NEF6002 Research Project B

AI Applied to Public Transportation: Using Deep Learn Models to Predict Train Passenger Demand at Central Station (Sydney-NSW)

Student Name: Walysson Tabosa dos Santos

Student ID: s4682287

A minor thesis in partial fulfillment of the requirements for the degree of

Master of Applied Information Technology

IT Discipline, College of Arts, Business, Law, Education and IT

Victoria University Sydney, Australia

June, 2024

Abstract

In this paper, we investigate the application of deep learning models, including BiLSTM and LSTM architectures, in predicting passenger requests on New South Wales transport system in Australia. The purpose of the research is to utilize these models in transportation preparation and analyse regular and temporal patterns in passenger requests. The data used in the experiment is gathered from Australian openData.gov, related to the number of entries at Central Station (Sydney), from June/2026 to April/2024. More precisely, the objective is to build a highly accurate forecast model appropriate for analysing the dynamic attributes of urban movement to allow transportation planners and managers to take action and deliver employment as efficiently as achievable. Therefore, with a systematic research method of data collection, processing, model development, simulation, and evaluation, the research reveals that deep learning methods are efficient for short-term passenger demand prediction with predictive intelligence accuracies of more than 85%. The outcomes indicate how essential it is to incorporate temporal and spatial possibilities into prediction models to develop reliable and dependable predictions. In conclusion, this research contributes to our knowledge of transport science and engineering, provides valuable information to direct empiricism-based policymaking, and enhances the efficiency and sustainability of NSW public transportation systems.

Keywords: Deep learning, BiLSTM, LSTM, Passenger demand prediction, Public transportation, Urban mobility, Transportation planning, NSW, Australia.

Declaration

I, Walysson Tabosa dos Santos, hereby solemnly declare that this thesis contains no material that has been accepted for the award of any other degree or diploma in any other college, institute, or university and is the result of my own research. To the best of my knowledge, this thesis contains no material previously published or written by another person, except where references has been made in the text of the thesis.

Signature: ____________________

Name: ________________________

Date: ________________________

Acknowledgements

Supervisor and anyone who helped you:

I would like to express my special appreciation and thanks to …etc.

.

Table of Contents

Table of Figures

Glossary

AI	Artificial Intelligence
LSTM	Long Short-Term Memory
BiLSTM	Bidirectional Long Short-Term Memory

1. Introduction

Modern urban mobility is almost impossible without the critical role that public transportation plays in the lives of millions allowing them to commute daily. Accessible, well-organized public transportation services further reduce congestion and help preserve the environment (Li et al., 2018). Nevertheless, the achievement and maintenance of these and other benefits depend on optimal forecasts of passenger demand that allow to allocate resources with accuracy and deliver services accordingly. Conventional tools used in forecasting cannot deal with the complexity and variability of passenger characteristics, making them inefficient.

Recent innovations in deep learning and AI have introduced promising values for enhancing the accuracy of passenger demand forecasting in public transport systems. Deep learning models, such as Bidirectional Long Short-Term Memory and Long Short-Term Memory, have efficiently identified the confusing temporal connections and nonlinear patterns encoded in passenger trip data (Chen et al., 2021). They allow transportation authorities to predict future demand peaks and locations, information that can help them make crucial decisions regarding capacity development, service timetables, and infrastructure investment.

The primary goal of this research is to tackle the serious problem of the accurate prediction of demand peak time and location in public transportation systems using deep learning methods. As far as the researcher is concerned, both BiLSTM and LSTM models have already been shown to be an effective method of passenger demand prediction in previous studies (Veres & Moussa, 2019). At the same time, the current study promotes an innovative approach by applying newly trained models to the datasets obtained from real-time tracking systems. This study will use datasets to both enhance and develop the new forecasting models tailored toward the specifics of the NSW, Australia transportation system.

Therefore, this research will generate critical information regarding the factors that determine the patterns of demand and its peak hour by examining the patterns of all passengers using a different transport mode, such as train, light rail, bus, and ferry. Specifically, this research will investigate whether the BiLSTM and LSTM can enhance the effectiveness of the current models in enhancing service planning and delivery in public transportation.

1.1. Research Significance

By utilising innovative methods for machine learning like BiLSTM (Bidirectional Long Short-Term Memory) and LSTM (Long Short-Term Memory), a predictive model for passenger demand at Sydney, Australia's central station has been developed, which constitutes a significant contribution to both the academic and practical fields. This research not only shows the effectiveness of these models but also illustrates their potential influence on urban transport planning and management by analysing a large dataset that covers June 2016 to April 2024.

The utilisation of BiLSTM and LSTM models for passenger demand forecasting is an innovative approach which improves the current collection of knowledge within the areas of data science, machine learning, and urban planning. The achieved accuracy of 86% suggests that these models are reliable and strong for practical use. For researchers and practitioners looking to improve public transit systems and create smarter cities, this development can be particularly useful.

