CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE

Project Reference No.: *47S_BE_0343*

College *: Biluru Gurubasava Mahaswamiji Institute of Technology, Mudhol*
Branch *: Computer Science and Engineering*
Guide(s) *: Prof. Varun P. Sarvade and Prof. Vinayak A. Telsang*
Student(S) *: Mr. Gopal R. Chenni*
Mr. Pradnya S. Bilagi
Mr. Dhanush J. Yadav
Mr. Niranjan S. Purad

Keywords:

Artificial Intelligence. Artificial Neural Network, Long Short-Term Memory, Recurrent Neural Network, Decision Tree, Prediction

Introduction:

KSRTC (Karnataka State Road Transport Corporation) is one of the largest public transport corporations in India. As per KSRTC data for the month of December 2020, 12.93 lakh passengers travel per day and 40% buses have incurred miner and major accidents in the same month, and also KSRTC reported 2000 litters of fuel wastage per 2 to 4 kms. Crowd prediction for bus lines utilizing artificial intelligence is a groundbreaking application that harnesses AI algorithms like Decision Tree, ANN, RNN, and LSTM Models and data analytics to forecast and manage passenger loads and bus capacities. By amalgamating offline data, and machine learning models, this technology aims to revolutionize the public transportation sector. Through predictive analysis, AI algorithms can anticipate and project the volume of passengers at various bus stops or along specific routes at different times of the day. The primary objective is to optimize bus services by efficiently allocating resources, adjusting schedules, and enhancing passenger experience. By anticipating crowdedness, transportation authorities can implement strategies like deploying additional buses, altering routes, or regulating frequencies to mitigate overcrowding and improve overall efficiency

Objectives of Project:

- Develop an AI model that can accurately predict whether bus should be allocated for the given source and destination.
- To improve public transportation efficiency and passenger experience on KSRTC buses using artificial intelligence models like Decision Tree, Artificial Neural Network, Recurrent Neural Network and Long Short-Term Memory.
- Provide a comparative analysis of all the above-mentioned models in terms of accuracy and loss.

Methodology:

