International Journal of Education and Life Sciences (IJELS)
Vol. 4 No. 6 (2026): 761-772

Machine Learning-Based Patient No-Show Prediction for Hospital Appointment Systems

Fitria[1*], Nitami Lestari Putri[2]
Accounting Information Systems, Politeknik Negeri Banjarmasin
Smart City Information System, Politeknik Negeri Banjarmasin
Corresponding Author: Fitria fitria@poliban.ac.id

ARTICLEINFO
Keywords: Patient No-Show, Machine Learning, Decision Tree, Hospital Information System, Healthcare Analytics

Received : 27, April
Revised : 20, May
Accepted: 8, June

ABSTRACT
Patient no-show is a major challenge in hospital appointment systems because it affects operational efficiency, physician utilization, and healthcare service quality. This study aims to implement a machine learning approach for predicting patient no-show in Hospital Management Information Systems (HMIS) using the Decision Tree algorithm. The dataset used in this study consisted of 200 hospital appointment records containing appointment schedules, physician information, visit reasons, and appointment status. The research methodology included exploratory data analysis, data preprocessing, feature engineering, categorical data encoding, and machine learning modeling. The analysis showed that no-show appointments represented a considerable portion of the dataset, with higher no-show frequencies occurring on certain appointment days and sessions. The Decision Tree model was trained using appointment-related features, including doctor_id, reason_for_visit, day_name, and session. The evaluation results demonstrated that the model achieved an accuracy of 60% with acceptable recall performance for detecting no-show patients. The findings indicate that machine learning has the potential to support predictive analytics and operational decision-making in hospital appointment systems.

DOI: https://doi.org/10.59890/ijels.v4i6.23
ISSN: 3026-2429
https://journal.multitechpublisher.com/index.php/ijels

INTRODUCTION

Hospital Management Information Systems (HMIS) play an important role in improving healthcare services and supporting operational efficiency in hospitals (Aldali, 2024; Van et al., 2020). One of the major challenges in hospital appointment management is patient no-show, which occurs when patients fail to attend scheduled appointments without prior notification. This issue causes inefficient physician utilization, increased idle service time, operational losses, and reduced healthcare service effectiveness.

Patient no-show has become a significant concern in healthcare management because it directly affects scheduling systems, medical resource allocation, and hospital financial performance. Hospitals require effective strategies to predict patient attendance behaviour to minimize operational inefficiencies. Along with the rapid development of artificial intelligence and machine learning technologies, healthcare data can be utilized to identify patterns and predict patient behaviour more effectively.

Machine learning techniques have been widely applied in healthcare analytics for disease prediction, patient monitoring, and healthcare decision support systems. In appointment management systems, machine learning models can analyse historical appointment data to predict the likelihood of patient no-show. Such predictive systems may help hospitals optimize appointment scheduling and improve healthcare service quality.

This study aims to implement a machine learning-based approach for predicting patient no-show in hospital appointment systems using the Decision Tree algorithm. The study also seeks to analyse patient appointment patterns and evaluate the performance of the developed prediction model.

LITERATURE REVIEW

Hospital Management Information System

Hospital Management Information Systems are integrated systems designed to manage healthcare services, patient administration, medical records, and operational processes within hospitals (Olamijuwon & Zouo, 2024; Sarla, 2024; Shiyanbade, 2024). HMIS supports decision-making processes by providing accurate and timely information for healthcare management.

Effective appointment management within HMIS is essential to ensure optimal physician utilization and patient service quality. However, patient no-show remains one of the major operational challenges in healthcare systems.

Machine Learning in Healthcare

Machine learning is a branch of artificial intelligence that enables systems to learn patterns from historical data and make predictions automatically (Anjum et al., 2021; Hassan, 2024; Tenepalli & Navamani, 2024). In healthcare, machine learning has been applied in disease diagnosis, healthcare analytics, patient risk prediction, and operational management.

Several studies have demonstrated that machine learning algorithms can be utilized to predict patient no-show behavior based on historical appointment data. Prediction models can assist hospitals in reducing scheduling inefficiencies and improving healthcare resource management.

Decision Tree Algorithm

Decision Tree is a supervised machine learning algorithm commonly used for classification tasks. The algorithm creates a tree-like structure based on decision rules derived from training data. Decision Tree is widely used because it is easy to interpret and capable of handling categorical data effectively (Bodria et al., 2023; Liu, 2022; Nagahisarchoghaei et al., 2023).

