VISVESVARAYA TECHNOLOGICAL UNIVERSITY,
BELAGAVI, 590 014

Shri. B. V. V. Sangha's

BILURU GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY
MUDHOL-587313

2023-2024

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Project Report On

"CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE"

Student Project Program-47th Series

(KSCST Reference No. 47S_BE_0343)

Submitted By:

MR. GOPAL CHENNI	USN: 2LB20CS005
MISS. PRADNYA BILAGI	USN: 2LB20CS015
MR. DHANUSH YADAV	USN: 2LB20CS003
MR. NIRANJAN PURAD	USN: 2LB20CS013

Under the Guidance of

Prof. Varun P. Sarvade
Assist. Professor, Dept. of Computer Science & Engineering,
BGMIT, Mudhol.

Shri. B. V. V. Sangha's

BILURU GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY

MUDHOL-587313

Department of Computer Science & Engineering

CERTIFICATE

This is to certify that the Project work entitled **"CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE"**, is a bonfide work carried out by **Mr. Gopal Chenni USN:2LB20CS005, Miss. Pradnya Bilagi USN:2LB20CS015, Mr. Dhanush Yadav USN:2LB20CS003, Mr. Niranjan Purad USN:2LB20CS013,** in partial fulfilment for the award of **Bachelor of Engineering** in **Computer Science and Engineering, BGMIT, Mudhol** from **Visvesvaraya Technological University, Belagavi** during the year 2023-24. The Project work report has been approved as it satisfies the academic requirements in respect of project prescribed for the aforesaid degree.

Guide	HOD	Principal
Prof. Varun P. Sarvade	**Prof. Manjunath S. Gabasavalagi**	**Dr. Shravankumar B. Kerur**

Name of the Examiners: **Signature with Date:**

1.
2.

ACKNOWLEDGEMENT

The sense of contentment and elation that accompanies the successful completion of the Project work titled **"CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE",** would be incomplete without mentioning the names of those people who helped us in accomplishing this work. People whose constant guidance and encouragement resulted in its realization.

First and foremost, we deeply express our sincere gratitude to our guide **Prof. Varun P. Sarvade,** Assistant Professor Department of Computer Science & Engineering, BGMIT, Mudhol, for his able guidance and coordinator, regular source of encouragement and assistance throughout this work.

We would like to thank **Prof. Manjunath S. Gabasavalagi,** Head of the Department, Computer Science & Engineering, BGMIT, Mudhol, for his valuable suggestions and advices.

We would like to thank **Dr. Shravankumar B. Kerur, Principal, BGMIT, Mudhol,** for his moral support towards completion of our Project work.

We thank all Faculty members of Department of Computer Science &Engineering for their constant support and encouragement.

Last, but not the least, we would like to thank our parents and friends who provided us their valuable suggestions.

DECLARATION

We, Mr. Gopal Chenni USN:2LB20CS005, Miss. Pradnya Bilagi USN:2LB20CS015, Mr. Dhanush Yadav USN:2LB20CS003, and Mr. Niranjan Purad USN:2LB20CS013 the students of VIII semester B.E in Computer Science and Engineering, Biluru Gurubasava Mahaswamiji Institute of Technology, Mudhol, affiliated to Visvesvaraya Technological University, Belagavi. Hereby declare that we have independently carried out project titled **"CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE"**, and submitted in partial fulfilment of the requirements for the award of degree in Bachelor of Engineering during the academic year 2023-24.

Mr. Gopal Chenni,

USN:2LB20CS005.

Miss. Pradnya Bilagi,

USN:2LB20CS015.

P. S. Bilagi

Mr. Dhanush Yadav,

USN:2LB20CS003.

Mr. Niranjan Purad,

USN:2LB20CS013.

ABSTRACT

Inefficient crowd prediction in bus lines leads to overcrowding, long wait times, missed connections, and suboptimal resource allocation. This hinders passenger satisfaction, service reliability, and overall public transportation efficiency. In current scenario, the Bus stand controllers maintain a bus routine manual and they don’t possess any software for this maintenance. The proposed work will provide solution for controllers, passengers and also public transportation owners a better way for handling the crowd, and optimize fuel consumption. The proposed work also aims to improve travel experience, reduce waiting times, improve operational efficiency, and reduced congestion. We have used Decision tree, ANN, LSTM , RNN models to predict the accuracy value or score. and we predict model accuracy, and model loss. In this project Decision tree, ANN model, LSTM model and RNN model gives 99%, 51%, 95% and 96% of accuracy score respectively.