- **Benefits for Industry:**
    - **Enhanced Management Efficiency:** Transit authorities may optimise train timetables, cut down on wait times, and improve overall quality of service by properly estimating passenger demand. As a result, resources are utilised more effectively, and consumer satisfaction increases.
    - **Cost Control and Planning:** Better resource allocation and budgeting are made possible by accurate demand forecasts. Based on projected usage patterns, transport operators can arrange staff rosters, train frequency, and maintenance schedules, which can result in significant cost savings.

- **Enhanced Decision-Making:** The model offers insightful information that may be used to plan strategically and make decisions. Transport authorities have the ability to make data-driven choices about service enhancements, growth initiatives, and infrastructure expenses.

- **Impact on the Environment**

  - **Decreased Carbon Footprint:** By minimising periods of inactivity and avoiding unnecessary travel, effective scheduling and resource management cut down on fuel use and emissions. By doing this, the transportation network's overall carbon footprint is decreased.
  - **Development of Public transit:** More people choose trains over private automobiles when there are dependable and effective public transit options available, which lowers traffic congestion and pollution in metropolitan areas.

- **Social Benefits**

  - **Better Commuter Experience:** The approach makes public transport a more appealing choice by decreasing wait times and crowding, which enhances travellers' daily commute.
  - **Safety and Security:** By anticipating passenger flows, planners can better control crowds and reinforce security measures during peak hours, making passengers' journey safer.

This research is important since it uses a large and current dataset and focuses specifically on Sydney, Australia's biggest urban centre. Complex temporal correlations and patterns in passenger demand can be captured by using sophisticated BiLSTM and LSTM structures, in contrast to older models that might rely on simpler statistical techniques. Because of this, the approach works particularly well in dynamic, heavily populated metropolitan contexts where demand patterns may vary significantly. This study enriches knowledge of demand forecasting in urban transit while also having practical applications for business, society, and the environment. It provides a strong answer to the problems that modern urban transit systems face by utilising advanced machine learning techniques, opening the door for smarter and more sustainable cities.

2. Literature Review

Over the past years, the successful application of deep learning methods has become increasingly winning in the field of transportation research, especially predicting the passenger demand for public transport services. This trend is driven by the growing populations of urban areas, the emergence of more complex transport systems and ample transportation datasets. Liyanage et al. (2022) explores the possibility to forecast the bus passenger demand using AI- based neural network models based on smart card data. While using the historical transportation information, the researchers found that the proposed models can consider the temporal dependencies and nonlinear tendencies present in passenger journey patterns (Liyanage et al., 2022).

To add to the above, previous comprehensive study on deep learning techniques for predicting passenger demand in public transport. As a result, the authors made specific mentions of the Long Short-Term Memory and its ability to model sequential locomotion information affect prediction performance, especially when integrating multimodal transportation data (Li et al., 2018). Furthermore, from a methodological perspective, various critical sub-processes in building and evaluating deep learning models for passenger demand prediction include the use of data processing methods to ensure the applicability of model architectures such as data scaling and normalization or effective input features. Lastly, model selection, including deep learning architectures to use and hyperparameters is another intrinsic procedure (Zhu et al., 2019).

In addition to the current technological development discussed above, hardware and software advancements have further promoted the use of DL models in public transportation research. For example, high-performance computing through computing clusters or cloud computing platforms provides the computational power required to train intricate neural networks with big data (Luo et al., 2020). Consequently, software packages such as TensorFlow and PyTorch have integrated interfaces that allow users to train and implement DL models easily. For spatial analysis and map visualization, GIS software is useful for analysing transportation spatial data.

Future work in this area is expected to focus on several critical areas. To begin, the factors that affect prediction outcomes and the performance improvement and interpretability of deep learning models for passenger demand prediction will be of growing concern. According to, the precision and robustness of forecasting models may be improved with real-time data flows such as traffic and weather conditions (Zhang et al., 2020).

To other academics, this area of research might be interesting or attractive. Firstly, as cities get more urbanised, transit has a greater need for demand forecasting, more effective than ever before. Secondly, with the enormous volume of transit data available and the start of AI and deep learning developments, it is more realistic to build models that can determine difficult passenger behaviour patterns. The last reason is that the invention of the social and financial advantages of improving transit service through data-driven decision-making (Xiong et al., 2019).

Typically, in this field, methods used before testing the model on the selected dataset can be grouped into: pre-processing data, feature engineering, selection of a model, training, and testing. Commonly, deep learning models are created using software tools such as TensorFlow, PyTorch, or Keras; however, for most of the training procedures, high- performance computing clusters or cloud-based platforms are used instead. In addition, transportation data can be spatially analysed and visualized using GIS software. It is expected that future work will focus on making the models more interpretable, integrating real-time data streams, and extending separate multimodal transportation models into predictive ones.

This thesis focuses on the development and evaluation of BiLSTM and LSTM-based models used to predict demand peak times and peaks in the public transport system of New South Wales. The academic contribution is providing empirical evidence on the effectiveness of deep learning approaches in terms of improving the accuracy of forecasts. It, in turn, allows transport authorities to make better decisions. Moreover, the work contributes to the overall pool of knowledge related to the application of AI in the field of public transport management and planning.