In healthcare analytics, Decision Tree algorithms have been successfully applied for patient classification, disease prediction, and healthcare operational analysis.

Hypothesis

H1: Machine learning can be used to predict patient no-show in hospital appointment systems.

H2: Appointment schedule characteristics influence patient no-show behavior.

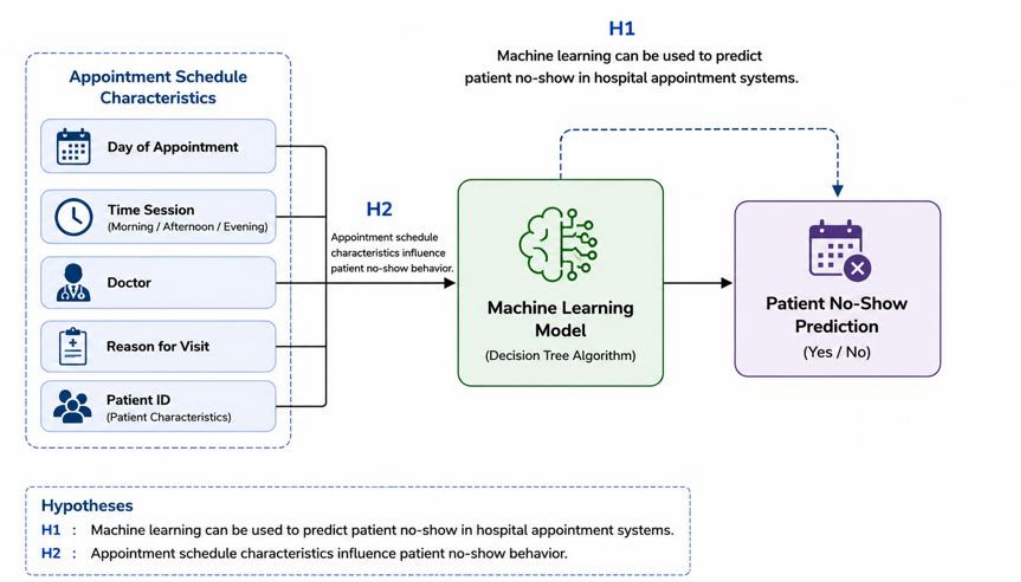

Figure 1. Conceptual Framework for Patient No-Show Prediction Using Machine Learning

METHODOLOGY

Dataset

The dataset used in this research consists of 200 hospital appointment records with the following variables:

Table 1. Variables of Hospital Appointment Dataset

Variable	Description
appointment_id	Appointment ID
patient_id	Patient ID
doctor_id	Doctor ID
appointment_date	Appointment date
appointment_time	Appointment time
reason_for_visit	Reason for visit
status	Appointment status

The appointment status includes:

- Completed
- No-show
- Scheduled
- Cancelled

For machine learning implementation, the study focused only on Completed and No-show records.

Research Stages

The research stages consist of:

