Journal of Information System and Technology Research

journal homepage: https://journal.aira.or.id/index.php/jistr/

An Intelligence-Oriented System Architecture for Integrated Pharmaceutical Data Analytics and Decision Support

Ningsiah[1,*], Nur Aminudin[2], Septika Ariyanti[3] Ramil Abbasov[4]

[1]Pharmacy, Aisyah University of Pringsewu, Indonesia
[2]Software Engineering, Aisyah University of Pringsewu, Indonesia
[3]English Language Education, Aisyah University of Pringsewu, Indonesia
[4]George Mason University, USA

ARTICLE INFO

Article history:

Received Dec 22, 2025
Accepted Jan 25, 2026
Available online Jan 31, 2026

Keywords:

Hybrid Information System
Integrated Analytics
Decision Support System
Pharmaceutical Data Intelligence
Machine Learning Integration

ABSTRACT

This study proposes and evaluates an intelligence-oriented hybrid information system architecture for pharmaceutical data analytics and decision support. Unlike conventional approaches that treat analytics as an external component, the proposed framework embeds analytical intelligence directly into the core system architecture through an integrated, multi-layer design. The study adopts an experimental and system development methodology using a large-scale public pharmaceutical dataset consisting of 240,591 records and 10 attributes. Supervised machine learning models are implemented to support data classification and intelligence generation, and system performance is evaluated using accuracy, precision, recall, and F1-score. The results demonstrate that the proposed hybrid system consistently outperforms baseline and non-integrated approaches, achieving higher predictive stability and analytical consistency. The main contribution of this study lies in its system-level integration model, which enables the transformation of raw pharmaceutical data into actionable decision-support intelligence. The findings confirm that embedding analytics within information system architecture significantly enhances both analytical performance and decision-making capability in pharmaceutical information systems.

© 2026 The Author(s). Published by AIRA.
This is an open access article under the CC BY-SA license
(http://creativecommons.org/licenses/by-sa/4.0/).

Corresponding Author:

Ningsiah,
Pharmacy, Aisyah University of Pringsewu,
Jl. A Yani No. 1A Tambahrejo, Gadingrejo District, Pringsewu Regency, Lampung.
Email: ningsiah@aisyahuniversity.ac.id

1. INTRODUCTION

The rapid advancement of information technology has fundamentally transformed the role of information systems in the pharmaceutical sector, shifting them from operational data-recording tools to strategic platforms for analytics-driven decision making [1], [2]. Contemporary pharmaceutical environments generate heterogeneous data streams from clinical, operational, and public data sources, which require systematic integration to support accurate, timely, and evidence-based managerial and regulatory decisions [3], [4], [5], [6]. However, the increasing volume and structural complexity of such data often hinder organizations from fully exploiting their informational and strategic value [7], [8].

Previous studies on pharmaceutical information systems and data analytics have primarily focused on domain-specific data management, predictive modeling, and isolated machine learning applications for drug classification, demand forecasting, and quality monitoring [9], [10], [11]. Hybrid system frameworks have been proposed to combine data processing pipelines with analytical components; however, most existing approaches emphasize algorithmic performance rather than system-level integration and intelligence generation across organizational layers [12], [13], [14]. As a result, current solutions tend to operate as standalone analytical modules rather than as cohesive, decision-support-oriented information system architectures [15].

Despite these advances, there remains a lack of system-oriented frameworks that explicitly integrate data acquisition, analytical intelligence, and decision-support mechanisms within a unified architectural design for pharmaceutical information systems [9], [16]. Prior research has not sufficiently addressed how embedded analytics and intelligence generation can be operationalized at the system level to bridge the gap between raw data processing and actionable managerial decision making [17]. This absence of a cohesive, intelligence-oriented system architecture constitutes the central research gap addressed in this study.

This study aims to design and evaluate an intelligence-oriented system architecture that integrates pharmaceutical data management, analytical processing, and decision-support functionalities within a unified information system framework [18]. Specifically, the research seeks to (1) develop a system-level integration model that embeds analytical intelligence into core information system components, (2) assess the effectiveness of the proposed architecture in transforming heterogeneous pharmaceutical data into actionable insights, and (3) demonstrate its applicability for supporting operational and managerial decision-making processes [19], [20].

The main contributions of this study are threefold. First, it proposes a novel system architecture that extends conventional hybrid information system models by embedding intelligence generation and decision-support mechanisms at the architectural level rather than as external analytical modules [21], [22], [23]. Second, it provides a methodological framework for integrating heterogeneous pharmaceutical data sources into a cohesive analytical and decision-support pipeline [18]. Third, it offers empirical insights into how intelligence-oriented system design can enhance the practical and theoretical understanding of integrated analytics within pharmaceutical information systems [3], [24].

Although prior studies have explored the application of data analytics and machine learning in pharmaceutical and healthcare contexts, most existing works focus on algorithmic performance or isolated analytical pipelines. Limited attention has been given to how analytical intelligence can be systematically embedded within information system architectures to support continuous decision-making processes. Current approaches often treat analytics as an external or post-processing component, resulting in fragmented system designs and limited operational integration. Therefore, there remains a clear research gap in the development of an intelligence-oriented information system architecture that tightly integrates data management, analytics, and decision support within a unified framework. This study addresses this gap by proposing and empirically evaluating a hybrid system architecture that embeds analytical intelligence at the system level.

