DEVELOPMENT OF AN ENTERPRISE FORECASTING INFORMATION SYSTEM FOR DAILY PASSENGER DEMAND PREDICTION ON THE YOGYAKARTA–SOLO COMMUTER RAIL LINE USING LSTM

Marta Ardiyanto[1*], Ridwan Dwi Irawan[2], Esti Dwi Rahmawati[3], Muhammad Ilham 'Aziiz Alfarobi[4]
Universitas Duta Bangsa Surakarta[1,2,3,4]
*Correspondence Email : marta.ardiyanto@udb.ac.id

ABSTRACT

This study aims to develop an enterprise forecasting information system for daily passenger demand prediction on the Yogyakarta–Solo Commuter Rail Line using Long Short-Term Memory. The proposed system integrates passenger demand forecasting and dashboard-based decision support to assist operational monitoring. The research used an experimental and system development approach. Daily passenger volume data were processed through data cleaning, chronological ordering, historical feature engineering, normalization, and sequence formation. Three LSTM feature scenarios were evaluated: historical features, historical and calendar features, and historical, calendar, and tourism event features. The forecasting model was then integrated into a web-based dashboard that provides historical trends, forecasting results, actual versus predicted visualization, model evaluation metrics, and downloadable prediction reports. The results show that the LSTM model with historical features achieved the best performance, with an MAE of 3,713.15, RMSE of 4,281.68, and MAPE of 14.57%. The system testing results indicate that the main dashboard functions run according to the expected outputs. Therefore, the proposed system can support data-driven passenger demand monitoring and operational decision-making for commuter rail services.

KEYWORDS

Enterprise forecasting information system; LSTM; passenger demand prediction; predictive dashboard; Yogyakarta–Solo Commuter Rail

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International

INTRODUCTION

The Yogyakarta–Solo Commuter Rail Line is one of the strategic public transportation modes connecting the Special Region of Yogyakarta and Central Java. This service supports community mobility for education, work, trade, tourism, and intercity travel. In the context of rail-based public transportation, passenger volume prediction is an important aspect because forecasting results can be used to support operational planning, service capacity management, schedule arrangement, and service quality improvement. Daily and weekly passenger demand forecasting in urban rail transit plays a role in helping station authorities diagnose operational problems and develop service scheduling patterns (Chuwang & Chen, 2022). In addition, accurate passenger flow prediction is important to support operational management, service control, and the reliability improvement of urban rail transit systems (Zhu et al., 2024).

Passenger volume on the Yogyakarta–Solo Commuter Rail Line has fluctuating characteristics because community movement does not always follow a constant pattern. The number of passengers may change according to weekdays, weekends, national holidays, collective leave days, school holidays, and tourism or cultural activities along the Yogyakarta–Solo corridor. This condition indicates that passenger prediction should not only consider historical data, but also calendar factors and certain events that may influence community mobility. Calendar components, seasonality, holidays, and specific events are important factors in time-series modeling because they can affect demand patterns in certain periods (Taylor & Letham, 2018). In addition, external variables and temporal contexts can be used to enrich public transportation demand prediction models so that they become more representative of community mobility dynamics (Alexandre et al., 2023; Rocco di Torrepadula et al., 2024; Zhu et al., 2024).

Various approaches have been used in passenger prediction studies, ranging from statistical models and machine learning to deep learning. Machine learning and deep learning approaches are widely used in public transportation demand prediction because they are able to capture nonlinear and temporal patterns in travel data (Alexandre et al., 2023; Rocco di Torrepadula et al., 2024; Zhu et al., 2024). Hou et al. (2022) developed a combined TCN-LSTM model to predict passenger flow in urban rail transit by considering nonlinear characteristics, correlations, and periodicity in the data. Zheng et al. (2022) proposed a graph-temporal fused deep learning approach to predict origin-destination passenger flows in urban rail transit. The LSTM model is also relevant for time-series data because it can learn sequential patterns and long-term dependencies in data (Hochreiter & Schmidhuber, 1997; Siswanto et al., 2024; Zhang & Zhang, 2024). With these characteristics, LSTM can be used to model daily passenger patterns influenced by historical values, weekly patterns, and calendar variations.