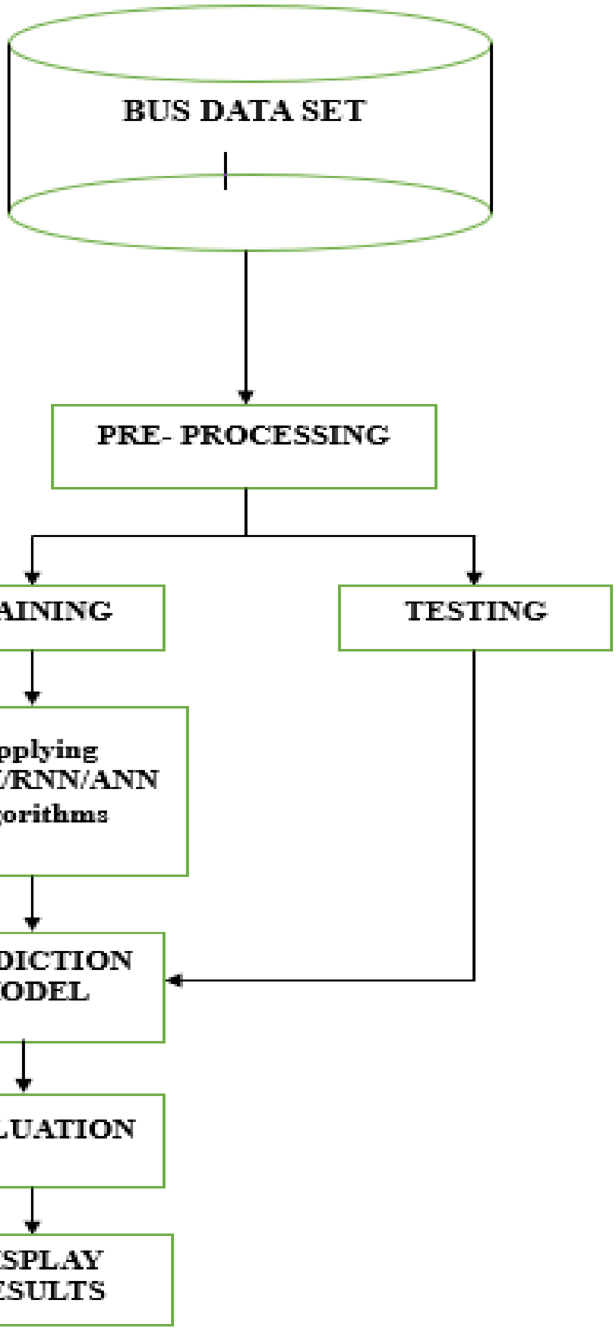
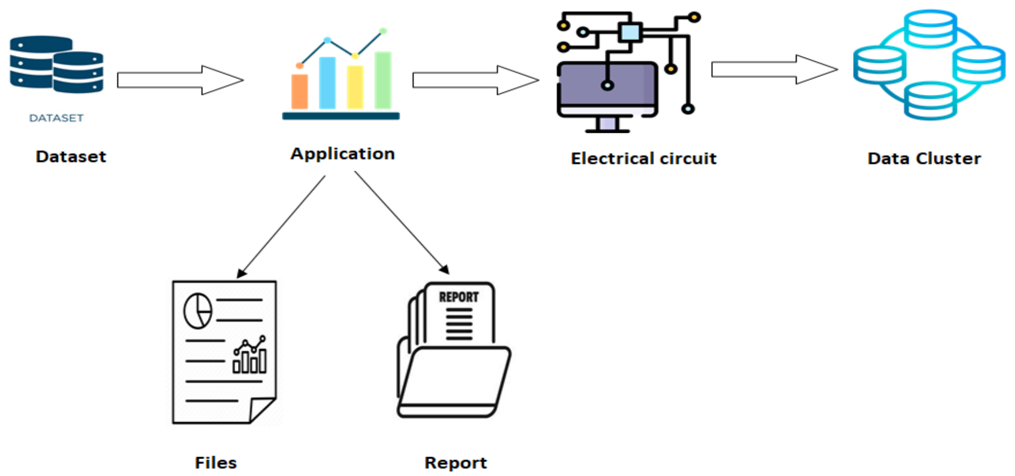


Figure 7.1: Flow Chart

Figure 7.2:



Architecture of Passenger Flow

Public transport is a key element to ensure urban mobility and allow citizens to move easily and efficiently in urban areas. However, managing the influx of passengers in public transport is a daily challenge for transport companies. In this context, predicting the number of passengers on bus lines is essential to optimize the planning, management and organization of the public transport service. In this project, we propose a methodology to predict the number of passengers on bus lines.

Here we have to use KSRTC bus data set and it consist the following parameters

Table 7.1 Sample Data Set

Source	**Destination**	**Trip**	**Route**	**Slot No**	**Adults**	**Childs**	**Revenue**	**Passenger count**	**Label**
1	2	3	1	1	50	4	2589	54	1
1	4	3	1	1	50	10	3289	60	1
1	7	2	2	2	48	4	5478	52	0
1	13	2	3	3	49	7	5100	56	1
1	10	2	3	3	54	4	4989	58	1

After collecting the data to compare, the next step is to load and pre-process the data, to ensure that it is in the correct format and that it was clean and error-free. The following steps will be used for preprocessing:

- Data cleaning: Removal of missing values, duplicates and correction of errors.
- Feature scaling: Rescale the data so that each feature has a similar scale. Common scaling techniques include min-max scaling, standardization, and normalization.
- Feature engineering: Create new features from existing features that might be more informative for the model. These can be transformations, aggregations or combinations of features.
- Data encoding: Convert categorical data into numerical data. It can be point coding, ordinal coding or binary coding.

- Data Splitting: Split your data into training and testing sets. The training set is used to train the model, and the test set is used to evaluate the performance of the model on new, unpublished data. Another step in the data mining process, when the end goal is to predict the outcome, is to create visualizations that help understand the outcome and discover the relationships between attributes and the outcome.