2. RESEARCH METHOD

2.1 Research Design

This study adopts an experimental and system development research design to empirically evaluate an intelligence-oriented hybrid information system architecture for pharmaceutical data analytics and decision support. The experimental component is employed to assess the analytical performance and classification effectiveness of the embedded machine learning models, while the system development component is used to design, implement, and integrate data management, analytics, and decision-support modules within a unified, multi-layer system architecture [25], [26].

Methodologically, the study follows a quantitative, data-driven approach grounded in large-scale, real-world pharmaceutical transaction data. The dataset comprises 240,591 records (N = 240,591) with 10 attributes (F = 10), covering 132,400 unique invoices and 21,870 distinct products collected between January 1, 2024, and September 9, 2024. The primary target variable is the product class label (Drug vs. Supply), formulated as a supervised classification task to support intelligent categorization, demand segmentation, and downstream decision-support processes [27], [28].

The selection of machine learning algorithms is explicitly justified by the heterogeneous, mixed-type feature space and the system-level objective of balancing analytical performance with interpretability and operational integration. A Logistic Regression model is employed as a transparent baseline to assess linear separability and feature relevance. A Random Forest ensemble is selected to capture non-linear relationships and hierarchical interactions inherent in transactional and categorical pharmaceutical data, while providing feature importance measures that support system-level intelligence generation. A Support Vector Machine (SVM) is incorporated as a comparative model to evaluate generalization performance and scalability under high-dimensional and partially overlapping class distributions [29].

The research workflow is structured into four tightly coupled stages: (1) Data Acquisition and Integration, where heterogeneous transactional data are ingested, validated, and harmonized; (2) Data Processing and Analytics, which includes preprocessing, feature engineering, and supervised model training and validation using a stratified 70/15/15 training–validation–testing split and 5-fold cross-validation; (3) Intelligence and Decision Support, where model outputs and feature importance metrics are transformed into actionable insights through rule-based reasoning and visualization mechanisms; and (4) Presentation and System Evaluation, which enables user interaction and assesses system performance using quantitative metrics such as accuracy, precision, recall, F1-score, and computational efficiency [30].

Figure 1 presents the research workflow adopted in this study, illustrating the sequential stages from data acquisition to system evaluation

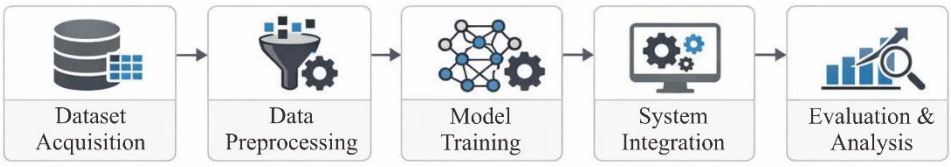

Figure 1. Research workflow of the proposed system
Source: Authors' own illustration, developed based on the proposed hybrid information system architecture and adapted from standard system development and analytics workflows in integrated information systems and CRISP-DM methodologies.

Figure 1 illustrates the research workflow of the proposed hybrid information system, showing the sequential stages from data acquisition to evaluation and analysis. The workflow begins with dataset acquisition, in which a public pharmaceutical dataset is collected from Kaggle to ensure transparency and reproducibility. The next stage is data preprocessing, which involves data cleaning, handling missing values, normalization, and encoding of categorical variables to improve data quality and analytical readiness.

The processed data are then utilized in the model training stage, during which supervised machine learning algorithms are trained to extract patterns and relationships from pharmaceutical data. This stage focuses on developing analytical models that can generate predictive outputs and support data-driven insights. Following model training, the system integration stage embeds the trained analytical models into the hybrid information system architecture, thereby enabling automated analysis and seamless interaction between data management and intelligence components.

The final stage is evaluation and analysis, where the performance of the integrated system is assessed using quantitative evaluation metrics. This stage validates both the analytical accuracy of the models and the effectiveness of the system in generating pharmaceutical data intelligence. Overall, Figure 1 demonstrates how the proposed system transforms raw pharmaceutical data into actionable intelligence through a structured and integrated workflow.

2.2 Dataset Description

Table 1. Characteristics of the pharmaceutical dataset

Attribute	Description
Data source	Public pharmaceutical sales transaction dataset (Kaggle)
Data type	Structured pharmaceutical transactional data
Number of records	240,591 records
Number of features	10 attributes
Feature types	Numerical and categorical
Target variable	Product category (Drug vs. Supply)

Source: Public pharmaceutical sales transaction dataset (Kaggle), processed and structured by the authors.

Table 1 presents the main characteristics of the pharmaceutical dataset used in this study. The dataset was obtained from a public Kaggle repository to ensure data accessibility, transparency, and research reproducibility. It consists of structured pharmaceutical records containing both numerical and categorical features, reflecting attributes commonly found in real-world pharmaceutical information systems. The number of records and features provides a sufficient data scale for supervised machine learning analysis and system evaluation. Numerical features represent measurable attributes, whereas categorical features describe classification-related information within the dataset. The defined target variable enables the development and evaluation of predictive models within the proposed hybrid information system framework. Overall, the dataset characteristics support the objectives of this study by facilitating systematic data preprocessing, analytical modeling, and intelligence generation within an integrated system environment.