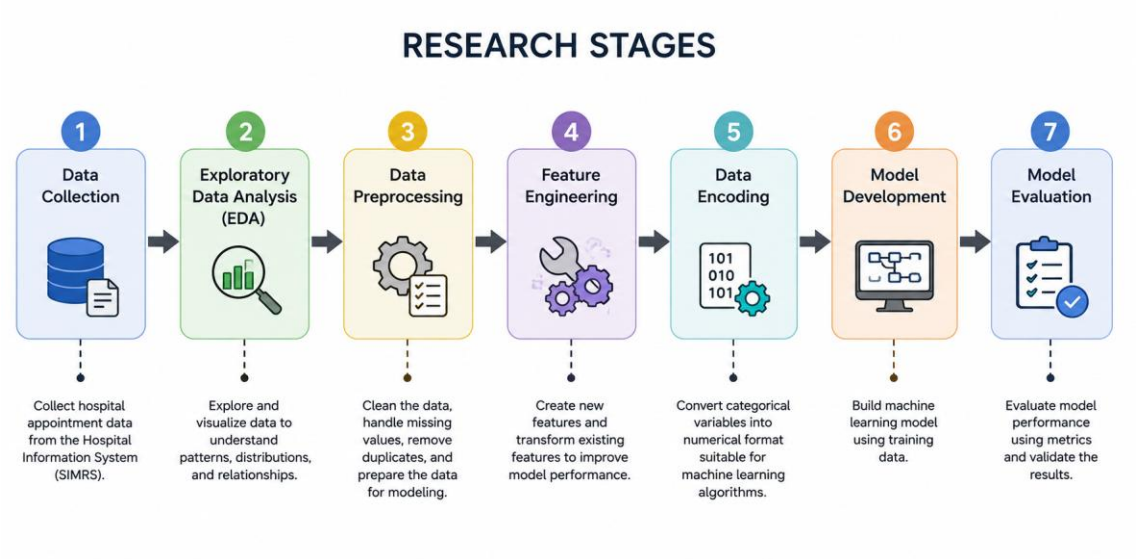

Figure 2. Research Stage

Exploratory Data Analysis

EDA was conducted to analyze the characteristics of the appointment data and identify no-show patterns.

The analysis showed that the dataset contains:

- 52 No-show records
- 46 Completed records
- 51 Scheduled records
- 51 Cancelled records

The hospital appointment dataset also contains 10 doctors and 48 patients.

Further analysis showed that Wednesday had the highest number of no-show appointments. Additionally, no-show appointments were more frequent during Morning and Afternoon sessions.

Data Preprocessing

Data preprocessing was conducted to prepare the dataset before implementing the machine learning model. This stage included data cleaning, transformation, and feature preparation to improve data quality and model performance. First, the dataset was checked for missing values and duplicate records to ensure data consistency. The appointment date and appointment time

variables were transformed into appropriate datetime formats to facilitate temporal analysis.

Furthermore, feature engineering was performed by generating additional variables, including day_name, hour, and session. The session variable was categorized into Morning, Afternoon, and Evening based on appointment time. The study also focused only on appointment records with Completed and No-show status, while Scheduled and Cancelled records were excluded from the modeling process.

Categorical variables such as doctor_id, reason_for_visit, day_name, and session were converted into numerical representations using one-hot encoding. Finally, the dataset was divided into training and testing data using an 80:20 ratio before implementing the Decision Tree algorithm for patient no-show prediction.

Machine Learning Model

This study implemented the Decision Tree algorithm to predict patient no-show behavior in hospital appointment systems. The Decision Tree algorithm was selected because of its ability to handle categorical data effectively and provide interpretable classification results. The model was developed using appointment-related variables extracted from the hospital appointment dataset.

Before model training, the dataset was divided into training and testing data using an 80:20 ratio. The training dataset was used to build the prediction model, while the testing dataset was utilized to evaluate model performance. Several appointment-related features were used as input variables, including doctor_id, reason_for_visit, day_name, and session. These variables were selected because they represent appointment schedule characteristics and healthcare service patterns that may influence patient attendance behavior.

The Decision Tree model was trained to classify appointment status into two categories, namely Completed and No-show. After the training process, the model performance was evaluated using accuracy, precision, recall, and F1-score metrics to measure the effectiveness of patient no-show prediction.

Model Evaluation

The performance of the Decision Tree model was evaluated using several classification metrics, including accuracy, precision, recall, and F1-score. Accuracy was used to measure the overall correctness of the prediction model, while precision measured the model's ability to correctly identify no-show patients among all predicted no-show cases. Recall was utilized to evaluate the model's capability to detect actual no-show patients, which is important in hospital operational management to minimize appointment inefficiencies. In addition, the F1-score was used to provide a balanced evaluation between precision and recall. These evaluation metrics were applied to the testing dataset to assess the effectiveness of the machine learning model in predicting patient no-show behavior within the hospital appointment system.

RESEARCH RESULT

Exploratory Data Analysis Results

The distribution of appointment status is shown in Figure 1.