REQUIREMENTS SPECIFICATION:

HARDWERE USED

- Processor: 12th Gen Intel(R) Core(TM) i5-1235U 1.30 GHz
- Memory: 8.00 GB
- Hard Disk space: 2GB Others: Computer peripherals, such as keyboard and mouse
  SOFTWERE USED
- Operating System: Windows 10
- Google Colab

Results and Conclusions:

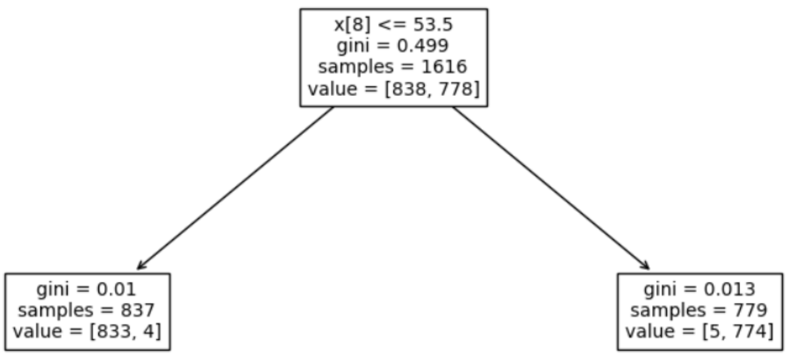
In this section we present result analysis on Decision tree, Artificial Neural Network, Long Short Term Memory, and Artificial Neural Network Model for predicting the Model Accuracy, Model Loss, Model Value loss, and Model Value Accuracy Score.

<u>RESULTS USING DECISION TREE</u>

Figure 8.1: Using score generation of Decision Tree

```
The DT test score is 0.994430693069307
Requirement already satisfied: graphviz in /usr/local/lib/python3.10/dist-packages (0.20.3)
```

The



Decision Tree score is 99% and decision tree contains samples and values of the taring and testing data sets .

<u>RESULTS USING ARTIFICIAL NEURAL NETWORK MODEL</u>

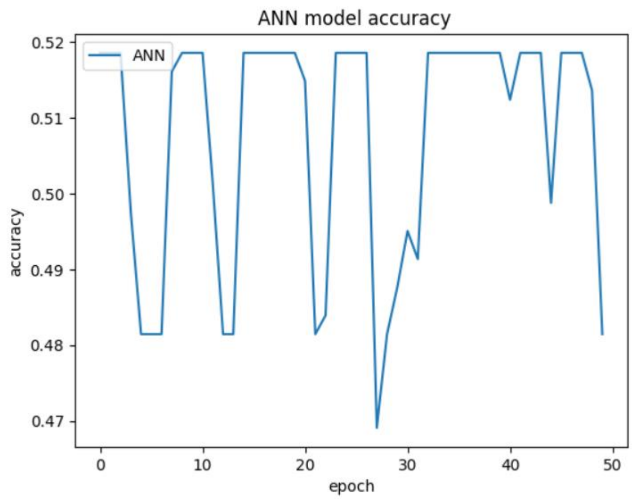


Figure 8.2: Epoch v/s Model Accuracy

ANN Model gives 51% Model accuracy for 50 epochs. with batch size 400, callback list and validation data