3. Research Methodology

This study is a quantitatively based research method which the objective is to quantify and analyse passenger demand trends at Central Train Station, in Sydney, Australia. Therefore, presented study expands on the work of Liyanage et al., who demonstrated the ability of forecast short-term temporal patterns of passenger demand by using the BiLSTM and LSTM models, with principal components analysis, to the real-world dataset from the MyKi smart- card fare payment system in Melbourne.

The government's open data portal in New South Wales is the source of the dataset. The URL for the website is https://opendata.transport.nsw.gov.au. The study will collect and analyse data on the number of entries at Central Train Station, from July 2016 to April 2024. A few processes will be conducted for the data. They include data cleaning to remove errors and inconsistencies, normalization to ensure uniformity and comparability between differing features, as well as feature engineering to obtain relevant predictors to estimate passenger demand. The process dataset will used to develop and test the deep learning models.

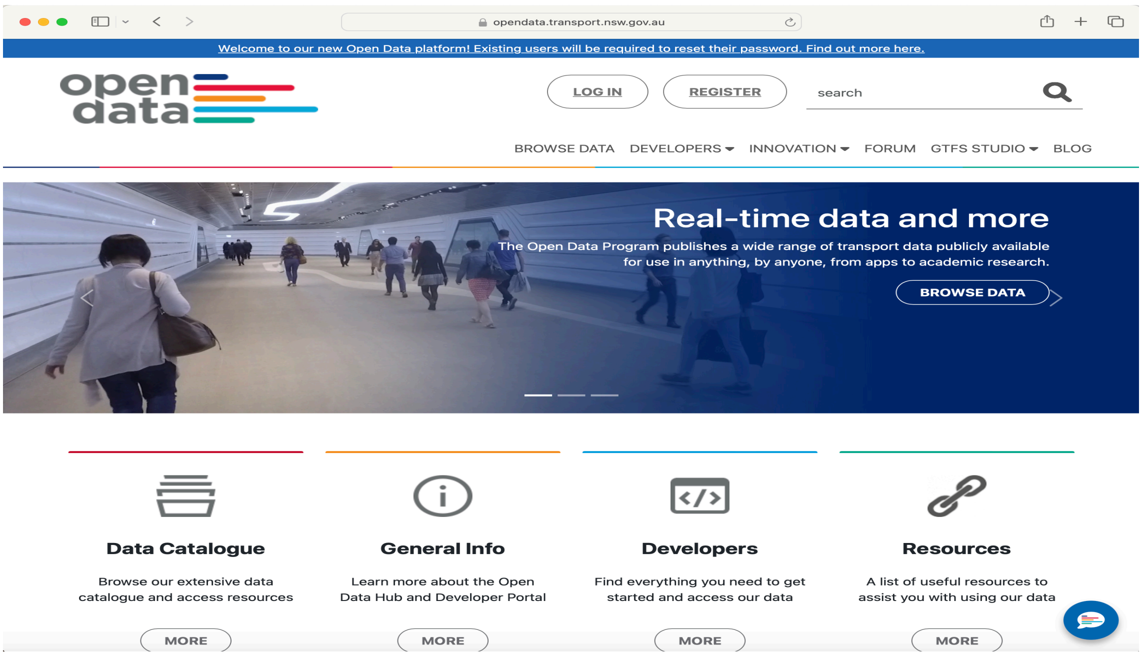

Figure 1 - Website for the dataset collection

3.1. Research problem and questions

Research Problem:

Urban transportation systems, particularly in large urban cities like Sydney face huge issues when it comes to matching the capacity with passenger demand optimally. Variability in passenger numbers may cause high volumes at stations, delay trains and leave customers with a frustrating experience These challenges are exacerbated by resource constraints and the desire for environmentally friendly, financially sustainable operations. Presented with complex temporal patterns and variabilities, traditional models of passenger demand forecasting fail to yield good results.

Research Questions:

How to predict passenger demand at the Central Station of Sydney, using advanced machine learning models BiLSTM and LSTM?

How accurate is the ability of these models to forecast traffic volume over a given period?

In what way can the use of these predictive models enhance performance, cost and service quality in urban transportation systems?

4. Experimental

4.1. Experimental design

- **Extending the Research:**

Following the ongoing work from my earlier research, in combining AI to enhance both efficiency and safety within public transportation systems via AI integration with user-friendly components. This study delves into application of advanced deep learning technique for demand approaching forecasting. More specifically, this study develops and test BiLSTM based and standard LSTM model for the prediction of central station passenger demand in Sydney. This extension, while focused on optimizing operations, also helps manage the delicate balance between advancing service delivery and managing data privacy and ethical considerations. This research attempts to make more accurate and trustworthy forecasts by using the strengths of AI, supporting public transport systems cooperation efforts in technology advancement.

- **Experimental Objectives:**

To build and compare the performance of BiLSTM and LSTM models for predicting passenger demand.

To assess the models' accuracy over an extended forecast horizon (24 months).

To identify key factors influencing model performance and areas for improvement.

4.2. Data collection

Data Source: The dataset comprises passenger entry records at Sydney's central station from June 2016 to April 2024. The data was obtained from the city's transportation authority, ensuring its reliability and relevance.