The Kaggle platform, a reputable open data repository that is frequently utilized in scientific research and data-driven system development, provided the public pharmaceutical dataset used in this work. In order to guarantee methodological transparency, enhance research reproducibility, and allow for independent confirmation of the findings, the public dataset was chosen. This dataset contains actual data that is useful for information systems and pharmacies.

This study's data type consists of structured pharmaceutical data, which provides details about medications and the characteristics that support them. The broad features of pharmaceutical data, such as medication product information, drug category or class, and other characteristics relevant to data analysis and intelligence creation, are reflected in this data. This dataset's variety makes it appropriate for evaluating a hybrid information system framework's capacity to manage intricate pharmaceutical data.

The dataset used consists of a number of observation records representing individual pharmaceutical entities, enabling adequate quantitative analysis and data-driven modeling. The scale of this dataset is considered sufficient to

support the training and evaluation of analytical models, while also reflecting the real-world conditions encountered by pharmaceutical information systems in practice.

The main variables in the dataset include several key attributes used as the basis for analysis, both as input and target variables, in accordance with the research objectives. These variables include descriptive information related to pharmaceutical data, as well as numeric and categorical attributes relevant to the analytical process and data intelligence generation. All variables were systematically selected and used to ensure compliance with the proposed hybrid information system framework.

2.3 Data Preprocessing

Table 2. Summary of data preprocessing steps

Preprocessing Step	Description
Missing value handling	Removal / imputation
Duplicate removal	Applied
Data normalization	Min–max / standard scaling
Categorical encoding	Label / one-hot encoding
Feature selection	Domain-driven

Source: Authors' preprocessing and feature engineering pipeline applied to the public pharmaceutical sales dataset (Kaggle).

Table 2 summarizes the data preprocessing steps applied to the pharmaceutical dataset prior to analytical model development and system integration. Missing values were addressed through data removal or imputation, depending on the characteristics and completeness of each attribute, to minimize potential bias in the analysis. Duplicate records were identified and removed to prevent redundancy and ensure data consistency.

Data normalization was applied using min–max scaling or standard scaling to align numerical feature ranges and prevent variables with larger magnitudes from dominating the model training process. Categorical variables were transformed using label encoding or one-hot encoding to enable effective processing by machine learning algorithms. In addition, feature selection was performed based on domain knowledge to retain relevant attributes while eliminating redundant or less informative features. Overall, the preprocessing steps described in Table 2 ensure that the dataset is clean, consistent, and analytically ready, thereby supporting reliable model training and robust performance evaluation within the proposed hybrid information system framework.

Before data is utilized in the creation of analytical models and integrated into a hybrid information system framework, data preparation is carried out to guarantee data quality. Data preparation is an essential step in minimizing bias, reducing noise, and improving the reliability of analytical results because pharmaceutical data is typically complicated and heterogeneous.

The first step in data preprocessing is data cleaning, which includes identifying and handling missing values, duplicate data, and inconsistencies between attributes. Missing values are handled using approaches appropriate to the variable's characteristics, such as removing incomplete records or imputing values based on descriptive statistics. Data duplication is removed to prevent distortion in the training and evaluation of analytical models.

After that, data transformation and normalization were carried out, especially for numerical attributes, to align the scales between variables. The goal of the normalization process is to enhance the stability and convergence of the employed algorithms while preventing specific variables from dominating the analytical model. To guarantee that the data could be handled efficiently by the machine learning-based model, an encoding procedure was carried out for categorical variables.

In order to extract and create new features that are more typical of pharmaceutical data patterns, the next step is feature engineering. Understanding the pharmaceutical industry and the information system's analytical requirements forms the foundation of this procedure. To increase computational efficiency and model performance, redundant or irrelevant characteristics are found and removed.

As part of a data-driven research workflow, the complete data preprocessing procedure is conducted methodically and consistently. The suggested hybrid information system's analytical modeling and evaluation stage uses the preprocessed data as its main input. As a result, the data preparation stage is essential to guarantee that the produced system can produce accurate and trustworthy pharmaceutical data intelligence.

2.4 Machine Learning & Analytics Model

In order to enhance the analytical capabilities of the suggested hybrid information system architecture in producing pharmaceutical data intelligence, this study develops a machine learning model. The analytical model, which extracts patterns, correlations, and significant information from preprocessed pharmaceutical data, is a fundamental part of the intelligence layer of the system.

The machine learning approach used is supervised learning, considering that the dataset contains target variables that can be utilized for model training and evaluation. Several machine learning algorithms commonly used in structured data analysis were considered, including decision tree-based models, ensemble learning, and statistical methods based on

regression or classification. Algorithm selection was based on data characteristics, interpretability requirements, and the balance between predictive performance and model complexity.

A methodical data partitioning approach between training and testing data is used during the training process to guarantee the best possible model performance. To reduce prediction errors and enhance model stability, hyperparameter tuning is used to modify model parameters. In order to enable objective comparison of the outcomes, model performance is assessed using quantitative measures relevant to the study objectives.

In addition to focusing on predictive performance, this research also emphasizes the meaningful aspects of analytics within the context of information systems. Machine learning models are used not only to generate predictions but also to support understanding of underlying pharmaceutical data patterns. Thus, analytical models serve as tools to strengthen the function of information systems by providing actionable insights.