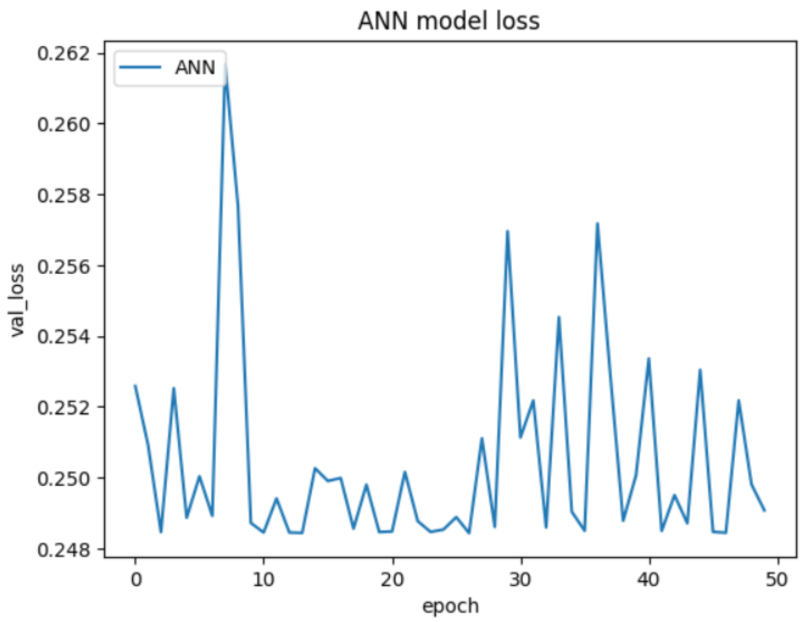


Figure 8.3: Epoch v/s Model Loss

ANN Model gives 25% Model loss for 50 epochs with batch size 400, callback list and validation data

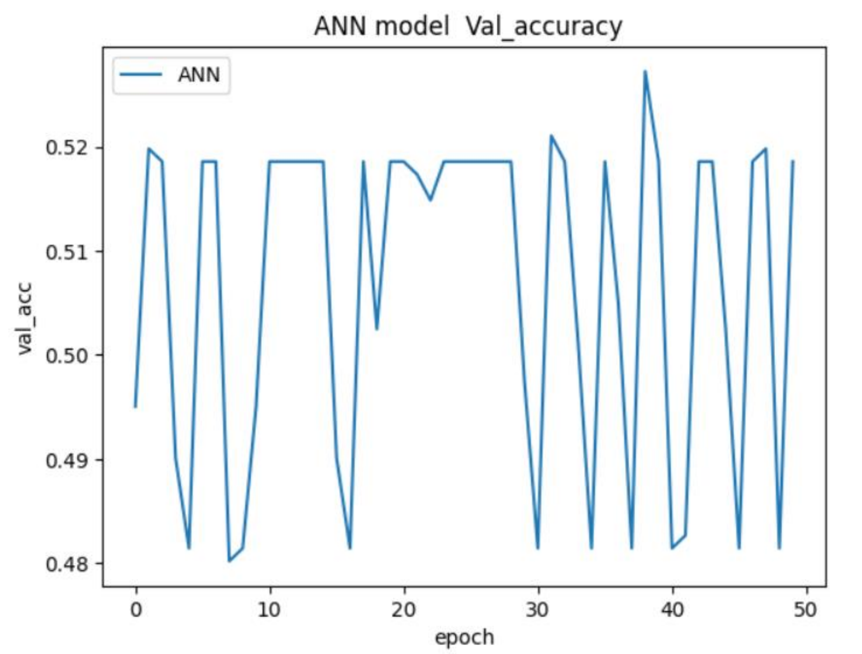


Figure 8.4: Epoch v/s Model Value Accuracy

ANN Model gives 53% Model Value accuracy for 50 epochs, with batch size 400, callback list and validation data

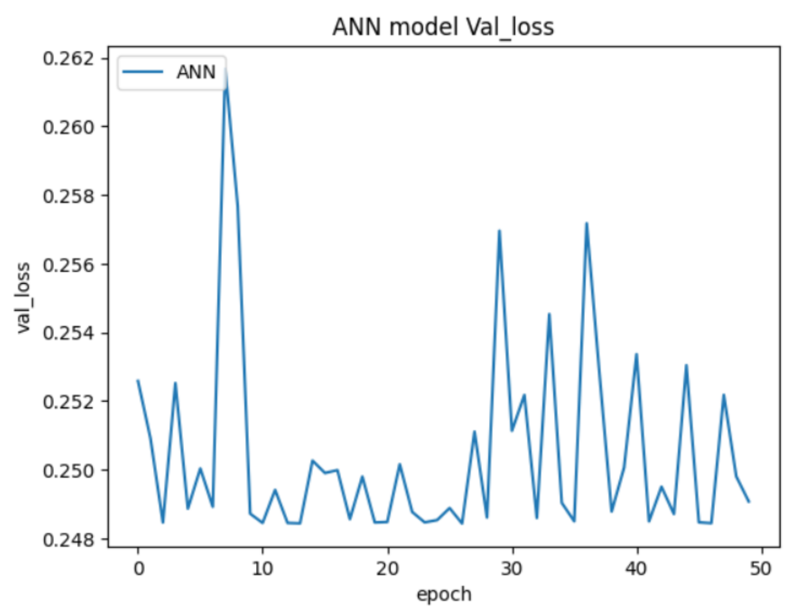


Figure 8.5: Epoch v/s Model Value Loss

ANN Model gives 24% Model Value loss for 50 epochs, with batch size 400, callback list and validation data