However, most previous studies have focused more on improving algorithm performance or comparing prediction models. Forecasting results often stop at evaluation values such as MAE, RMSE, or MAPE, and have not been fully integrated into an information system that can support monitoring and decision-making. In an implementation context, prediction model results need to be visualized through a dashboard so that they can be more easily understood by non-technical users. Dashboards support the presentation of analytical information through structured, concise, interactive visualizations that facilitate data interpretation (Bach et al., 2023). In addition, machine

learning-based public transportation prediction systems can support demand analysis, operational planning, and data-driven decision-making (Alexandre et al., 2023; Zhu et al., 2024). Therefore, integrating prediction models into an enterprise information system is important so that forecasting results do not remain merely numerical outputs, but can also serve as a basis for decision support.

Based on these problems, this study proposes the development of a daily passenger prediction dashboard for the Yogyakarta–Solo Commuter Rail Line using Long Short-Term Memory with the integration of calendar features. The calendar features used include day of week, weekend indicator, public holiday indicator, collective leave indicator, tourism event indicator, as well as historical features such as lag passenger volume and rolling average. The LSTM model is used as the forecasting engine, while the dashboard is developed to display historical data, prediction results, actual and predicted comparisons, evaluation metrics, and relevant calendar information. Thus, this study does not only focus on building a prediction model, but also on integrating the model into an interactive dashboard as a decision support system.

The main contribution of this study is the development of an LSTM-based daily passenger prediction information system for the Yogyakarta–Solo Commuter Rail Line with the integration of calendar features. This study contributes to two aspects. First, from the predictive modeling aspect, this study develops an LSTM model based on historical and calendar features to estimate daily passenger volume. Second, from the enterprise information system aspect, this study integrates the prediction model into a web-based system that includes data management, passenger trend visualization, prediction results, model evaluation, prediction reports, and support for operational decision-making in public transportation.

RESEARCH METHOD

This study uses an experimental and system development approach to develop an LSTM-based daily passenger prediction information system for the Yogyakarta–Solo Commuter Rail Line. The experimental approach is used to build and evaluate the prediction model, while the system development approach is used to integrate the prediction model into an interactive dashboard as part of an enterprise information system. The system development process refers to the principles of the System Development Life Cycle, as SDLC provides systematic stages from requirement analysis, design, implementation, testing, to system evaluation (Maryani et al., 2022). The main focus of this study is to develop a predictive dashboard that not only displays historical data, but also presents prediction results, evaluation metrics, and calendar-related information that can be used as a basis for operational decision support.

The first stage is data collection. The main data used in this study consist of historical daily passenger volume data for the Yogyakarta–Solo Commuter Rail Line. Each observation represents the number of passengers on a specific date. In addition to the main data, this study also uses supporting data, including the national calendar, weekday and weekend information, national holidays, collective leave days, and tourism events or local activities taking place along the Yogyakarta–Solo corridor. These supporting data are used because public transportation passenger volume can be influenced by temporal and external factors, such as holidays, special events, and tourism activities. External variables and temporal contexts can enrich public transportation demand prediction models so that they

become more representative of community mobility dynamics (Alexandre et al., 2023; Rocco di Torrepadula et al., 2024; Zhu et al., 2024).

The second stage is data preprocessing. At this stage, the date attribute is converted into a standard format, and the data are then sorted chronologically based on the observation date. The data are also checked for missing values, duplicate dates, and format inconsistencies. If missing values are found, they are handled according to the characteristics of the data, such as by revalidating the original data source or applying an appropriate imputation method. After the data are cleaned, passenger volume data are merged with calendar and event data based on the date attribute. This process produces an integrated dataset containing passenger volume information, calendar characteristics, and event indicators.

The third stage is feature engineering. The constructed features consist of calendar features and historical features. The calendar features include day of week, month, week of year, is weekend, is public holiday, is collective leave, and tourism event indicator. The historical features include lag 1 day, lag 7 days, rolling mean 7 days, and rolling mean 14 days. Lag features are used to capture the influence of passenger volume in previous periods, while rolling mean features are used to capture short-term trends and weekly patterns. The use of historical and calendar features is relevant to the characteristics of time-series data, which often contain temporal patterns, seasonality, and holiday effects. Trend, seasonality, and holiday effects are important components in time-series modeling because they help the model recognize changes in data patterns across specific time periods (Taylor & Letham, 2018).

The fourth stage is normalization and sequence formation. Since LSTM works with sequential patterns, the dataset is transformed into sequences based on a specific lookback window. In this study, data from several previous days are used as input to predict passenger volume for the following day. Before being used in the model, numerical variables are normalized using Min-Max Scaling so that they are transformed into a uniform value range. Normalization is applied to make the model training process more stable because deep learning models are sensitive to differences in data scale. The use of LSTM in public transportation prediction is relevant because this model can learn sequential patterns and long-term dependencies in time-series data (Hochreiter & Schmidhuber, 1997; Siswanto et al., 2024; Zhang & Zhang, 2024).