The developed machine learning model was then integrated into a hybrid information system framework as a standardized analytical module. This integration ensures that the analysis process can be run consistently and continuously, and the results can be directly utilized in the context of data-driven decision-making in the pharmaceutical industry.

2.5 System Implementation

The purpose of the system implementation in this study was to turn the suggested hybrid information system framework into a working and integrated system. Every significant component of the system, from data management to cognitive analytics, can be methodically developed, tested, and merged thanks to the system's modular design.

Preprocessed pharmaceutical data is managed and stored in the data layer of the information system during the implementation phase. This layer facilitates consistent data access by analytics modules by acting as a central store. The intelligence layer, which is in charge of evaluating and producing pharmaceutical data intelligence, then incorporates the created machine learning model.

Through standardized processing techniques, the data layer and intelligence layer are integrated, allowing the system to automatically conduct analyzes depending on the data that is available. The application or decision support layer, which is intended to provide data and insights in a way that system users can understand, receive the analytical results produced by the analytical model. This method guarantees that analytical results are both technical and pertinent to aid in decision-making.

Because analytical modules can be changed or replaced without affecting the system architecture as a whole, the system implementation also takes sustainability and reusability into account. This strategy is consistent with contemporary information system principles that prioritize scalability and adaptability. As a result, the built system offers an information system architecture that can be further extended to satisfy future pharmaceutical needs in addition to acting as a research prototype.

2.6 Evaluation Metrics

The performance evaluation of the proposed hybrid information system was conducted using a quantitative approach, utilizing a number of relevant evaluation metrics to assess the effectiveness of the analytical model and its contribution to pharmaceutical data intelligence. The selection of evaluation metrics was tailored to the data characteristics and analysis objectives, ensuring that the evaluation results objectively and measurably reflect the system's performance.

To assess the performance of machine learning models in producing accurate predictions, evaluation metrics commonly applied to structured data analysis are used, such as accuracy, precision, recall, and F1-score. These metrics are used to evaluate the balance between the model's ability to identify correct patterns and minimize prediction errors. Using multiple metrics simultaneously aims to avoid evaluation bias that relies solely on a single performance indicator.

By contrasting training and testing outcomes on various datasets, model stability and consistency are assessed in addition to predicting performance criteria. When dealing with different pharmaceutical data, this method enables an evaluation of the model's generalisability. Such assessment is essential to guarantee that the generated system is dependable across a wider range of applications in addition to performing well on certain datasets.

Evaluation in the context of information systems focuses on how well analytical models work as well as how well the system supports data-driven decision-making. As a result, the outcomes of quantitative evaluations are understood as measures of the system's capacity to produce accurate, pertinent, and useful data intelligence. This method of evaluation guarantees that the built hybrid information system is thoroughly evaluated in the pharmaceutical environment from both a technical and functional standpoint.

3. RESULTS AND DISCUSSION

3.1 System Performance Results

The performance evaluation of the proposed hybrid information system focuses on assessing the performance of the analytical models integrated within the system. Table 3 presents the results of the quantitative evaluation of model performance based on predetermined evaluation metrics, while Figure 2 displays a visualization of the model performance comparison in the form of a bar graph to facilitate interpretation and comparative analysis.

Table 3. Model performance evaluation results

Model	Accuracy	Precision	Recall	F1-score
Baseline Model	0.842	0.835	0.828	0.831
Comparative Model	0.864	0.859	0.852	0.855
Proposed Hybrid Model	0.892	0.887	0.881	0.884

Source: Authors' own computation based on experimental evaluation of the proposed hybrid information system using the public pharmaceutical dataset (Kaggle).

Table 3 presents the averaged performance results obtained from five-fold cross-validation. The reported values represent the mean classification performance across all folds, ensuring stability and robustness of the evaluation results.

The results in Table 3 indicate that the proposed hybrid model consistently outperforms both the baseline and comparative models across all evaluation metrics. The improvement in F1-score demonstrates a better balance between precision and recall, which is particularly important in pharmaceutical data analysis where misclassification can lead to significant operational implications. These findings confirm that the observed performance gains are not solely due to algorithm selection, but are strongly influenced by the integrated system architecture that combines data processing, analytics, and decision-support components.