RESULTS USING LONG SHORT TERM MEMORY MODEL

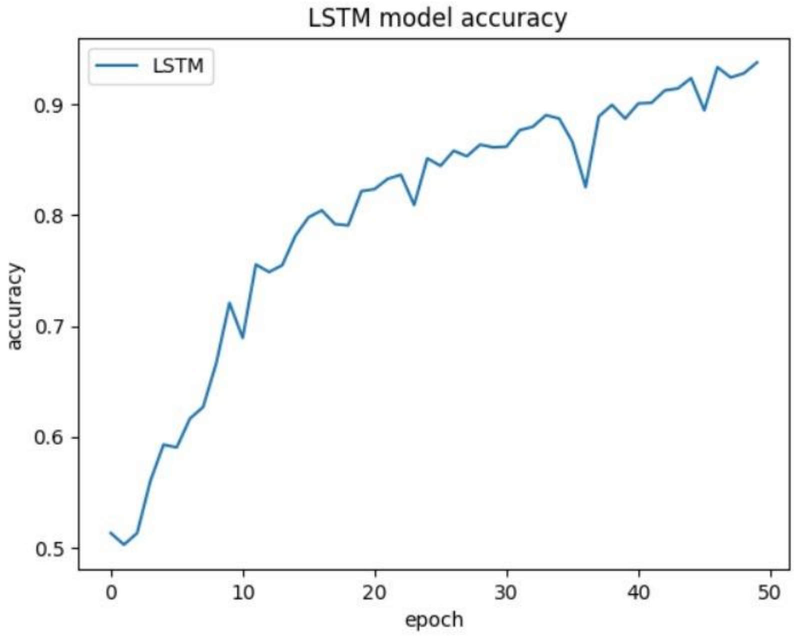


Figure 8.6: Epoch v/s Model Accuracy

LSTM Model gives 95% Model accuracy for 50 epochs with batch size 400, callback list and Validation data.

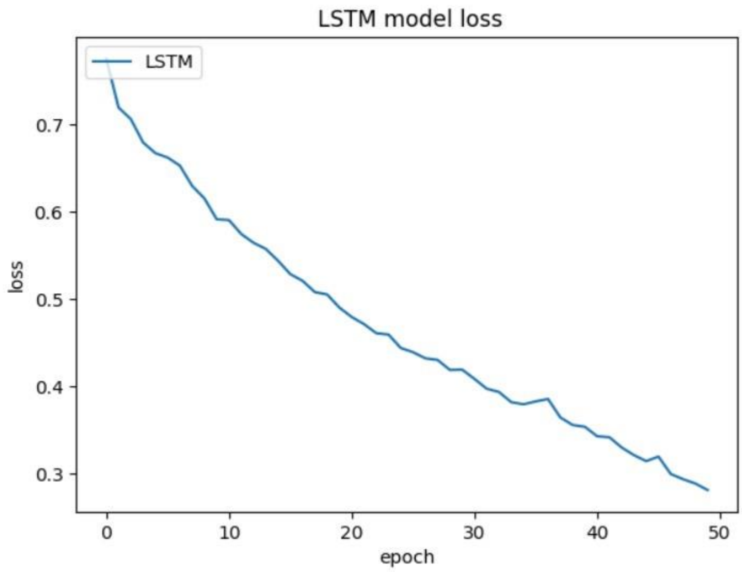


Figure 8.7: Epoch v/s Model Loss

LSTM Model gives 21% Model loss for 50 epochs with batch size 400, callback list and validation data