The fifth stage is the splitting of training and testing data. The data are split chronologically to align with the characteristics of time-series data. Random splitting is not used because it may cause information leakage from future data into past data. Data from the earlier period are used as training data, while data from the later period are used as testing data. This strategy is used to simulate real forecasting conditions, where the model is trained using historical data and then used to predict subsequent periods. Chronological data splitting is important in time-series prediction because the evaluation process must preserve temporal order so that the testing results represent actual forecasting conditions (Rocco di Torrepadula et al., 2024; Zhu et al., 2024).

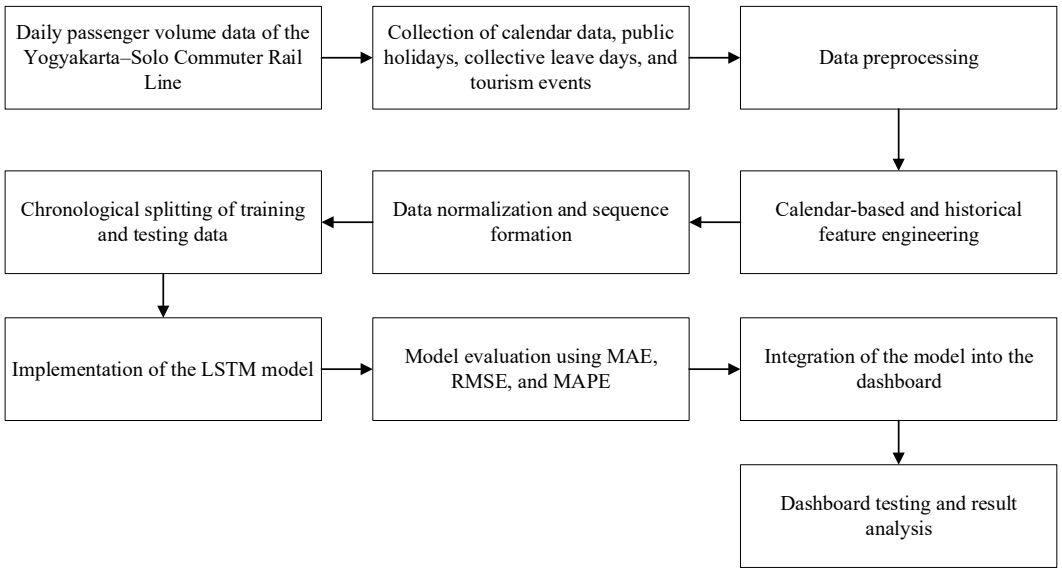
The sixth stage is the implementation of the LSTM model. The LSTM model is built using sequential input consisting of historical and calendar features. The model architecture consists of an LSTM layer to learn temporal patterns, a dropout layer to reduce the risk of overfitting, and a dense layer as the prediction output. The model is trained using the Adam optimizer and the Mean Squared Error loss function. LSTM is selected because of its ability to capture long-term dependencies in sequential data, making it suitable for predicting daily passenger volume that contains temporal patterns and calendar variations (Hochreiter & Schmidhuber, 1997; Siswanto et al., 2024; Zhang & Zhang, 2024).

The seventh stage is model evaluation. The evaluation is conducted using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). MAE is used to measure the average absolute error between actual and predicted values. RMSE is used to measure prediction error by giving a higher penalty to larger errors. MAPE is used to measure prediction error in percentage form, making it easier to interpret. The selection of evaluation metrics must be considered carefully because each metric has different characteristics in measuring forecasting accuracy (Hyndman & Koehler, 2006).

The eighth stage is dashboard development. The dashboard is designed to integrate the LSTM model results into an easy-to-use visual interface. The dashboard modules include passenger data overview, historical trend visualization, prediction results for the next 7 to 30 days, actual versus predicted comparison charts, holiday and event information, model evaluation metrics, and prediction result export features. The dashboard is also designed to support analysis so that users can observe changes in passenger volume based on temporal patterns and calendar contexts. Dashboards support the presentation of analytical information through structured, concise, interactive visualizations that facilitate data interpretation (Bach et al., 2023). Therefore, the dashboard functions not only as a visualization medium, but also as a predictive analytics-based decision support system.