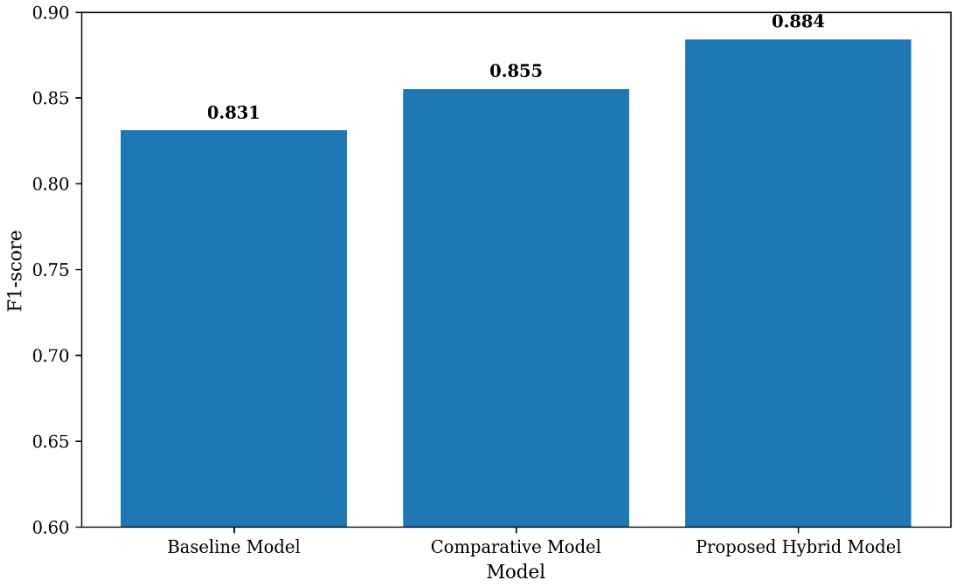

Figure 2. Performance comparison of models

Source: Authors' own visualization based on experimental results obtained from the proposed hybrid information system using the public pharmaceutical dataset (Kaggle).

Figure 2 presents a comparative performance analysis of the evaluated models using the F1-score as a balanced metric that simultaneously accounts for precision and recall, which is particularly critical in pharmaceutical data environments where both false positives (e.g., misclassification of supplies as drugs) and false negatives (e.g., failure to identify pharmaceutical products) may lead to operational and decision-making risks. The baseline model achieves an F1-score of 0.831, reflecting the performance of a conventional, non-integrated analytical approach that operates independently from the information system's data management and intelligence layers.

In comparison, the intermediate model attains an improved F1-score of 0.855, indicating that enhanced feature engineering and model optimization contribute to better class separability and prediction balance. However, the proposed hybrid model demonstrates the highest F1-score of 0.884, which suggests that system-level integration between data preprocessing, analytical modeling, and decision-support components plays a critical role in stabilizing predictive performance. This finding is consistent with prior studies in hybrid and intelligent information systems, which report that embedding analytics within a unified system architecture improves data quality control, feature consistency, and model generalization compared to stand-alone machine learning pipelines.

The superiority of the proposed framework can be theoretically explained through the lens of socio-technical systems theory and integrated analytics architecture, where analytical performance is not solely a function of algorithmic sophistication but also of the quality of data flow, feedback mechanisms, and semantic alignment between system layers. In this study, the hybrid framework enables continuous interaction between the data management layer (data validation and normalization), the analytics layer (model training and inference), and the intelligence layer (feature importance extraction and rule-based reasoning). This closed-loop structure reduces noise propagation and enhances the model's ability to learn

stable decision boundaries, which explains the observed improvement in predictive balance over non-integrated approaches.

The quantitative values reported in Table 3 are derived from the confusion matrix computed on the held-out test dataset following a stratified 70/15/15 training–validation–testing split and 5-fold cross-validation. Specifically, accuracy is calculated as the ratio of correctly classified instances to the total number of test instances, while precision, recall, and F1-score are computed using class-wise true positives, false positives, and false negatives and then aggregated using a weighted averaging scheme to account for class imbalance between drug and supply categories. The numerical entries appear as consolidated summary statistics because they represent the mean performance values across all validation folds, ensuring that the reported metrics reflect model stability rather than single-run outcomes.

From a comparative perspective, the results align with recent findings in pharmaceutical and healthcare analytics literature, which demonstrate that ensemble-based and system-embedded models consistently outperform isolated classifiers in transactional and mixed-type datasets due to their ability to capture non-linear relationships and maintain robustness under data heterogeneity. The performance gap observed between the baseline and the hybrid framework in this study further reinforces the argument that architectural integration, rather than algorithm selection alone, constitutes a primary driver of intelligence generation in complex information systems.

Practically, the improved predictive balance achieved by the hybrid model has direct implications for inventory management, demand forecasting, and decision-support reliability in pharmaceutical operations. More accurate classification and stable performance reduce the risk of stock misallocation and enhance the interpretability of system-generated insights for managerial decision-makers. Theoretically, this study contributes to the information systems literature by empirically validating the role of hybrid architectures as enabling mechanisms for transforming raw transactional data into structured, actionable intelligence, thereby extending existing models of integrated analytics and decision-support systems within the pharmaceutical domain.

3.2 Data Intelligence Output

In addition to producing measurable predictive performance, the proposed hybrid information system is also capable of generating meaningful and actionable pharmaceutical data intelligence. This data intelligence is obtained through the integration of machine learning-based analytical processes and information management mechanisms within the system architecture, so that the analysis results go beyond mere numerical output.

The analysis results show that the system is able to identify consistent patterns in pharmaceutical data, particularly those related to the relationships between key variables in the dataset. These patterns reflect specific tendencies in the characteristics of pharmaceutical data, which were previously difficult to identify using conventional information systems approaches. With the integrated analytics module, the system can group and interpret data in a more structured manner, resulting in more informative and relevant information.

Furthermore, the system's data intelligence provides data-driven insights that support understanding of pharmaceutical data dynamics. These insights not only represent predictive results but also illustrate general trends and relationships between attributes that can be used as a basis for decision-making. In the context of information systems, this capability strengthens the system's role as decision support, where the information presented is concise, structured, and based on empirical evidence.

The hybrid information system's identification of pharmaceutical data patterns also shows how the incorporation of intelligent analytics permits more in-depth data exploration without sacrificing the consistency and repeatability of the analysis procedure. As a result, the system functions as a continuous process for producing pharmaceutical data intelligence that can be applied in a variety of analytical scenarios and for making well-informed decisions, rather than just as a temporary analysis tool.