LIST OF CONTENTS

SI No	TITLE	Page No.
Chapter 1	**Scope and Objectives**	**01**
	1.1 Objectives	
	1.2 Scope	
Chapter 2	**Problem Statement**	**02**
Chapter 3	**Introduction**	**03-09**
	3.1 Artificial Intelligence	
	3.2 How it works	
	3.3 Decision Tree	
	3.4 Artificial Neural Network	
	3.5 Long Short Term Memory	
	3.6 Recurrent Neural Network	
Chapter 4	**Literature Survey**	**10-13**
Chapter 5	**Methodology**	**14-17**
	5.1 Architecture	
	5.2 Flow Chart	
	5.3 Algorithm	
Chapter 6	**Specification Requirements**	**18**
	6.1 Hardware Used	
	6.2 Software Used	
Chapter 7	**Result Analysis**	**19-26**
Chapter 8	**Advantages and Disadvantages**	**27**
	Conclusion And Future Scope	**28**
	Reference	**29**
	Appendix A	**30-36**
	Appendix B	**37-40**
	Appendix C	**41-45**

LIST OF FIGURES

Figure No.	Figure Name	Page No.
Figure :3.3.1	Decision Tree	03
Figure :3.4.1	ANN Model	04
Figure :3.5.1	LSTM Model	07
Figure :3.6.1	RNN Model	08
Figure :5.1.1	Architecture of Passenger Flow	14
Figure :5.2.1	Flow Chart	16
Figure: 7.1	Using score generation of Decision Tree	19
Figure :7.2.1	Epoch v/s Model Accuracy	20
Figure :7.2.2	Epoch v/s Model Loss	20
Figure :7.2.3	Epoch v/s Model Value Accuracy	21
Figure :7.2.4	Epoch v/s Model Value Loss	21
Figure :7.3.1	Epoch v/s Model Accuracy	22
Figure :7.3.2	Epoch v/s Model Loss	22
Figure :7.3.3	Epoch v/s Model Value Accuracy	23
Figure: 7.3.4	Epoch v/s Model Value Loss	23
Figure :7.4.1	Epoch v/s Model Accuracy	24
Figure :7.4.2	Epoch v/s Model Loss	24
Figure: 7.4.3	Epoch v/s Model Value Accuracy	25
Figure: 7.4.4	Epoch v/s Model Value Losss	25

LIST OF TABLES

Table No	Table Name	Page No
Table 5.1	Sample Data Set	14
Table 7.4	Comparison between four models using Accuracy and Loss Result	26

CHAPTER 1

SCOPE AND OBJECTIVES

1.1 Scope

- For this project we have considered four cities namely Mudhol, Jamakhandi, Vijapura, and Bagalkote as sources for KSRTC bus lines and only destination from these sources is considered.
- The dataset was prepared using the offline data obtained from KSRTC offices of above-mentioned cities.

1.2 Objective

- 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.

CHAPTER 2

PROBLEM STATEMENT

Public transportation plays a vital role in urban mobility, but overcrowding on buses often leads to inconvenience, discomfort, and safety concerns for passengers. To address this challenge, we aim to develop an artificial intelligence-based crowd prediction system for a bus line. The system will utilize historical data to forecast bus demand accurately.

CHAPTER 3

INTRODUCTION

3.1 ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is a branch of computer science that deals with creating systems and machines capable of performing tasks that typically require human intelligence. These tasks can include understanding natural language, recognizing patterns, learning from experience, and making decisions. AI encompasses a broad range of techniques and approaches, including machine learning, neural networks, natural language processing, computer vision, and robotics. One of the key goals of AI is to develop algorithms and systems that can mimic or surpass human-level intelligence across various domains.

3.2 HOW IT WORKS

Artificial intelligence (AI) operates through a series of critical phases, the most common of which are data gathering, preprocessing, model training, evaluation, and deployment. This is a broad synopsis of the procedure like identifying problem statement, data collection to monetizing the results.

3.3 DECISION TREE

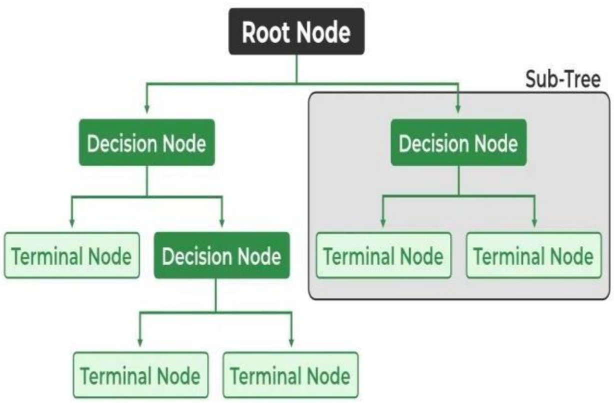

Figure 3.3.1: Decision Tree

In the context of decision trees, "score" generally refers to the impurity measure used to evaluate the quality of a split. The most common impurity measures are Gini impurity and entropy. These measures are used to quantify the homogeneity of a node's class distribution