The final stage is dashboard testing. Testing is conducted to ensure that each dashboard function operates according to system requirements. The testing process includes data reading, historical data visualization, prediction model execution, prediction result display, evaluation metric display, period-based filtering, and prediction result export. In addition to functional testing, result interpretation is also conducted by comparing actual and predicted charts to examine the extent to which the model can follow passenger volume fluctuation patterns. The final output of this study is an LSTM-based daily passenger prediction information system for the Yogyakarta–Solo Commuter Rail Line that can support public transportation monitoring and operational decision-making.

Figure 1. Research Method Flowchart



RESULT AND DISCUSSION

This section presents the results of developing an LSTM-based daily passenger prediction information system for the Yogyakarta–Solo Commuter Rail Line with the integration of calendar features. The discussion is organized into six main parts: passenger volume pattern analysis, LSTM model evaluation results, SDLC-based system requirement analysis, use case design, predictive dashboard implementation, and system testing. With this structure, the research results not only present the performance of the prediction model, but also demonstrate how the prediction model is integrated into an enterprise information system as a decision support tool for public transportation operations.

1. Passenger Volume Pattern Analysis

An initial analysis was conducted on the historical daily passenger volume data of the Yogyakarta–Solo Commuter Rail Line to identify general passenger movement patterns. Based on the time-series data visualization, passenger volume shows a fluctuating pattern from day to day. This fluctuation indicates that the number of passengers is not only influenced by historical values from previous periods, but also by calendar characteristics, such as weekdays, weekends, national holidays, collective leave days, and tourism events or certain local activities. This condition is consistent with the characteristics of public transportation data, which often contain temporal patterns, seasonality, and external factors (Alexandre et al., 2023; Zhu et al., 2024).

Based on day category grouping, differences can be observed between weekdays and weekends. The average passenger volume on weekends tends to differ from weekdays because travel characteristics on weekends are more frequently associated with recreational, family, trade, and tourism activities. In addition, national holidays and collective leave days may also affect passenger volume due to increased community mobility during holiday periods. Therefore, features such as day of week, is weekend, is public holiday, and is collective leave are used as supporting variables in the prediction model.

In addition to calendar factors, tourism events or local activities along the Yogyakarta–Solo corridor may also influence passenger volume. Cultural events, festivals, school holidays, and tourism-related activities can increase community mobility toward certain activity centers. These conditions indicate that external factors should be considered in passenger demand prediction because they can affect temporal and seasonal mobility patterns in public transportation (Alexandre et al., 2023; Rocco di Torrepadula et al., 2024; Zhu et al., 2024).

Table 1 presents the summary statistics of passenger volume based on day category.

Table 1. Average Passenger Volume Based on Day Category

Day Category	Number of Data	Average Passenger Volume	Minimum	Maximum
Weekday	261	23,328	14,922	45,536
Weekend	104	28,146	14,888	42,209
Public Holiday	17	29,758	15,909	41,488
Collective Leave	11	32,695	18,410	45,536
Tourism Event	179	25,960	17,595	45,536

Based on Table 1, it can be observed that certain day categories have different average passenger volumes. If the average passenger volume on weekends, public holidays, or tourism event days is higher than on weekdays, this indicates a potential influence of calendar factors and events on the increase in passenger numbers. Conversely, if a certain category has a lower average value, it may indicate that not all holidays or events significantly increase passenger volume.

2. LSTM Model Evaluation Results

Long Short-Term Memory is used as the main forecasting engine in the developed system. This model is selected because it is capable of learning sequential patterns in time-series data and capturing temporal dependencies from historical passenger data. In this study, the model input is constructed from a combination of historical and calendar features, such as lag passenger volume, rolling mean, day of week, weekend indicator, public holiday indicator, collective leave indicator, and tourism event indicator.

Model evaluation is conducted using Mean Absolute Error, Root Mean Square Error, and Mean Absolute Percentage Error. MAE is used to measure the average absolute error between actual and predicted values. RMSE is used to measure the magnitude of prediction error by giving a higher penalty to larger errors. MAPE is used to evaluate prediction error in percentage form, making it easier to interpret in an operational context.