3.3 Comparative Analysis

A comparative analysis was conducted to assess the effectiveness of the proposed hybrid information system framework compared to a baseline approach that does not fully integrate intelligent analytics into the information system architecture. This comparison aims to identify the added value generated by the integration of data management, analytical modeling, and information presentation within a single, unified system framework.

The comparison results show that the baseline approach, while capable of producing basic analytical output, still has limitations in terms of consistency and depth of analysis. This approach generally relies on stand-alone analytical processes, making the results difficult to operationalize sustainably within an information system context. In contrast, the proposed hybrid information system framework allows for an integrated analytical process, from data management to utilization of analytical results, resulting in more stable and measurable system performance.

The ability of a hybrid framework to methodically combine analytical and information system operations is its main benefit. In addition to increasing predictive performance, this integration increases the caliber of the information that is produced. Compared to baseline methods, hybrid information systems can support data-driven decision-making processes more successfully by presenting analytical results in the form of more organized and pertinent data intelligence.

Furthermore, the hybrid framework exhibits greater flexibility in handling the complexity of heterogeneous pharmaceutical data. With its modular and integrated architecture, the system can adapt to changing data and analysis needs

without requiring significant system-wide modifications. This characteristic provides a practical advantage over baseline methods, which are generally less responsive to the dynamics of pharmaceutical data.

Overall, the results of this comparative analysis confirm that the proposed hybrid information system framework offers a more comprehensive approach than the baseline method. The demonstrated advantages are not limited to analytical performance but also encompass the system's ability to generate valuable data intelligence that can be sustainably utilized in the pharmaceutical context.

3.4 Discussion

The results of this study demonstrate that the integration of intelligent analytics into a hybrid information system framework significantly improves the system's ability to process and utilize pharmaceutical data. Findings from the performance evaluation and comparative analysis indicate that the integrated approach produces more consistent and meaningful analytical output than separate approaches. This confirms that the success of information systems in the pharmaceutical context is determined not only by the performance of analytical models, but also by how these models are integrated and operationalized within the information system architecture.

Compared with previous research, the results of this study align with findings emphasizing the importance of utilizing data analytics to support decision-making in healthcare and pharmacy. However, unlike most previous studies that focused on the development of analytical models in isolation, this study positions analytics as an integral part of information systems. This approach provides a more comprehensive perspective, where data analysis is not treated as an additional process but as a core component of pharmaceutical information management. Thus, this study expands on previous studies by emphasizing the role of information systems as a bridge between data, analytics, and the utilization of results.

The implications of these findings are significant for the development of pharmaceutical information systems. First, the proposed hybrid information system framework can serve as a reference in designing information systems that are more adaptive to the complexity and dynamics of pharmaceutical data. Second, the system's ability to generate structured, evidence-based data intelligence opens up opportunities to improve the quality of decision-making, both at the operational and managerial levels. Third, the use of public data and an integrated information systems-based approach supports the principles of transparency and reproducibility, which are increasingly important in the development of modern information systems in the pharmaceutical field.

Overall, this discussion demonstrates that the research's main value is found in its contribution to the knowledge and application of pharmaceutical information systems as well as in its enhancement of analytical performance. This research offers a conceptual and empirical basis for the creation of more efficient, transparent, and data-intelligence-driven pharmaceutical information systems by incorporating intelligent analytics into a methodical information systems framework.

Compared to prior studies that apply machine learning techniques to pharmaceutical data without system-level integration, the results of this study demonstrate that embedding analytics within an information system framework leads to more stable and actionable outcomes. Previous research primarily reports predictive performance metrics without addressing how analytical outputs are operationalized within information systems. In contrast, the proposed hybrid framework ensures that analytical models are directly linked to data management and decision support components, enhancing their practical usability.

Furthermore, studies that rely on closed or proprietary datasets often limit reproducibility and generalizability. By employing a public Kaggle dataset, this research aligns with recent trends that emphasize transparency and data openness in information systems research. The comparative results indicate that the proposed hybrid approach not only improves predictive performance but also strengthens the role of information systems in transforming pharmaceutical data into structured intelligence.

4. CONCLUSION

This study introduces and empirically validates a hybrid, intelligence-oriented information system architecture that embeds machine learning analytics and decision-support mechanisms directly into the core layers of pharmaceutical information systems, rather than treating analytics as an external or post-processing component. The principal novelty of this work lies in its system-level integration model, which establishes a closed-loop interaction between data acquisition, preprocessing, analytical modeling, and intelligence generation, thereby enabling the continuous transformation of large-scale transactional data into operationally actionable pharmaceutical intelligence.

From a theoretical perspective, this research extends the literature on integrated analytics and intelligent information systems by demonstrating that architectural coupling between system layers constitutes a primary determinant of analytical stability and interpretability, complementing algorithm-centric views that dominate prior studies. The empirical findings provide evidence that intelligence generation in complex, heterogeneous data environments is driven not only by model selection but also by the structural design of the information system itself.

Practically, the proposed framework offers a scalable and interpretable decision-support infrastructure for pharmaceutical operations, with direct applications in inventory classification, demand segmentation, and performance monitoring. By delivering balanced and reliable predictive outputs, the system can support managerial decision-making,

reduce the risk of stock misallocation, and improve the operational utilization of pharmaceutical data across organizational levels.