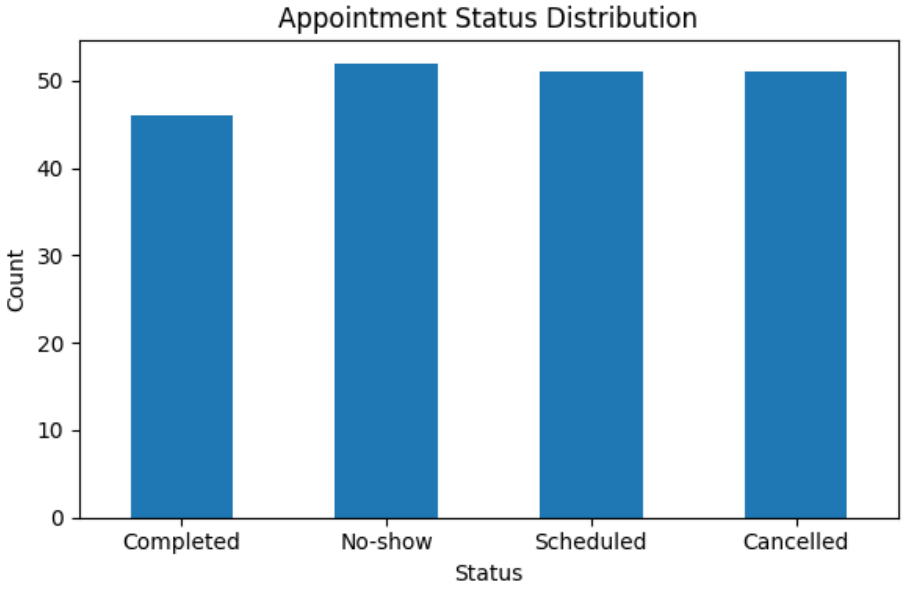

Figure 3. Distribution of Hospital Appointment Status

Figure 1 illustrates the distribution of appointment status in the hospital appointment dataset. The results show that the dataset contains four appointment categories, namely Completed, No-show, Scheduled, and Cancelled. The No-show category represents a considerable portion of the dataset, indicating that patient no-show is a relevant issue in hospital appointment systems.

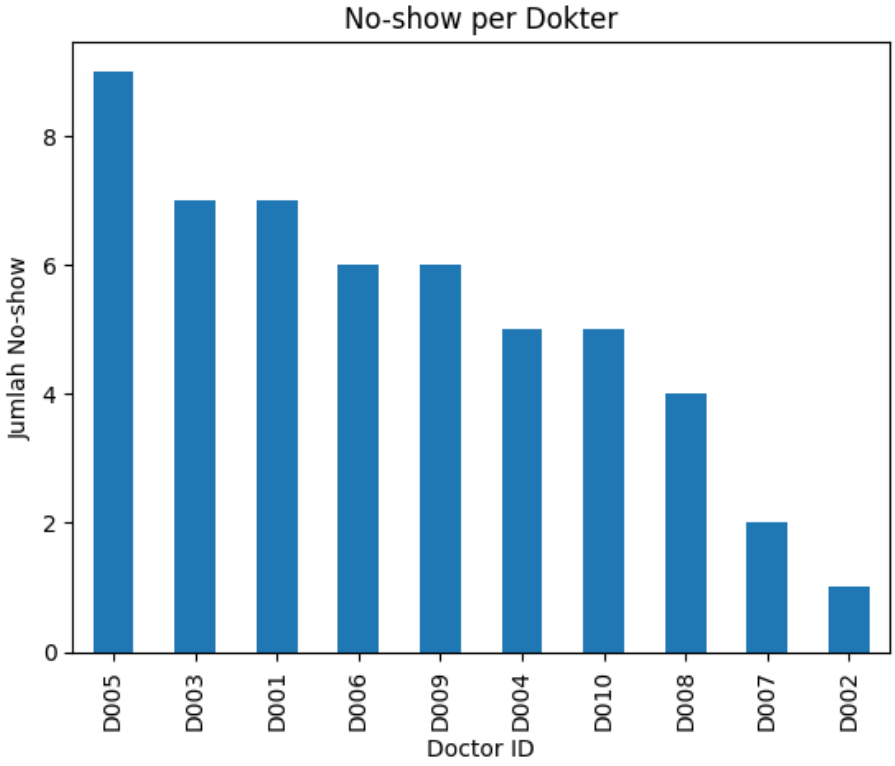

Figure 4. No-show Appointments by Doctor

Figure 4 presents the distribution of no-show appointments based on doctor identification. The results indicate that several doctors experienced higher no-show frequencies compared to others. Doctor D005 recorded the highest number of no-show appointments, followed by D003 and D001. This finding suggests that appointment attendance behavior may vary depending on physician-related factors and appointment scheduling patterns.