Tabel 2. LSTM Forecasting Model Evaluation Result

Model	Feature Scenario	MAE	RMSE	MAPE
LSTM	Historical Features	3,713.15	4,281.68	14.57 %
LSTM	Historical + Calendar Features	4,505.45	6,373.97	17.27%
LSTM	Historical + Calendar + Tourism Event Features	4,798.52	6,283.41	19.11%

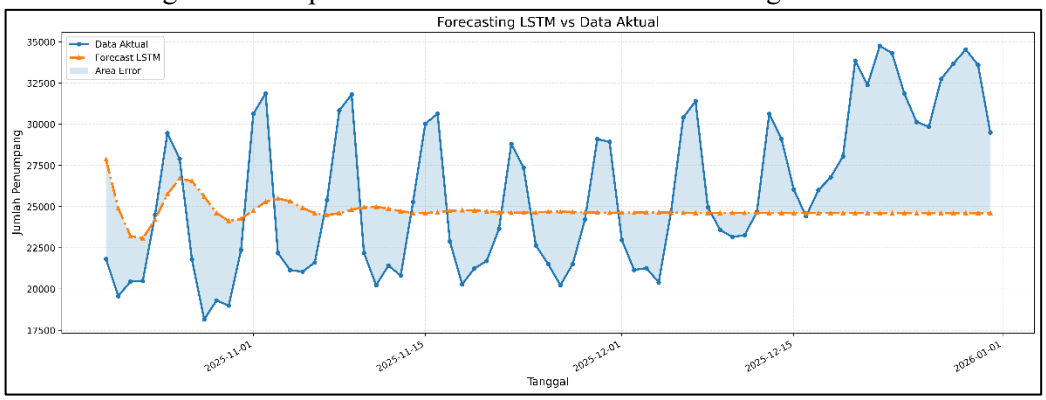
Based on Table 2, the scenario with the lowest MAE, RMSE, and MAPE values indicates the best prediction performance. The evaluation results show that the LSTM model with the Historical Features scenario achieved the best performance, with an MAE of 3,713.15, RMSE of 4,281.68, and MAPE of 14.57%. The Historical + Calendar Features scenario produced an MAE of 4,505.45, RMSE of 6,373.97, and MAPE of 17.27%, while the Historical + Calendar + Tourism Event Features scenario produced an MAE of 4,798.52, RMSE of 6,283.41, and MAPE of 19.11%. These results indicate that historical features, such as lag and rolling mean, are more dominant in helping the LSTM model recognize daily passenger volume patterns on the Yogyakarta–Solo Commuter Rail Line.

The addition of calendar and tourism event features in this experiment did not reduce the prediction error. This condition indicates that the calendar and event features used in this study still need to be designed more specifically in order to provide stronger contributions to the model. Tourism event features, for example, have not yet been differentiated based on event scale, location, duration, proximity to stations, and potential visitor volume. Therefore, these results indicate that the integration of calendar and event features remains relevant as part of a predictive information system, but the quality of feature representation needs to be improved to provide a more significant impact on prediction accuracy.

In addition to numerical evaluation, the prediction results are also analyzed through a comparison chart between actual and predicted values. The actual and predicted comparison chart is used to examine the model's ability to follow daily passenger volume fluctuation patterns. If the prediction line follows the direction of the actual data movement,

the model can be considered reasonably capable of recognizing general passenger demand patterns. However, if large differences occur on certain dates, these conditions may be associated with external factors that are not fully captured in the dataset, such as large-scale events, operational disruptions, or changes in community travel behavior.

Figure 2. Comparison of Actual and Predicted Passenger Volume



Based on Figure 2, the LSTM model can be analyzed by examining the closeness of the pattern between the actual data and the predicted results. Differences in prediction values during certain periods can serve as input for improving the forecasting system by continuously updating it with more recent data and more complete external variables. Therefore, model evaluation does not only function to measure the level of prediction accuracy, but also serves as a basis for developing a forecasting system that is more adaptive to changes in community mobility patterns.

3. Enterprise System Requirement Analysis Based on SDLC

The development of the passenger forecasting information system was carried out using the System Development Life Cycle approach. In the requirement analysis stage, the system was designed to meet user needs in managing passenger data, processing calendar features, running the LSTM prediction model, displaying prediction results, and providing forecasting reports as a basis for decision-making. Therefore, the developed system does not only function as a visualization dashboard, but also as an enterprise information system based on predictive analytics.

The requirement analysis identified three main actors: Administrator, Operator or Analyst, and Decision Maker or Manager. The Administrator is responsible for managing users, passenger data, calendar data, public holidays, collective leave days, and tourism events. The Operator or Analyst is responsible for data input, preprocessing, feature construction, prediction model execution, and model evaluation. Meanwhile, the Decision Maker or Manager uses the dashboard to monitor passenger trends, view prediction results, review evaluation metrics, and download prediction reports to support operational decision-making.