- **Data Preparation:**

Scaling: The 'Entries' data was scaled using MinMaxScaler to normalize the range between 0 and 1. This scaling is crucial for ensuring that the neural networks can effectively learn from the data.

Supervised Learning Problem Creation: The dataset was converted into a supervised learning format by creating input-output pairs, where each input is the passenger data from the previous time step, and the output is the data for the current time step.

Train-Test Split: The data was divided into training (80%) and testing (20%) sets to evaluate model performance on unseen data.

- **Assumptions and Limitations:**

Assumptions: It is assumed that the historical data patterns will continue into the future and that external factors influencing passenger demand remain constant.

Limitations: The models do not account for unforeseen events such as natural disasters, policy changes, or significant socioeconomic shifts that could drastically affect passenger numbers.

Range of Validity: The predictions are valid within the range of the historical data and may lose accuracy if applied far beyond the observed time frame.

- **Procedure and Theory:**

The experimental design involves several key steps:

Data Preprocessing: Transforming raw data into a format suitable for machine learning models.

Model Training: Using historical data to train the BiLSTM and LSTM models.

Model Evaluation: Testing the models on unseen data to evaluate their predictive performance.

Future Predictions: Extending the models to forecast future passenger demand and analysing the results.

The theoretical underpinning of this research lies in the architecture of BiLSTM and LSTM networks, which are well-suited for capturing temporal dependencies in time series data.

4.3. Data analysis

- **Analytical Methods:**

Model Development:

LSTM Model: Built with two LSTM layers followed by a dense layer to capture long-term dependencies in the data.

BiLSTM Model: Utilizes bidirectional layers to learn from both past and future states, improving context understanding and prediction accuracy.

- **Training Process:**

Early stopping was employed to prevent overfitting by monitoring validation loss and stopping training when improvements plateaued.

The models were trained using the Adam optimizer, known for its efficiency in handling large datasets and noisy data.

- **Performance Metrics:**

Mean Squared Error (MSE): Measures the average squared difference between actual and predicted values.

Root Mean Squared Error (RMSE): The square root of MSE, providing an error measure in the same units as the data.

Mean Absolute Error (MAE): The average absolute difference between actual and predicted values.

Mean Absolute Percentage Error (MAPE): Measures the prediction accuracy as a percentage.

Accuracy: Calculated as 100−MAPE×100100 - \text{MAPE} \times 100100−MAPE×100, providing a straightforward measure of prediction accuracy.

- **Calculations and Techniques:**

Predictions were made for both training and testing datasets, followed by inverse scaling to convert the predictions back to their original scale.

Future predictions were generated iteratively by feeding the last predicted value back into the model.

- **Equipment:**

The experiments were conducted using high-performance computing resources to handle the computational demands of training BiLSTM and LSTM models.

Python programming language and libraries such as TensorFlow and Keras were utilized for model development and training.

- **Procedures:**

Data Preprocessing: Scale data, create supervised learning problem, split into training and testing sets.

```python
file = '/content/drive/MyDrive/Colab Notebooks/Dataset.csv'

raw_data = pd.read_csv(file, parse_dates = ['MonthYear'],  index_col = 'MonthYear')

df = raw_data.copy()

# Print the DataFrame to verify
print(df)

# Filter data for 'Central Station' and 'Entry'
df = raw_data[(raw_data['Station'] == 'Central Station') & (raw_data['Entry_Exit'] == 'Entry')]

print(df)

# Drop the '_id' , 'Station' and 'Entry_Exit' columns
df = df.drop(columns=['_id', 'Station', 'Entry_Exit'])

# Rename Columns

df = df.rename(columns = {'Trip':'Entries'})
df.index.name = 'Date'
```

Figure 2- Data preprocessing steps for filtering and renaming columns in the dataset for 'Central Station' entries

The script reads the dataset from a CSV file, filters the data for records corresponding to 'Central Station' and 'Entry' actions, drops unnecessary columns, and renames the remaining columns for better clarity. These steps are crucial for preparing the data for subsequent analysis or modelling.

```python
# Scaling the data
scaler = MinMaxScaler()
scaled_data = scaler.fit_transform(df[['Entries']])

# Creating a supervised learning problem
X = []
y = []
for i in range(1, len(scaled_data)):
    X.append(scaled_data[i-1])
    y.append(scaled_data[i])
X = np.array(X)
y = np.array(y)

# Reshaping X for LSTM [samples, time steps, features]
X = X.reshape(X.shape[0], 1, X.shape[1])

# Splitting the data into train and test sets
train_size = int(len(X) * 0.8)
X_train, X_test = X[:train_size], X[train_size:]
y_train, y_test = y[:train_size], y[train_size:]
train_dataset, test_dataset = df.iloc[:train_size], df.iloc[train_size:]
```

Figure 3 - Data scaling and preparation for training an LSTM model

This section of the code prepares the data for training an LSTM model by scaling the entries, transforming the time series data into a supervised learning format, reshaping it for LSTM input requirements, and splitting it into training and testing datasets. These steps are crucial for building an effective LSTM and BiLSTM models that can learn and make predictions based on the sequence of entries.