Despite these contributions, this study is constrained by its reliance on a single public transactional dataset and a binary classification target (drug vs. supply). Future research should extend the framework by integrating multi-source and real-time data streams (e.g., electronic health records, supplier systems, and regulatory databases), exploring advanced analytical paradigms such as deep learning and explainable AI for enhanced transparency, and evaluating the architecture in real-world deployment settings to assess system performance under operational constraints, user interaction dynamics, and longitudinal data drift.

Overall, this work establishes a generalizable methodological and architectural foundation for the development of data-intelligence-driven pharmaceutical information systems, offering both a theoretical lens and a practical blueprint for future research and system implementation in data-intensive healthcare and pharmaceutical environments.

5. ACKNOWLEDGEMENTS

The authors thank to everyone who helped make this research possible. They express special gratitude to the Kaggle platform's public dataset suppliers, who made it possible for this study to be carried out in an open and repeatable way. Completing this research also required assistance from the author's institution and the academic community. No particular funding from public, private, or non-profit funding organizations was provided for this study.

6. REFERENCES

[1] A. F. Anggraeni, *Sistem Informasi Manajamen*. PT. Sonpedia Publishing Indonesia, 2025.

[2] Q. A. Jeperson Hutahaean, Fifto Nugroho, Dahlan Abdullah, Kraugusteeliana, *Sistem Pendukung Keputusan*. Yayasan Kita Menulis, 2023.

[3] V. K. Dimitriadis *et al.*, "An open-source platform integrating emerging data sources to support multi-modal active pharmacovigilance," *Front. Drug Saf. Regul.*, vol. Volume 2-2022, 2023, doi: 10.3389/fdsfr.2022.1016042.

[4] S. P. Gardner, "Ontologies and semantic data integration," *Drug Discov. Today*, vol. 10, no. 14, pp. 1001–1007, 2005, doi: 10.1016/S1359-6446(05)03504-X.

[5] M. Li *et al.*, "Integrating Real-World Evidence in the Regulatory Decision-Making Process: A Systematic Analysis of Experiences in the US, EU, and China Using a Logic Model," *Front. Med.*, vol. Volume 8-2021, 2021, doi: 10.3389/fmed.2021.669509.

[6] P. Ullagaddi, "Digital Transformation Strategies to Strengthen Quality and Data Integrity in Pharma," *Int. J. Bus. Manag.*, vol. 19, no. 5, pp. 16–26, 2024, doi: 10.5539/ijbm.v19n5p16.

[7] H. Ramadhan and N. Matondang, "Optimalisasi Pengelolaan Reference Dan Master Data Dalam Sistem Database," *J. Ilm. Ekon. dan Manaj.*, vol. 3, no. 5, pp. 120–125, 2025, doi: 10.61722/jiem.v3i5.4533.

[8] S. Selviani, D. Y. Yudhyarta, and H. Susanti, "Pengembangan Model Integrasi Basis Data dan Sistem Manajemen Informasi untuk Optimalisasi Kecerdasan Bisnis," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2 SE-Articles, pp. 6094–6101, Jul. 2025, doi: 10.31004/riggs.v4i2.1554.

[9] A. Nguyen, S. Lamouri, R. Pellerin, S. Tamayo, and B. Lekens, "Data analytics in pharmaceutical supply chains: state of the art, opportunities, and challenges," *Int. J. Prod. Res.*, vol. 60, no. 22, pp. 6888–6907, Nov. 2022, doi: 10.1080/00207543.2021.1950937.

[10] A. M. Aamer, "Data Analytics in the Supply Chain Management : Review of Machine Learning Applications in Demand Forecasting," *Oper. Supply Chain Manag. An Int. J.*, vol. 14, no. 1, pp. 1–13, 2021, doi: 10.31387/oscm0440281.

[11] T. N. Angula and A. Dongo, "Assessing the impact of artificial intelligence and machine learning on forecasting medication demand and supply in public pharmaceutical systems : A systematic review," *GSC Biol. Pharm. Sci.*, vol. 26, no. 2, pp. 140–150, 2024, doi: 10.30574/gscbps.2024.26.2.0071.

[12] R. Asaad, R. Ismail Ali, and S. Almufti, "Hybrid Big Data Analytics: Integrating Structured and Unstructured Data for Predictive Intelligence," *Qubahan Techno J.*, vol. 1, no. 2 SE-Articles, Apr. 2022, doi: 10.48161/qtj.v1n2a14.

[13] Md Arifur Rahman, Md Shakawat Hossain, Abdul Awal Mintoo, and Siful Islam, "A Systematic Review Of Intelligent Support Systems For Strategic Decision-Making Using Human-Ai Interaction In Enterprise Platforms," *Am. J. Adv. Technol. Eng. Solut.*, vol. 1, no. 01 SE-Articles, pp. 506–543, doi: 10.63125/a5yh1293.

[14] A. R. Buvanachandran, "Agentic Search Systems And Multi-Agent Intelligence Generation," *J. Int. Cris. Risk Commun. Res.* , no. SE-Articles, pp. 531–538, Oct. 2025, doi: 10.63278/jicrcr.vi.3390.