Table 3. User Requirement of the Enterprise Forecasting Information System

Actor	Main Requirement	System Function
Administrator	Managing system master data	User management, passenger data management, calendar and event data management

Operator/Analyst	Processing data and running prediction	Data preprocessing, feature engineering, LSTM forecasting, model evaluation
Decision Maker/Manager	Monitoring and decision support	Dashboard monitoring, actual vs predicted analysis, report export

Based on Table 3, the system requirements indicate that the passenger forecasting system must be able to support business processes ranging from data management and prediction processing to the presentation of forecasting results. These requirements show that the developed system has enterprise characteristics because it involves multiple actors, data management, analytical processing, and information outputs that can be used to support operational decisions.

4. Use Case Design as an Interpretation of Information System Development

The use case diagram was developed as part of the analysis and system design stages in the SDLC approach (Rospricilia & Ma'ady, 2024). The use case is used to describe the interactions between actors and the main functions of the system. In this study, the use case serves as an interpretation of the requirements for an enterprise information system that integrates historical data, calendar features, the LSTM prediction model, an analytical dashboard, and forecasting reports.

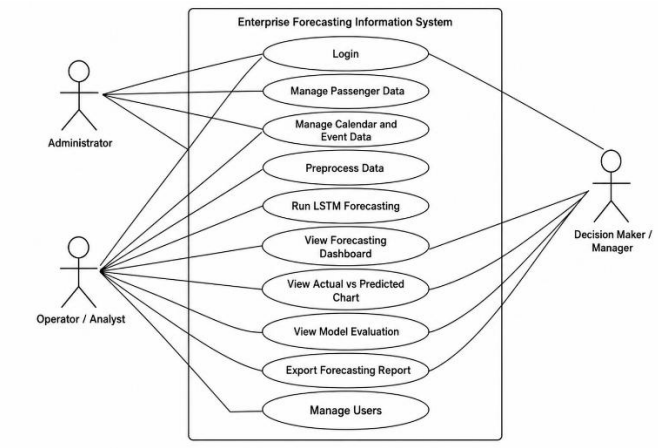


Figure 3. Use Case Diagram of the Enterprise Forecasting Information System

The use case diagram in Figure 3 shows that the system has several main functions, including login, passenger data management, calendar and event data management, data preprocessing, LSTM prediction execution, forecasting dashboard viewing, actual and predicted chart viewing, model evaluation viewing, forecasting report downloading, and user management. Each function is connected to the relevant actors according to their access rights and responsibilities.

Based on Figure 3, the use case shows that the developed system does not only provide prediction functions, but also includes data management, data processing, result visualization, model evaluation, and reporting. This strengthens the position of this study as the development of an enterprise information system, rather than merely the application of the LSTM model to time-series data. The use case also shows that the prediction results are directed toward supporting decision-making processes, particularly in monitoring passenger volume and planning the operations of the Yogyakarta–Solo Commuter Rail Line service.

5. Implementation of the Enterprise Forecasting Dashboard

The implementation stage produced a passenger forecasting dashboard that integrates the LSTM model into a web-based system. The dashboard was designed to enable users to view information in a concise, interactive, and easily understandable manner. The main modules developed include the dashboard overview, historical trend visualization, passenger volume prediction, actual versus predicted comparison, calendar and event information, model evaluation, and prediction report export.

The dashboard overview presents key information, such as the total number of passengers, average daily passenger volume, highest passenger volume, lowest passenger volume, and passenger trends within a specific period. The historical trend visualization module is used to display passenger volume movements over time. The prediction module presents forecasting results for a certain period, such as the next 7, 14, or 30 days. The actual versus predicted module is used to compare actual values with the prediction results generated by the LSTM model. The model evaluation module displays MAE, RMSE, and MAPE values as indicators of forecasting performance.