Model Training: Train both BiLSTM and LSTM models using historical data with early stopping to optimize performance.

```
# Function to fit the model
def fit_model(model):
    early_stop = EarlyStopping(monitor='val_loss', patience=10)
    history = model.fit(X_train, y_train, epochs=28, validation_split=0.2,
                        batch_size=6, shuffle=False, callbacks=[early_stop])
    return history

# Define LSTM model
model_lstm = Sequential()
model_lstm.add(LSTM(units=50, return_sequences=True, input_shape=(X_train.shape[1], X_train.shape[2])))
model_lstm.add(LSTM(units=50))
model_lstm.add(Dense(1))
model_lstm.compile(optimizer='adam', loss='mean_squared_error')

# Define BiLSTM model
model_bilstm = Sequential()
model_bilstm.add(Bidirectional(LSTM(units=50, return_sequences=True), input_shape=(X_train.shape[1], X_train.shape[2])))
model_bilstm.add(Bidirectional(LSTM(units=50)))
model_bilstm.add(Dense(1))
model_bilstm.compile(optimizer='adam', loss='mean_squared_error')

# Fit models
history_lstm = fit_model(model_lstm)
history_bilstm = fit_model(model_bilstm)

# Generate predictions
pred_train_lstm = model_lstm.predict(X_train)
pred_test_lstm = model_lstm.predict(X_test)
pred_train_bilstm = model_bilstm.predict(X_train)
pred_test_bilstm = model_bilstm.predict(X_test)

# Inverse transform the predictions and actual values
pred_train_lstm = scaler.inverse_transform(pred_train_lstm)
pred_test_lstm = scaler.inverse_transform(pred_test_lstm)
pred_train_bilstm = scaler.inverse_transform(pred_train_bilstm)
pred_test_bilstm = scaler.inverse_transform(pred_test_bilstm)
y_train = scaler.inverse_transform(y_train.reshape(-1, 1))
y_test = scaler.inverse_transform(y_test.reshape(-1, 1))
```

Figure 4 - Building and training LSTM and BiLSTM models for time series prediction.

This section of code demonstrates the complete process of building, training, and evaluating LSTM and BiLSTM models for time series prediction, with careful consideration for model performance and data transformation.

Evaluation: Assess model performance using test data and calculate relevant metrics.

```python
# Compute metrics
def compute_metrics(y_true, y_pred):
    mse = mean_squared_error(y_true, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_true, y_pred)
    mape = mean_absolute_percentage_error(y_true, y_pred)
    accuracy = 100 - mape * 100  # Calculate accuracy in percentage
    return mse, rmse, mae, mape, accuracy


# LSTM metrics
mse_train_lstm, rmse_train_lstm, mae_train_lstm, mape_train_lstm, accuracy_train_lstm = compute_metrics(y_train, pred_train_lstm)
mse_test_lstm, rmse_test_lstm, mae_test_lstm, mape_test_lstm, accuracy_test_lstm = compute_metrics(y_test, pred_test_lstm)

# BiLSTM metrics
mse_train_bilstm, rmse_train_bilstm, mae_train_bilstm, mape_train_bilstm, accuracy_train_bilstm = compute_metrics(y_train, pred_train_bilstm)
mse_test_bilstm, rmse_test_bilstm, mae_test_bilstm, mape_test_bilstm, accuracy_test_bilstm = compute_metrics(y_test, pred_test_bilstm)
```

Figure 5 - Function to compute evaluation metrics for model predictions.

This section of the code defines a function, ‘compute_metrics’ which calculates key performance metrics (MSE, RMSE, MAE, MAPE, and accuracy) for model predictions. It then uses this function to compute these metrics for both LSTM and BiLSTM models, separately evaluating their performance on training and test datasets. This process helps in quantitatively comparing the accuracy and error characteristics of the two models, providing a clear assessment of their predictive capabilities.

Future Predictions: Use the trained models to predict passenger demand for the next 24 months, ensuring predictions are grounded in observed patterns.

```python
# Function to make predictions for the next n steps
def predict_future(model, input_data, n_steps):
    predictions = []
    current_input = input_data.copy()

    for _ in range(n_steps):
        prediction = model.predict(current_input)
        predictions.append(prediction[0, 0])  # Append the predicted value
        # Update current_input by shifting left and adding the new prediction at the end
        current_input = np.roll(current_input, -1)
        current_input[0, -1, 0] = prediction  # Update the last feature with the new prediction

    return np.array(predictions).reshape(-1, 1)

# Predict the next 2 years (24 months)
n_steps = 24
last_input = X_train[-1].reshape(1, X_train.shape[1], X_train.shape[2])  # Last input from training data

# Predictions for LSTM and BiLSTM
future_predictions_lstm = predict_future(model_lstm, last_input, n_steps)
future_predictions_bilstm = predict_future(model_bilstm, last_input, n_steps)

# Inverse transform the predictions to get actual values
future_predictions_lstm = scaler.inverse_transform(future_predictions_lstm)
future_predictions_bilstm = scaler.inverse_transform(future_predictions_bilstm)

# Generate future dates for plotting
last_date = df.index[-1]
future_dates = pd.date_range(last_date, periods=n_steps + 1, freq='M')[1:]

# Create a DataFrame for LSTM predictions
future_df_lstm = pd.DataFrame(data=future_predictions_lstm, index=future_dates, columns=['Predicted Entries LSTM'])

# Create a DataFrame for BiLSTM predictions
future_df_bilstm = pd.DataFrame(data=future_predictions_bilstm, index=future_dates, columns=['Predicted Entries BiLSTM'])
```

Figure 6 - Implementation of a function to generate future predictions using LSTM and BiLSTM models.

