CSV files using a dedicated menu. This feature enables detailed analysis of multiple aspects based on the files available in each machine's repertoire, including the option to save the results as PNG images and easily navigate between the different sections of the program.

Additionally, improvements were introduced to increase interactivity and user experience, such as sound effects when interacting with program elements and the implementation of more complex analysis, such as linear regression and anomaly detection. The beginning of the program was enriched with a welcome sequence and a function tutorial, using video files for a more dynamic introduction, thanks to the use of the cv2 library.

The final touch in improving accessibility and usability was the creation of an independent executable for the program, as shown in Fig. 3. This transformation allows users to run the software without the need to interact with the source code or development environments, making the program more user-friendly and accessible to a broader audience.

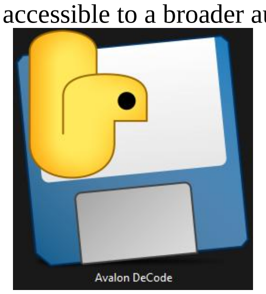

Fig 3. Result of the Executable of the program for the last increment.

J. Comparison of Results

The comparison of results is based on evaluation metrics obtained from surveys of general and expert users, focusing on usability, clarity of information and general satisfaction. The final phase of development was aimed at improving accessibility, transforming the script into an independent executable, which facilitated its use by avoiding the need to handle code directly. This incremental approach not only optimized existing functionalities but also improved the overall user experience, making it more accessible and satisfying for a wide range of users.

The developed system was evaluated based on implementation time, usability feedback, and deployment scalability. Pilot testing was conducted in two real industrial scenarios with flexographic machines. Feedback from operators highlighted ease of use and significant time savings in data processing. Two pilot tests conducted in real industrial environments with flexographic printing machines highlighted ease of use and significant time savings in data processing. The system was implemented in under two weeks, including initial data setup and basic operator training. Evaluation metrics focused on time reduction in data handling, error minimization during manual reporting, and ease of use in daily operations. On average, the system reduced manual processing time by approximately 40%, allowing operators to focus on quality control and process optimization. These outcomes highlight the platform's potential for rapid adoption in small to medium-sized enterprises seeking affordable, high-impact digital tools.

Compared to existing commercial tools, this system provided comparable visualization accuracy with lower cost and greater customization flexibility, making it ideal for small to medium-sized enterprises.

The use of the machine-learning module also granted more flexibility in the analysis that was being given for each machine operator. The usage of the machine-learning module and the IoT functionality makes the exportation of the analysis results visible for the operator even by not being besides the machine makes a boost in efficiency as well as better usage of time by not doing it manually and having to make the extraction from the machine and the cleaning afterwards.

These findings demonstrate the system's practical utility in industrial settings where speed, clarity, and non-expert usability are critical. The integration of email reporting and graphical summaries enables faster managerial response, contributing to leaner decision-making processes. Unlike many commercial data solutions, this system offers a fully customizable platform specifically designed for industrial environments with minimal setup costs and no licensing requirements. Its independence from cloud-based services or proprietary platforms ensures full data control, making it particularly valuable for small enterprises with limited budgets and strict data security requirements. These features—combined with the integration of AI-powered analysis, offline capabilities, and instant reporting—position this solution as a scalable and efficient alternative to generic software, tailored to the operational realities of local industry.

V. Conclusions

This research presents a comprehensive and effective solution for transforming large volumes of industrial machine data—typically stored in CSV format—into clear, actionable, and visually intuitive insights. The developed system not only streamlines the reading and processing of complex datasets but also enables real-time interpretation through graphical outputs, significantly enhancing operational decision-making in industrial contexts.

By integrating widely used Python libraries such as Pandas for data manipulation, Matplotlib for data visualization, and Tkinter for the development of graphical interfaces, the platform consolidates multiple advanced functionalities within a single, accessible environment. This convergence of tools introduces a new benchmark in industrial data analysis, where automation, usability, and scalability coexist.

Unlike conventional tools that often require technical expertise or costly licensing models, this solution is designed for accessibility. It empowers users—regardless of their programming background—to interact with complex data via an intuitive interface. Features such as graph generation, anomaly detection, linear regression, and email-based reporting enable broader adoption, particularly in small to medium-sized enterprises seeking digital transformation.

The project followed an incremental development methodology, validated through two real-world pilot implementations using flexographic printing machines. These deployments demonstrated that the system could be successfully adopted in under two weeks, delivering measurable gains including a 40% reduction in manual data processing time. Operators reported improved efficiency, data comprehension, and responsiveness, all without the need for manual extraction or post-cleaning procedures.

Incorporating Internet of Things (IoT) elements and basic machine learning capabilities further strengthened the platform's utility, enabling remote access to analytical outputs and supporting faster, data-driven decisions. These features position the system not only as a technical innovation, but as a practical, scalable tool that contributes to the continuous improvement of production processes and the democratization of data use in industry.

.

References

[1] F. Wang, M. Li, Y. Mei, y W. Li, «Time Series Data Mining: A Case Study With Big Data Analytics Approach», IEEE Access, vol. 8, pp. 14322-14328, 2020, doi: 10.1109/ACCESS.2020.2966553.

[2] A. A. Syed, F. L. Gaol, y T. Matsuo, «A Survey of the State-of-the-Art Models in Neural Abstractive Text Summarization», IEEE Access, vol. 9, pp. 13248-13265, 2021, doi: 10.1109/ACCESS.2021.3052783.