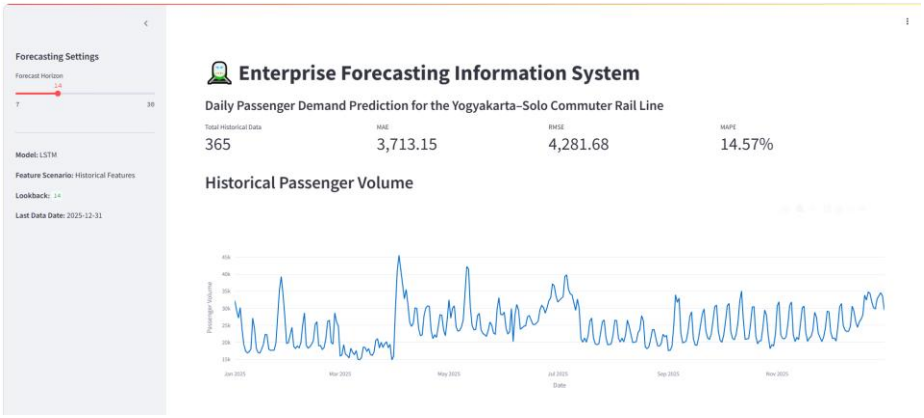


Figure 4. Dashboard Overview of Passenger Forecasting System



Figure 5. Forecasting Result Visualization

The dashboard implementation shows that LSTM prediction results can be presented in a visual form that is easier for users to understand. For operators, the dashboard helps monitor model performance and review prediction results. For decision makers, the

dashboard provides predictive information that can be used to anticipate potential increases or decreases in passenger volume. Therefore, the dashboard functions as both a visualization medium and a decision support component within the enterprise system.

6. System Testing

System testing was conducted to ensure that each function operates according to the requirements defined during the analysis stage. The testing was carried out using a black-box testing approach by examining the input, process, and output of each main module (Gustinov et al., 2023). The tested modules include login, passenger data management, calendar and event data management, LSTM prediction processing, dashboard visualization, model evaluation display, and prediction report export.

Table 4. Black-Box Testing Result

Test Scenario	Expected Result	Status
User logs into the system using valid account	The system displays the dashboard page according to user role	Valid
Administrator uploads passenger data	The system stores and displays passenger data correctly	Valid
Administrator manages calendar and event data	The system stores holiday, collective leave, and event data correctly	Valid
Operator runs preprocessing process	The system generates clean and structured dataset	Valid
Operator runs LSTM forecasting	The system displays passenger volume prediction results	Valid
User views actual versus predicted chart	The system displays comparison chart correctly	Valid
User views model evaluation	The system displays MAE, RMSE, and MAPE values	Valid
User exports prediction report	The system generates downloadable report file	Valid

Based on Table 4, all main system functions can be evaluated based on the conformity between the input and the expected output. The testing results show that the system can perform its main processes, ranging from data management and prediction processing to result visualization and report export. Therefore, the developed system has fulfilled the functional requirements as an enterprise forecasting information system.

7. Discussion

The research results show that the integration of the LSTM model and the enterprise information system can support daily passenger volume prediction on the Yogyakarta–Solo Commuter Rail Line. From the modeling perspective, the use of historical and calendar features helps the model recognize temporal patterns, weekly patterns, and variations related to holidays and certain events. The evaluation results using MAE, RMSE, and MAPE provide an overview of the prediction error produced by the model.

From the information system development perspective, the results of the use case design, dashboard implementation, and system testing show that the prediction model can be integrated into a more applicable system. The developed system does not only display prediction values, but also provides trend visualization, actual and predicted charts, calendar information, evaluation metrics, and prediction reports. Therefore, the prediction results become easier to understand and can be utilized by users to support operational decision-making.

This system development also shows that the SDLC approach is relevant for predictive analytics-based research. The requirement analysis stage produces the identification of actors and system functions. The design stage produces the use case diagram and system

module design. The implementation stage produces an LSTM-based prediction dashboard. The testing stage ensures that each main function operates according to system requirements. Through this workflow, this study contributes not only to the prediction model aspect, but also to the development of an enterprise information system that supports data-driven public transportation management.

Overall, this study shows that passenger volume prediction for the Yogyakarta–Solo Commuter Rail Line becomes more useful when it is integrated into an information system that can be used by stakeholders. The LSTM model functions as the forecasting engine, while the enterprise system functions as a medium for data management, result visualization, performance evaluation, and information provision for decision-making. Therefore, the proposed approach can serve as a basis for developing a passenger forecasting system that is more adaptive, integrated, and supportive of public transportation operational needs.

CONCLUSION

This study developed an enterprise forecasting information system for daily passenger demand prediction on the Yogyakarta–Solo Commuter Rail Line using the Long Short-Term Memory model. The system was designed not only to generate passenger volume predictions, but also to integrate forecasting results into a web-based dashboard that supports monitoring, visualization, evaluation, reporting, and operational decision-making. The development process followed the SDLC approach, including requirement analysis, system design, implementation, and functional testing.