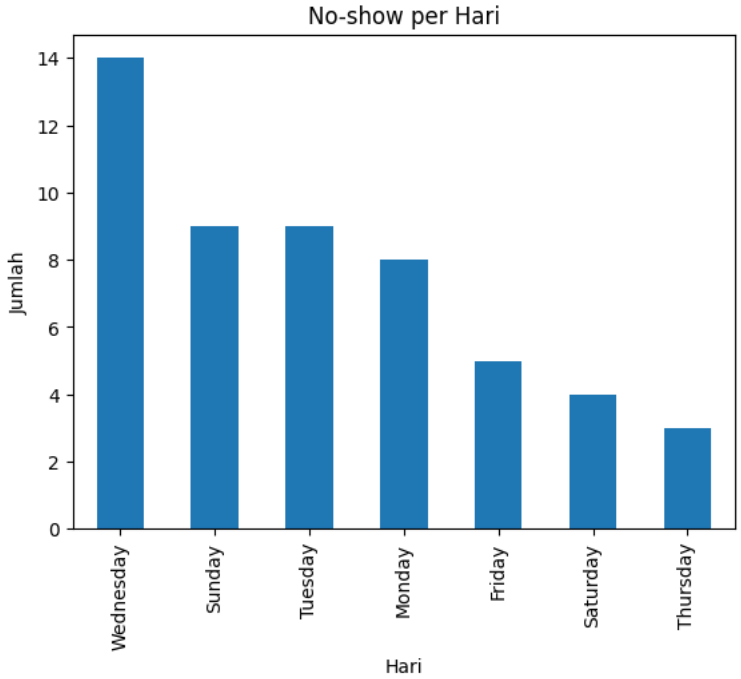

Figure 5. No-show Appointments by Day

Figure 5 illustrates the distribution of no-show appointments based on appointment days. The analysis shows that Wednesday had the highest number of no-show appointments compared to other days. This result indicates that appointment schedules may influence patient attendance behavior in hospital appointment systems.

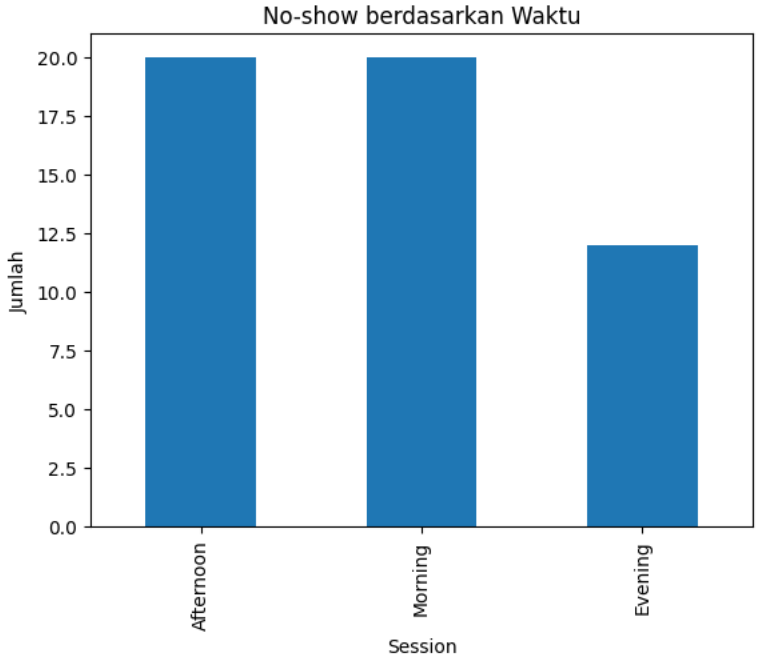

Figure 6. No-show Appointments by Session

Figure 6 shows the distribution of no-show appointments based on appointment sessions. The Morning and Afternoon sessions recorded higher no-show frequencies than the Evening session. This finding suggests that appointment time characteristics may affect patient attendance behavior.

Decision Tree Model Results

The Decision Tree model was successfully implemented to predict patient no-show behavior in hospital appointment systems. The model was trained using appointment-related variables including doctor_id, reason_for_visit, day_name, and session. Overall model accuracy: 60%

Table 2. Decision Tree Evaluation Results

Metric	Class 0	Class 1
Precision	0.70	0.50
Recall	0.58	0.62
F1-score	0.64	0.56

DISCUSSION

The implementation of machine learning for patient no-show prediction demonstrates that hospital appointment data can be utilized to identify patient attendance behavior patterns (Aghaeifar et al., 2023, 2023; Hartmann et al., 2024). Based on the exploratory data analysis results, the dataset contained a considerable number of no-show appointments, indicating that patient no-show is a relevant operational issue in hospital appointment systems. This condition may affect physician utilization, appointment scheduling efficiency, and healthcare service quality.