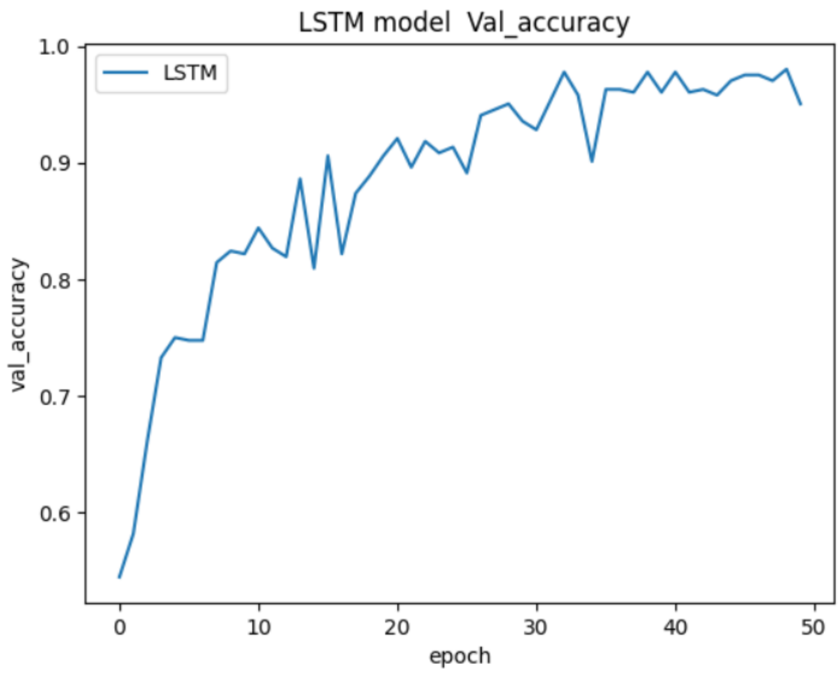


Figure 8.9: Epoch v/s Model Value Accuracy

LSTM Model gives 95% Model loss for 50 epochs with batch size 400, callback list and validation data

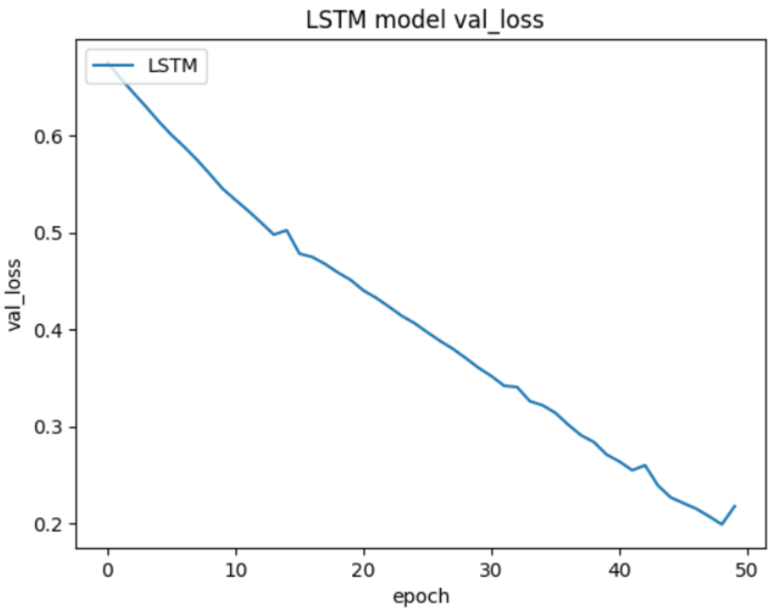


Figure 8.10: Epoch v/s Model Value Loss

LSTM Model gives 21% Model Value accuracy for 50 epochs with batch size 400, callback list and validation data