The experimental results show that the LSTM model with Historical Features produced the best forecasting performance, with an MAE of 3,713.15, RMSE of 4,281.68, and MAPE of 14.57%. The addition of calendar and tourism event features did not reduce the prediction error in this experiment, indicating that historical patterns such as lag and rolling mean were more dominant in explaining daily passenger volume. However, calendar and event features remain relevant for system interpretation and can be improved in future studies by considering event scale, location, duration, proximity to stations, and potential visitor impact.

The system implementation results show that the forecasting model can be integrated into an enterprise dashboard that presents historical passenger trends, forecasting results, actual versus predicted visualization, model evaluation metrics, and downloadable reports. Black-box testing indicates that the main system functions operate according to the expected outputs. Future research can extend this work by using multi-year datasets, improving event-based feature representation, adding external variables such as weather or service disruption data, and developing more adaptive forecasting models for public transportation planning.

REFERENCES

Alexandre, T., Bernardini, F., Viterbo, J., & Pantoja, C. E. (2023). Machine learning applied to public transportation by bus: A systematic literature review. *Transportation Research Record: Journal of the Transportation Research Board, 2677*(7), 639–660. doi: 10.1177/03611981231155189

Bach, B., Freeman, E., Abdul-Rahman, A., Turkay, C., Khan, S., Fan, Y., & Chen, M. (2023). Dashboard design patterns. *IEEE Transactions on Visualization and Computer Graphics, 29*(1), 342–352. doi: 10.1109/TVCG.2022.3209448

Chuwang, D. D., & Chen, W. (2022). Forecasting daily and weekly passenger demand for urban rail transit stations based on a time series model approach. *Forecasting, 4*(4), 904–924. doi: 10.3390/forecast4040049

Gustinov, M. D., Azani, N. W., Al Ghani, R., Auliani, S. N., Maharani, S., Hamzah, M. L., & Rizki, M. (2023). Analysis of web-based e-commerce testing using black box and white box methods. *International Journal of Information System and Innovation Management, 1*(1), 20–31. doi: 10.55583/ijisim.v1i1.687

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. doi: 10.1162/neco.1997.9.8.1735

Hou, Z., Du, Z., Yang, G., & Yang, Z. (2022). Short-term passenger flow prediction of urban rail transit based on a combined deep learning model. *Applied Sciences, 12*(15), 7597. doi: 10.3390/app12157597

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688. doi: 10.1016/j.ijforecast.2006.03.001

Maryani, Prabowo, H., Gaol, F. L., & Hidayanto, A. N. (2022). Comparison of the system development life cycle and prototype model for software engineering. *International Journal of Emerging Technology and Advanced Engineering, 12*(4), 155–162. doi: 10.46338/ijetae0422_19

Rocco di Torrepadula, F., Napolitano, E. V., Di Martino, S., & Mazzocca, N. (2024). Machine learning for public transportation demand prediction: A systematic literature review. *Engineering Applications of Artificial Intelligence, 137*, 109166. doi: 10.1016/j.engappai.2024.109166

Rospricilia, T. A., & Ma'ady, M. N. P. (2024). Pemodelan integration use case (IUC): Perancangan use case diagram (UML) untuk sistem-sistem yang terintegrasi. *INTEGER: Journal of Information Technology, 9*(2), 165–172. doi: 10.31284/j.integer.2024.v9i2.6345

Siswanto, J., Manongga, D., Sembiring, I., & Wijono, S. (2024). Deep learning based LSTM model for predicting the number of passengers for public transport bus operators. *Jurnal Online Informatika, 9*(1), 61–70. doi: 10.15575/join.v9i1.1245

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician, 72*(1), 37–45. doi: 10.1080/00031305.2017.1380080

Zhang, J., & Zhang, J. (2024). Enhancing intelligent mobility: LSTM neural networks for short-term passenger flow prediction in rail transit. In *Proceedings of the 1st International Conference on Data Science and Engineering* (pp. 410–415). SciTePress. doi: 10.5220/0012797600004547

Zheng, H., Chen, J., Huang, Z., Yang, K., & Zhu, J. (2022). Short-term online forecasting for passenger origin–destination flows of urban rail transit: A graph-temporal fused deep learning method. *Mathematics, 10*(19), 3664. doi: 10.3390/math10193664

Zhu, L., Chen, C., Wang, H., Yu, F. R., & Tang, T. (2024). Machine learning in urban rail transit systems: A survey. *IEEE Transactions on Intelligent Transportation Systems, 25*(3), 2182–2207. doi: 10.1109/TITS.2023.3319135