The exploratory analysis also revealed several important patterns related to patient no-show behavior. The analysis based on doctor identification showed that certain doctors experienced higher no-show frequencies than others. This finding suggests that physician-related factors and appointment management characteristics may influence patient attendance behavior. In addition, the

analysis of appointment days demonstrated that Wednesday recorded the highest number of no-show appointments, indicating that appointment schedules may affect patient attendance patterns.

Furthermore, the analysis based on appointment sessions showed that Morning and Afternoon sessions experienced higher no-show frequencies compared to Evening sessions (Aghaeifar et al., 2023; Barrera et al., 2020; Berquedich et al., 2020). This result indicates that appointment time characteristics may contribute to patient attendance behavior in hospital appointment systems. These findings support the second hypothesis stating that appointment schedule characteristics influence patient no-show behavior.

The Decision Tree algorithm was successfully implemented to classify patient appointment status into Completed and No-show categories. The model achieved an accuracy of 60%, indicating moderate classification performance. Although the overall accuracy was relatively moderate, the model demonstrated acceptable recall performance for detecting no-show patients. In healthcare operational management, recall is considered important because hospitals need to identify potential no-show patients in order to reduce scheduling inefficiencies and minimize idle physician time.

Several factors may have influenced the model performance obtained in this study. First, the dataset size was relatively small, consisting of only 200 appointment records. Limited data may reduce the model's ability to learn complex attendance behavior patterns. Second, the study used limited variables, mainly appointment-related features such as doctor_id, reason_for_visit, day_name, and session. Additional variables such as patient demographics, payment methods, previous appointment history, and waiting time may improve prediction performance in future studies.

Despite these limitations, the findings indicate that machine learning has strong potential for implementation in Hospital Management Information Systems (HMIS). Predictive models for patient no-show may assist hospitals in improving appointment scheduling efficiency, optimizing healthcare resource allocation, and supporting operational decision-making processes. The results of this study are also consistent with previous healthcare analytics studies indicating that machine learning can support predictive analysis and operational optimization in healthcare environments.

CONCLUSIONS AND RECOMMENDATIONS

This study successfully implemented a machine learning approach for predicting patient no-show in hospital appointment systems using the Decision Tree algorithm. The research results demonstrate that hospital appointment data can be utilized to identify patient attendance behavior patterns and support predictive analytics in healthcare management systems.

The exploratory data analysis revealed that patient no-show represents a considerable portion of the appointment dataset, indicating that no-show behavior is a relevant operational issue in hospital services. The analysis also showed that appointment schedule characteristics, including doctor assignment,

appointment day, and appointment session, may influence patient attendance behavior.

The Decision Tree model achieved an accuracy of 60%, indicating moderate classification performance in predicting patient no-show behavior. Although the model performance was relatively moderate, the findings demonstrate that machine learning has the potential to support hospital operational management by identifying potential no-show patients and improving appointment scheduling efficiency.

This study recommends that future research should utilize larger datasets and include additional variables such as patient demographics, payment methods, medical history, and previous appointment behavior to improve prediction performance. Future studies are also encouraged to compare multiple machine learning algorithms and integrate predictive models into Hospital management Information Systems (HMIS) and healthcare decision support systems.

Advanced Research

This study has several limitations that may influence the performance of the machine learning model. The dataset used in this research was relatively small and contained limited appointment-related variables. As a result, the Decision Tree model achieved moderate prediction performance. Future research is recommended to utilize larger and more diverse healthcare datasets in order to improve model generalization and prediction accuracy.

Further studies may also include additional variables such as patient demographics, waiting time, insurance type, payment methods, medical history, and previous appointment attendance behavior. The inclusion of these variables may provide a deeper understanding of patient no-show patterns and improve predictive performance.

In addition, future research may compare multiple machine learning algorithms such as Random Forest, XGBoost, Support Vector Machine, and Neural Networks to identify the most effective prediction model for healthcare appointment systems. Advanced healthcare analytics approaches, including deep learning and ensemble learning methods, may also be explored.

Future studies are also encouraged to integrate machine learning prediction models into Hospital Management Information Systems (HMIS) and healthcare dashboards to support real-time operational decision-making. Moreover, graph database technologies such as Neo4j and knowledge graph approaches may be implemented to analyze relationships among patients, physicians, appointments, and healthcare services in a more comprehensive manner.