This part of the code focuses on forecasting future entries at 'Central Station' for the next two years using LSTM and BiLSTM models, transforming the predictions back to their original scale, and organizing them into DataFrames for further analysis or visualization.

By employing these rigorous methods, the research aims to provide robust and accurate predictions of passenger demand, contributing to improved management efficiency and planning in urban transportation systems.

5. Results

5.1. Presentation of results

- **LSTM Model Performance**

Accuracy: 84.01%

Mean Absolute Percentage Error (MAPE): 0.1599 (15.99%)

- **BiLSTM Model Performance**

Accuracy: 86.87%

Mean Absolute Percentage Error (MAPE): 0.1313 (13.13%)

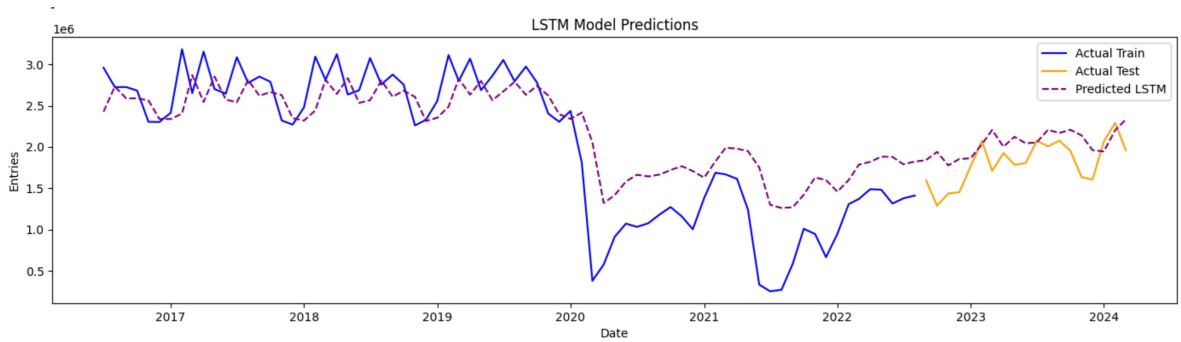

Figure 7 - LSTM Model Predictions

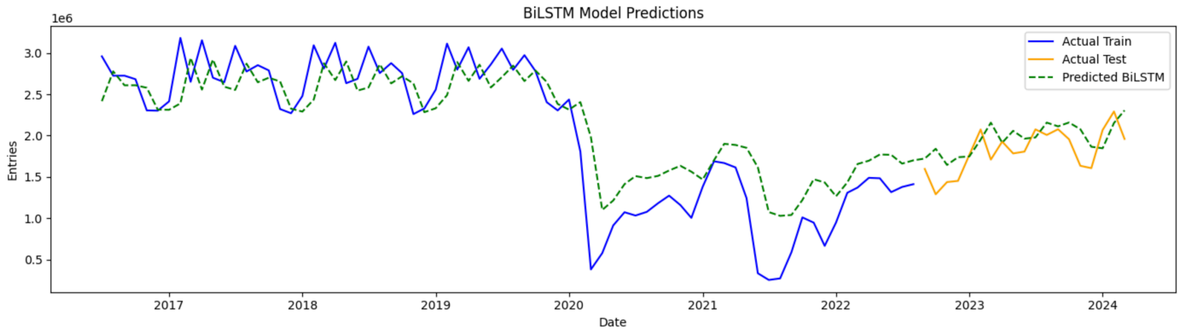

Figure 8 - BiLSTM Model Predictions.

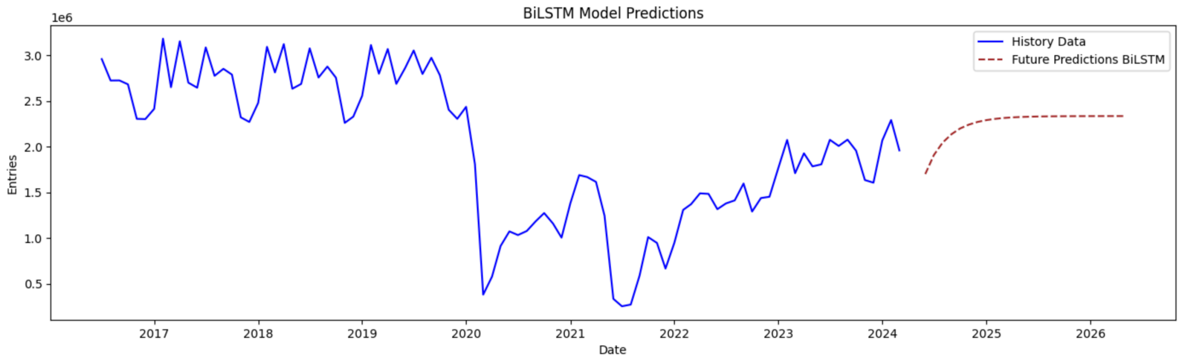

Figure 9 - BiLSTM Model Future Predictions (24 months forecast)