Result using RECURRENT NEURAL NETWORK MODEL

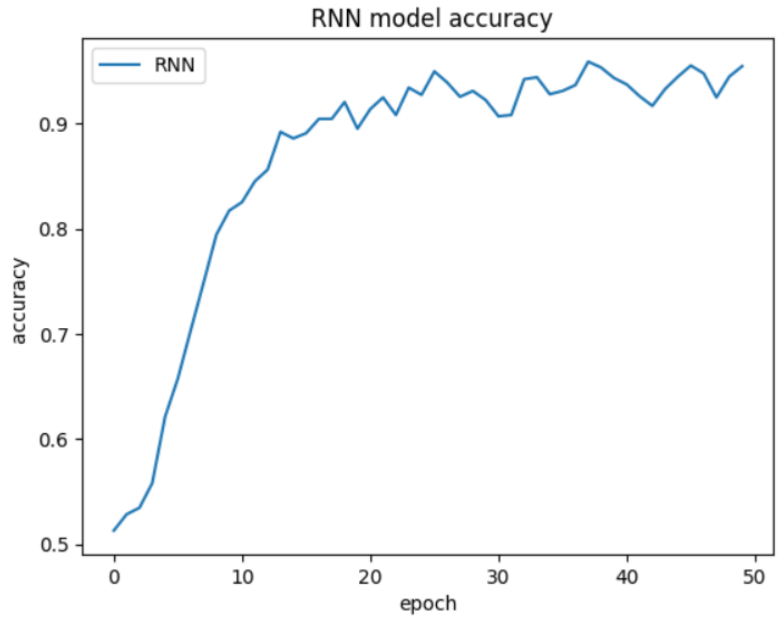


Figure 8.11: Epoch v/s Model Accuracy

RNN Model gives 96% Model Value accuracy for 50 epochs with batch size 400, callback list and validation data.

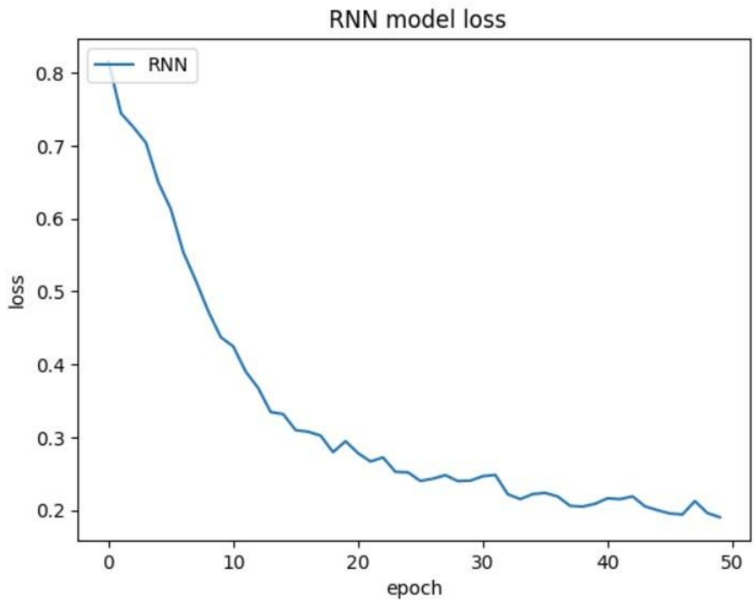


Figure 8.12: Epoch v/s Model Loss

RNN Model gives 20% Model loss for 50 epochs with batch size 400, callback list and validation Data.

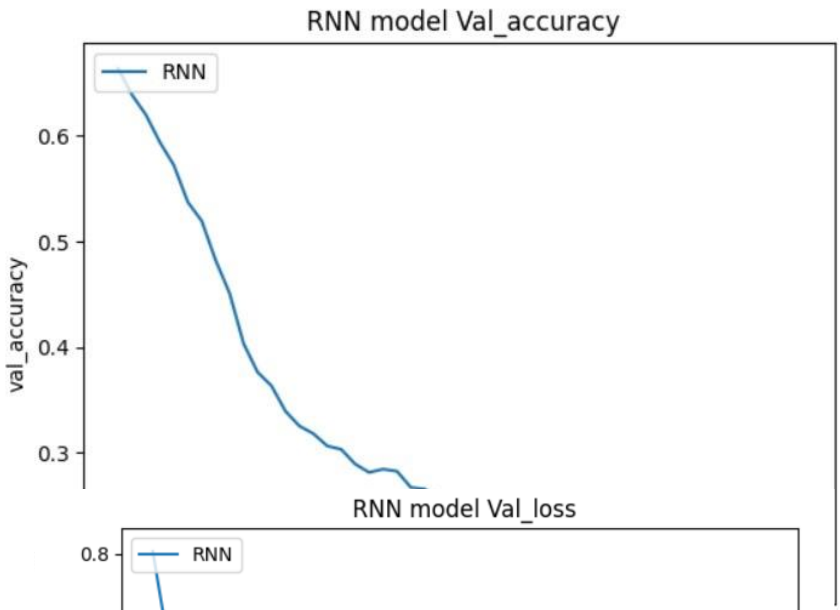


Figure 8.13: Epoch v/s Model Value Accuracy

RNN Model gives 99% Model Value accuracy for 50 epochs with batch size 400, callback list and validation data