Acknowledgment

The authors would like to express their sincere gratitude to Politeknik Negeri Banjarmasin for providing academic support and research facilities during this study. The authors also appreciate all parties who contributed to the completion of this research, particularly in supporting data analysis and research development in the field of healthcare information systems and machine learning applications.

REFERENCES

Aghaeifar, R., Servis, G., & Khasawneh, M. T. (2023). *Ensemble Learning for Addressing Class Imbalance in Cardiology Appointment Scheduling and Overbooking*. https://doi.org/10.21203/rs.3.rs-3359966/v1

Aldali, M. (2024). Artificial Intelligence Applications in Healthcare. In *Alqalam Journal of Medical and Applied Sciences*. https://doi.org/10.54361/ajmas.247323

Anjum, N., Kiran, M. S., Jabeen, F., Yang, Z., Huang, C., Noor, S., Imran, K., Ali, I., & Mohamed, E. M. (2021). *Intelligent COVID-19 Forecasting, Diagnoses and Monitoring Systems: A Survey*. https://doi.org/10.36227/techrxiv.15172488.v1

Barrera, D., Brailsford, S., Bravo, C., & Smith, H. (2020). Improving Healthcare Access Management by Predicting Patient No-Show Behaviour. In *Decision Support Systems*. https://doi.org/10.1016/j.dss.2020.113398

Berquedich, M., Kamach, O., Masmoudi, M., & Deshayes, L. (2020). An Immune Memory and Negative Selection to Visualizing Clinical Pathways From Electronic Health Record Data. In *Indonesian Journal of Electrical Engineering and Computer Science*. https://doi.org/10.11591/ijeecs.v19.i1.pp336-343

Bodria, F., Giannotti, F., Guidotti, R., Naretto, F., Pedreschi, D., & Rinzivillo, S. (2023). Benchmarking and Survey of Explanation Methods for Black Box Models. In *Data Mining and Knowledge Discovery*. https://doi.org/10.1007/s10618-023-00933-9

Hartmann, M., Ling, S., Exarchakou, A., Rachet, B., & Belot, A. (2024). *Predicting the Probabilities of Missed General Practice Appointments in England and Wales*. https://doi.org/10.21203/rs.3.rs-3836849/v1

Hassan, S. A. Z. (2024). An AI Healthcare Ecosystem Framework for Covid-19 Detection and Forecasting Using CronaSona. In *Medical \& Biological Engineering \& Computing*. https://doi.org/10.1007/s11517-024-03058-3

Liu, J. (2022). Deconstruction and Implementation of Strategic Human Resource Management Evaluation Algorithm Using Data Mining Technology. In *Mobile Information Systems*. https://doi.org/10.1155/2022/8890859

Nagahisarchoghaei, M., Nur, N., Cummins, L., Nur, N., Karimi, M. M., Nandanwar, S., Bhattacharyya, S., & Rahimi, S. (2023). An Empirical Survey on Explainable AI Technologies: Recent Trends, Use-Cases, and Categories From Technical and Application Perspectives. In *Electronics*. https://doi.org/10.3390/electronics12051092

Olamijuwon, J., & Zouo, S. J. C. (2024). The Impact of Health Analytics on Reducing Healthcare Costs in Aging Populations: A Review. In *International Journal of Management \& Entrepreneurship Research*. https://doi.org/10.51594/ijmer.v6i11.1690

Sarla, G. S. (2024). Challenges Faced by a Hospital Administrator. In *The Egyptian Journal of Internal Medicine*. https://doi.org/10.1186/s43162-024-00312-w

Shiyanbade, B. W. (2024). Administrative Structures Sustainability and Secondary Health Care Administration in Lagos, Nigeria. In *Global Journal of Social Sciences*. https://doi.org/10.4314/gjss.v23i1.5

Tenepalli, D., & Navamani, T. M. (2024). A Systematic Review on IoT and

Machine Learning Algorithms in E-Healthcare. In *International Journal of Computing and Digital Systems*. https://doi.org/10.12785/ijcds/160122

Van, N. T. T., Vrana, V., Duy, N. T., Minh, D. X. H., Dzung, P. T., Mondal, S. R., & Das, S. (2020). The Role of Human–Machine Interactive Devices for Post-Covid-19 Innovative Sustainable Tourism in Ho Chi Minh City, Vietnam. In *Sustainability*. https://doi.org/10.3390/su12229523