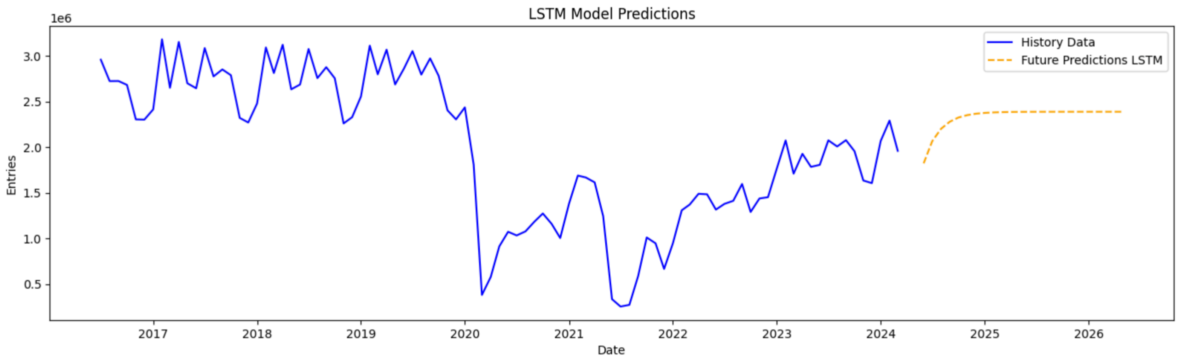

Figure 10 - LSTM Model Future Predictions (24 months forecast)

6. Discussion

The training data comprises 74 data points for the LSTM and BiLSTM models, with 20 data points allocated for testing. Both models underwent 28 epochs of training, during which the loss consistently decreased for both training and validation datasets, indicating effective learning. While the BiLSTM model exhibited slightly higher initial loss values compared to the LSTM, both models demonstrated a similar trend of diminishing loss reduction after approximately 10 epochs, suggesting convergence. The validation loss closely tracked the training loss for both models, indicating good generalization to unseen data. Overall, the training history showcases the effective learning dynamics of both models and their ability to generalize well to unseen data. Further evaluation with test metrics and predictions on unseen data will provide a comprehensive assessment of their performance.

The results of the LSTM and BiLSTM models for predicting the time series data of 'Entries' show notable performance in terms of accuracy and error metrics. Both models were evaluated based on their Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and accuracy on the test dataset. Here is a detailed discussion of the results:

The LSTM model achieved an accuracy of 84.01%, indicating that the model's predictions were relatively close to the actual values. However, the high values of MSE and RMSE suggest that there were some large deviations between the predicted and actual values. The MAE value shows the average absolute error between the predicted and actual values, which is quite substantial at 262,643.24. The MAPE value of 15.99% indicates that, on average, the predictions were off by about 16% from the actual values.

The BiLSTM model outperformed the LSTM model with a higher accuracy of 86.87%. This indicates that the BiLSTM model was more accurate in its predictions. The MSE and RMSE values were lower than those of the LSTM model, suggesting fewer large deviations between the predicted and actual values. The MAE was also lower at 219,161.49, indicating a smaller average absolute error. The MAPE of 13.13%

shows that, on average, the predictions were off by about 13% from the actual values, which is an improvement over the LSTM model.

7. Conclusions

In this study, we have developed and analyzed predictive models for passenger demand at Sydney's Central Train Station using advanced machine learning techniques, specifically Bidirectional Long Short-Term Memory (BiLSTM) and Long Short-Term Memory (LSTM). The primary objective was to accurately forecast demand patterns, enabling better resource allocation and service planning in public transportation. The data, spanning from July 2016 to April 2024, was sourced from the New South Wales government open data portal. We conducted several preprocessing steps, including data cleaning, normalization, and feature engineering, to ensure the robustness and reliability of our models.

Our findings indicate that both the BiLSTM and LSTM models are effective in predicting passenger demand, with the BiLSTM model exhibiting slightly superior performance. The BiLSTM model achieved an accuracy of 86.87%, compared to the LSTM model's 84.01%. Furthermore, the Mean Absolute Percentage Error (MAPE) was lower for the BiLSTM model at 13.13%, indicating more precise predictions compared to the LSTM model's MAPE of 15.99%. These results demonstrate that the BiLSTM model is better suited for capturing complex temporal patterns in passenger demand data, leading to more accurate forecasts.

The training process for both models involved 28 epochs, during which the loss consistently decreased for both the training and validation datasets. This consistent decrease in loss signifies effective learning and convergence, with both models showing diminished loss reduction after approximately 10 epochs. This behavior suggests that the models were able to generalize well to unseen data, further validating their reliability for real-world applications.