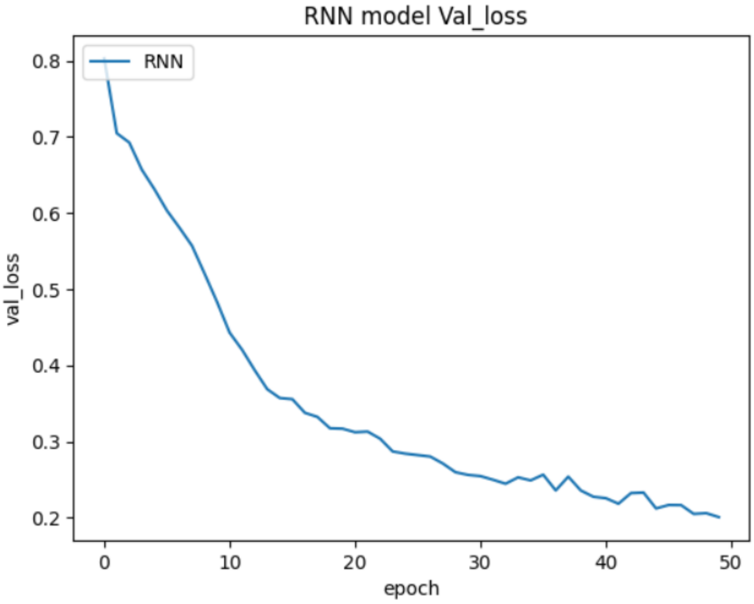


Figure 8.14: Epoch v/s Model Value Losss

RNN Model gives 18% Model Value accuracy for 50 epochs with batch size 400, callback list and validation data

Table 8.1 Comparison between four models using Accuracy and Loss Result

Result	ANN Model	LSTM Model	RNN Model
Model Accuracy	51%	95%	96%
Model Loss	25%	21%	20%

Model Value Accuracy	53%	95%	96%
Model Value Loss	24%	21%	18%

In this project Decision tree, ANN model, LSTM model and RNN model gives 99%, 51%, 95% and 96% of accuracy score respectively.

In conclusion, the successful implementation of AI for crowd prediction in bus lines holds immense potential for optimizing public transportation systems. By leveraging data-driven insights, transit authorities can proactively manage crowd levels, enhance passenger experiences, and ultimately improve the overall efficiency of urban mobility. As we continue to refine and expand upon this project, we aim to contribute to the advancement of smart transportation solutions for sustainable and accessible cities.

Innovation of the project:

- The successful implementation of AI for crowd prediction in bus lines holds immense potential for optimizing public transportation systems.
- By leveraging data-driven insights, transit authorities can proactively manage crowd levels,
- As we continue to refine and expand upon this project, we aim to contribute to the advancement of smart transportation solutions for sustainable and accessible cities.

Scope for Future Work:

- Multimodal data integration: Incorporating real-time data from various sources like traffic conditions, weather patterns, and social media trends can improve prediction accuracy.
- Advanced AI models: Exploring deep learning techniques like recurrent neural networks (RNNs) and convolutional neural networks (CNNs) can capture complex spatiotemporal patterns in ridership data.
- Dynamic bus scheduling: AI can optimize bus routes and schedules based on predicted passenger volumes, reducing wait times and overcrowding.
- Passenger information systems: Real-time crowd prediction can be integrated into mobile apps, allowing passengers to choose less crowded buses or alternative routes.
- Personalized recommendations: AI can suggest personalized travel options based on user preferences and real-time conditions, promoting a seamless commute.
- By leveraging AI's capabilities, bus line crowd prediction can become a powerful tool for improving public transport efficiency and passenger satisfaction.