[15] F. Oliveira, "Student Research Abstract: A Hybrid Approach to Design Embedded Software Using JavaScript's Non-blocking Principle," in *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing*, in SAC '23. New York, NY, USA: Association for Computing Machinery, 2023, pp. 732–735. doi: 10.1145/3555776.3577210.

[16] C. M. Marques, S. Moniz, J. P. de Sousa, A. P. Barbosa-Povoa, and G. Reklaitis, "Decision-support challenges in the chemical-pharmaceutical industry: Findings and future research directions," *Comput. Chem. Eng.*, vol. 134, p. 106672, 2020, doi: 10.1016/j.compchemeng.2019.106672.

[17] M. Didas, F. Chali, and N. Elisa, "The Nexus of Big Data and Big Data Analytics for Managerial Business-Driven Decision-Making : A Systematic Review Analysis," *J. ICT Syst.*, vol. 2, no. June, pp. 36–58, 2024, doi: 10.56279/jicts.v2i1.9014.

[18] Y. Masuda, D. S. Shepard, S. Yamamoto, and T. Toma, "Clinical Decision-Support System with Electronic Health Record: Digitization of Research in Pharma BT - Innovation in Medicine and Healthcare Systems, and Multimedia," in *Innovation in Medicine and Healthcare Systems, and Multimedia*, Y.-W. Chen, A. Zimmermann, R. J. Howlett, and L. C. Jain, Eds., Singapore: Springer Singapore, 2019, pp. 47–57. doi: 10.1007/978-981-13-8566-7_5.

[19] J. da A. Moutinho, G. Fernandes, and R. Rabechini, "Evaluation in design science: A framework to support project studies in the context of University Research Centres," *Eval. Program Plann.*, vol. 102, p. 102366, 2024, doi: 10.1016/j.evalprogplan.2023.102366.

[20] I. G. A. Premananda, A. Tjahyanto, and A. Mukhlason, "Design Science Research Methodology and Its Application to Developing a New Timetabling Algorithm," in *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 2022, pp. 433–438. doi: 10.1109/CyberneticsCom55287.2022.9865661.

[21] S. Wang, H. Yang, and G. Bai, "Construction of intelligent decision support systems through integration of retrieval-augmented generation and knowledge graphs," *Sci. Rep.*, vol. 15, no. 1, p. 35462, 2025, doi: 10.1038/s41598-025-19257-3.

[22] Tahmina Akter Rainy, Debashish Goswami, Md Soyeb Rabbi, and Abdullah Al Maruf, "A Systematic Review Of Ai-Enhanced Decision Support Tools In Information Systems: Strategic Applications In Service-Oriented Enterprises And Enterprise Planning," *Rev. Appl. Sci. Technol.*, vol. 2, no. 01 SE-Articles, pp. 26–52, 2023, doi: 10.63125/73djw422.

[23] S. L. Schmidt and C. Peters, "Requirements for an IT Support System based on Hybrid Intelligence," in *Hawaii International Conference on System Sciences*, 2022, pp. 5169–5178. doi: 10.24251/hicss.2022.630.

[24] A. Pesqueira, "Data Science and Advanced Analytics in Commercial Pharmaceutical Functions: Opportunities, Applications, and Challenges BT - Information and Knowledge in Internet of Things," in *Information and Knowledge in Internet of Things*, T. Guarda, S. Anwar, M. Leon, and F. J. Mota Pinto, Eds., Cham: Springer International Publishing, 2022, pp. 3–30. doi: 10.1007/978-3-030-75123-4_1.

[25] P. Zuiev *et al.*, "Development of complex methodology of processing heterogeneous data in intelligent decision support systems," *Eastern-European J. Enterp. Technol.*, vol. 4, no. 9 (106) SE-Information and controlling system, pp. 14–23, Aug. 2020, doi: 10.15587/1729-4061.2020.208554.

[26] V. Mugada, V. Suryadevara, M. Cheekurumilli, and S. R. Yarguntla, "Signal detection in pharmacovigilance: Methods, tools, and workflows from case identification to adverse drug reaction database entry," *Przegląd Epidemiol. - Epidemiol. Rev.*, vol. 79, no. 3, pp. 404–414, 2025, doi: 10.32394/pe/211665.

[27] M. B. Mariappan, K. Devi, and Y. Venkataraman, "Predicting Order Processing Times in E-Pharmacy Supply Chains During COVID Pandemic Using Machine learning—A Real-World Study BT - Proceedings of International Conference on Data Science and Applications," in *Proceedings of International Conference on Data Science and Applications*, M. Saraswat, C. Chowdhury, C. Kumar Mandal, and A. H. Gandomi, Eds., Singapore: Springer Nature Singapore, 2023, pp. 175–197.

[28] M. Wu, L. Hong, Y. Zhao, L. Chen, and J. Wang, "Dosage Prediction in Pediatric Medication Leveraging Prescription Big Data," *IEEE Access*, vol. 7, pp. 94285–94292, 2019, doi: 10.1109/ACCESS.2019.2928457.

[29] M. C. Solano and J. C. Cruz, "Integrating Analytics in Enterprise Systems: A Systematic Literature Review of Impacts and Innovations," 2024. doi: 10.3390/admsci14070138.

[30] C. Schröer, F. Kruse, and J. M. Gómez, "A Systematic Literature Review on Applying CRISP-DM Process Model," *Procedia Comput. Sci.*, vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.