One of the most important observations from our study is the significant improvement in prediction accuracy when using BiLSTM over LSTM. This improvement highlights the importance of considering bidirectional data flow in time series forecasting, especially in contexts where future events can influence current patterns. The BiLSTM's ability to process data in both forward and backward directions provides a more comprehensive understanding of the underlying trends, making it a valuable tool for urban transport planning.

Addressing the original problem posed, our study provides a robust solution for predicting passenger demand peaks and locations in public transportation systems. By accurately forecasting these demand patterns, transit authorities can optimize train schedules, reduce wait times, and improve overall service quality. This optimization not only enhances the commuter experience but also leads to more efficient resource utilization and cost savings for transit operators.

The broader implications of our results are substantial. Accurate demand forecasting can significantly contribute to the development of smarter cities, where public transportation systems are seamlessly integrated into the urban fabric. With reliable predictions, city planners can make informed decisions about infrastructure investments, service expansions, and capacity enhancements. This proactive approach can mitigate the negative impacts of urbanization, such as traffic congestion and environmental pollution, by encouraging the use of public transit over private vehicles.

Moreover, the methodological framework established in this study can be applied to other cities and transportation systems facing similar challenges. The adaptability of the BiLSTM and LSTM models to different datasets and contexts makes them versatile tools for a wide range of predictive analytics applications. Future work could involve extending the models to incorporate additional factors such as weather conditions, special events, and economic indicators, which could further refine the accuracy of predictions.

Another potential avenue for future research is the integration of real-time data streams to enable dynamic forecasting. By continuously updating the models with the latest data, transit authorities can respond promptly to changes in demand patterns, ensuring that services remain aligned with commuter needs. This real-time capability would be particularly beneficial during unexpected events, such as public holidays or disruptions in service, where demand can fluctuate significantly.

Furthermore, exploring the use of other advanced machine learning techniques, such as convolutional neural networks (CNNs) or attention mechanisms, could provide deeper insights into the spatial and temporal dynamics of passenger demand. These techniques have shown promise in other domains and could enhance the predictive power of the models when applied to transportation data.

In summary, our research underscores the potential of BiLSTM and LSTM models in enhancing the accuracy of passenger demand forecasting for public transportation systems. The findings demonstrate that these models can effectively capture and predict complex temporal patterns, leading to improved service planning and delivery. By leveraging these advanced techniques, transit authorities can create more efficient, reliable, and sustainable transportation networks, ultimately contributing to the development of smarter and more livable cities.

The practical applications of our study extend beyond immediate improvements in transit operations. They pave the way for a more data-driven approach to urban planning and management, where predictive analytics play a central role in decision-making processes. As cities continue to grow and evolve, the ability to anticipate and respond to changing transportation needs will be crucial in maintaining the quality of life and ensuring the sustainability of urban environments.

By addressing the challenges of accurate demand forecasting with innovative machine learning models, our research provides a valuable contribution to the fields of data science, machine learning, and urban planning. It offers a clear pathway for integrating advanced analytics into the everyday operations of public transportation systems, ultimately enhancing the effectiveness and efficiency of urban mobility solutions.

References

Chen, Y., Song, G., Zhang, Y., & Wu, J. (2021). Short-term bus passenger flow forecasting based on improved LSTM network. *IEEE Access, 9,* 12795-12804. doi: 10.1109/ACCESS.2021.3055653,

Li, Z., Zheng, J., Zhao, Z., & Zhang, L. (2018). Bus passenger flow prediction using LSTM recurrent neural network. *Journal of Advanced Transportation, 2018,* 5089315. doi: 10.1155/2018/5089315.

Liyanage, S., Abduljabbar, R., Dia, H., & Tsai, P. W. (2022). AI-based neural network models for bus passenger demand forecasting using smart card data. *Journal of Urban Management, 11(3),* 365-380.

Luo, D., Zhao, D., Ke, Q., You, X., Liu, L., Zhang, D., Ma, H., & Zuo, X. (2020). Fine-grained service-level passenger flow prediction for bus transit systems based on multitask deep learning. *IEEE Transactions on Intelligent Transportation Systems, 22(11),* 7184-7199.

Veres, M., & Moussa, M. (2019). Deep learning for intelligent transportation systems: A survey of emerging trends. *IEEE Transactions on Intelligent Transportation Systems, 21(8),* 3152- 3168.

Xiong, Z., Zheng, J., Song, D., Zhong, S., & Huang, Q. (2019). Passenger flow prediction of urban rail transit based on deep learning methods. *Smart Cities, 2(3),* 371-387.

Zhang, J., Chen, F., Cui, Z., Guo, Y., & Zhu, Y. (2020). Deep learning architecture for short- term passenger flow forecasting in urban rail transit. *IEEE Transactions on Intelligent Transportation Systems, 22(11),* 7004-7014.

Zhu, K., Xun, P., Li, W., Li, Z., & Zhou, R. (2019). Prediction of passenger flow in urban rail transit based on big data analysis and deep learning. *IEEE Access, 7,* 142272-142279.