[3] S. K. Sen, G. C. Karmakar, y S. Pang, «Critical Data Detection for Dynamically Adjustable Product Quality in IIoT-Enabled Manufacturing», IEEE Access, vol. 11, pp. 49464-49480, 2023, doi: 10.1109/ACCESS.2023.3276942.

[4] G. Park y M. Song, «Optimizing Resource Allocation Based on Predictive Process Monitoring», IEEE Access, vol. 11, pp. 38309-38323, 2023, doi: 10.1109/ACCESS.2023.3267538.

[5] D. Wu, H. Wang, H. Mohammed, y R. Seidu, «Quality Risk Analysis for Sustainable Smart Water Supply Using Data Perception», IEEE Trans. Sustain. Comput., vol. 5, n.o 3, pp. 377-388, jul. 2020, doi: 10.1109/TSUSC.2019.2929953.

[6] R. Takami, H. Shibata, y Y. Takama, «A Visual Analytics Interface for Formulating Evaluation Metrics of Multi-Dimensional Time-Series Data», IEEE Access, vol. 9, pp. 102783-102800, 2021, doi: 10.1109/ACCESS.2021.3098621.

[7] J. Rudnitckaia, H. S. Venkatachalam, R. Essmann, T. Hruška, y A. W. Colombo, «Screening Process Mining and Value Stream Techniques on Industrial Manufacturing Processes: Process Modelling and Bottleneck Analysis», IEEE Access, vol. 10, pp. 24203-24214, 2022, doi: 10.1109/ACCESS.2022.3152211.

[8] K. Nagorny, S. Scholze, A. W. Colombo, y J. B. Oliveira, «A DIN Spec 91345 RAMI 4.0 Compliant Data Pipelining Model: An Approach to Support Data Understanding and Data Acquisition in Smart Manufacturing Environments», IEEE Access, vol. 8, pp. 223114-223129, 2020, doi: 10.1109/ACCESS.2020.3045111.

[9] P. Dedousis, G. Stergiopoulos, G. Arampatzis, y D. Gritzalis, «A Security-Aware Framework for Designing Industrial Engineering Processes», IEEE Access, vol. 9, pp. 163065-163085, 2021, doi: 10.1109/ACCESS.2021.3134759.

[10]H. S. Kang, J. Y. Lee, y D. Y. Lee, «An Integrated Energy Data Analytics Approach for Machine Tools», IEEE Access, vol. 8, pp. 56124-56140, 2020, doi: 10.1109/ACCESS.2020.2981696.

[11]S. Xie, H. Wang, y J. Peng, «Energy Efficiency Analysis and Optimization of Industrial Processes Based on a Novel Data Reconciliation», IEEE Access, vol. 9, pp. 47436-47451, 2021, doi: 10.1109/ACCESS.2021.3068374.

[12]S. A. Shaikh y H. Kitagawa, «StreamingCube: Seamless Integration of Stream Processing and OLAP Analysis», IEEE Access, vol. 8, pp. 104632-104649, 2020, doi: 10.1109/ACCESS.2020.2999572.

[13]F. Deng et al., «A Multi-Sensor Data Fusion System for Laser Welding Process Monitoring», IEEE Access, vol. 8, pp. 147349-147357, 2020, doi: 10.1109/ACCESS.2020.3015529.

[14]C. Bernardis, M. F. Dacrema, F. B. P. Maurera, M. Quadrana, M. Scriminaci, y P. Cremonesi, «From Data Analysis to Intent-Based Recommendation: An Industrial Case Study in the Video Domain», IEEE Access, vol. 10, pp. 14779-14796, 2022, doi: 10.1109/ACCESS.2022.3148434.

[15]Y. Park y Y. Shin, «Tooee: A Novel Scratch Extension for K-12 Big Data and Artificial Intelligence Education Using Text-Based Visual Blocks», IEEE Access, vol. 9, pp. 149630-149646, 2021, doi: 10.1109/ACCESS.2021.3125060.

[16]S. Park y J.-H. Huh, «A Study on Big Data Collecting and Utilizing Smart Factory Based Grid Networking Big Data Using Apache Kafka», IEEE Access, vol. 11, pp. 96131-96142, 2023, doi: 10.1109/ACCESS.2023.3305586.

[17]J. Yu y X. Yan, «Modeling Large-Scale Industrial Processes by Multiple Deep Belief Networks With Lower-Pressure and Higher-Precision for Status Monitoring», IEEE Access, vol. 8, pp. 20439-20448, 2020, doi: 10.1109/ACCESS.2020.2968033.

[18]K. L.-M. Ang y J. K. P. Seng, «Big Data and Machine Learning With Hyperspectral Information in Agriculture», IEEE Access, vol. 9, pp. 36699-36718, 2021, doi: 10.1109/ACCESS.2021.3051196.

[19]A. Goldstein, L. Fink, y G. Ravid, «A Cloud-Based Framework for Agricultural Data Integration: A Top-Down-Bottom-Up Approach», IEEE Access, vol. 10, pp. 88527-88537, 2022, doi: 10.1109/ACCESS.2022.3198099.

[20]S. Salloum, T. Gaber, S. Vadera, y K. Shaalan, «A Systematic Literature Review on Phishing Email Detection Using Natural Language Processing Techniques», IEEE Access, vol. 10, pp. 65703-65727, 2022, doi: 10.1109/ACCESS.2022.3183083.

[A. Garcia-Robledo y M. Zangiabady, «Dash Sylvereye: A Python Library for Dashboard-Driven Visualization of Large Street Networks», IEEE Access, vol. 11, pp. 121142-121161, 2023, doi: 10.1109/ACCESS.2023.3327008.