For classification tasks, the Gini impurity at a node is calculated as

$$G_{ini} = 1 - \sum_{i=1}^{c} (pi)^2$$

Where c is the number of classes and p_i is the proportion of instances of class i in the node. A lower Gini impurity indicates a purer node

Entropy is another measure of impurity that can be used in decision trees. For a node in a classification tree, the entropy is calculated as

$$Entropy = -\sum_{i=1}^{c} pi\, log_2(pi)$$

Where pi is the same as above. As with Gini impurity, lower entropy values indicate a purer node

3.4 ARTIFICIAL NEURAL NETWORK

ANN stands for Artificial Neural Network, a computational model inspired by the biological neural networks present in animal brains. It's a fundamental concept in machine learning and has been widely used in various fields, including computer vision, natural language processing, and robotics.

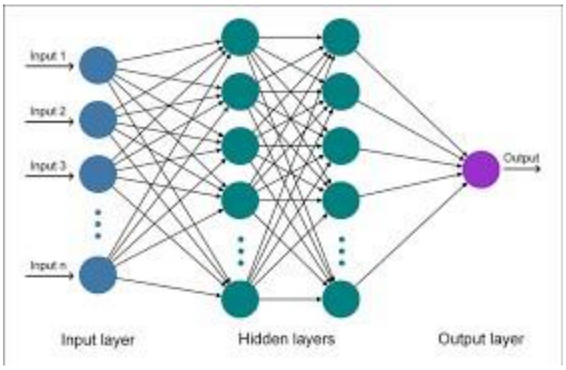

Figure 3.4.1: ANN Model

Here's a brief overview of how artificial neural networks work:

1. Neurons (Nodes): The basic building blocks of an artificial neural network are neurons or nodes. Each neuron receives input signals, performs a computation on those signals, and then produces an output signal.

2. Layers: Neurons are organized into layers. A typical neural network consists of an input layer, one or more hidden layers, and an output layer. The input layer receives the raw input data, the hidden layers process this data, and the output layer produces the final output.

3. Connections (Weights): Neurons in one layer are connected to neurons in the next layer through connections, which have associated weights. These weights determine the strength of the connection between neurons. During the training process, these weights are adjusted to minimize the error between the predicted output and the actual output.

4. Activation Function: Each neuron typically applies an activation function to the weighted sum of its inputs. Common activation functions include sigmoid, tanh, ReLU (Rectified Linear Unit), and softmax.

5. Feedforward Propagation: The process of passing input data through the network to generate an output is called feedforward propagation. During this process, the input data is multiplied by the weights, passed through the activation functions, and propagated forward layer by layer until the output is produced.

6. Backpropagation: Once the output is generated, the network compares it to the actual output (ground truth) and calculates the error. Then, using an optimization algorithm like gradient descent, the network adjusts the weights of the connections backward from the output layer to the input layer. This process is called back propagation and aims to minimize the error between the predicted and actual outputs.

7. Training: The process of adjusting the weights of the neural network to minimize the error on the training data is called training. This involves feeding the training data through the network, performing feed forward propagation, calculating the error, and then performing back propagation to update the weights. This process is repeated iteratively until the network converges to a set of weights that produce satisfactory results.

ANNs are highly flexible and can learn complex patterns and relationships in data, making them powerful tools for a wide range of tasks in machine learning and artificial intelligence.

Artificial Neural Networks (ANNs) are mathematical models used for machine learning and are composed of interconnected nodes that mimic the structure and function of biological neurons. The basic building block of an ANN is the perceptron, which takes an input, applies weights and a bias, and then passes the result through an activation function to produce an output.

The output y^ of a perceptron with input x and weights w (including bias as w 0) can be calculated as

$$y^\wedge = f(\sum_{i=0}^{n} w_i \, x_i$$

Where f is the activation function (e.g., sigmoid, ReLU), n is the number of inputs, w i are the weights, and x i are the inputs

For a single-layer feed forward neural network, the output y^ is the result of passing the inputs through a series of perceptrons.

$$y^\wedge = f(\sum_{j=0}^{m} w_j{}^{(2)} \, h_j)$$

Where m is the number of perceptrons in the hidden layer, w j^2 are the weights connecting the hidden layer to the output, and h j is the outputs of the hidden layer perceptrons.

3.5 LONG SHORT TERM MEMORY

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) architecture, designed to capture long-term dependencies and address the vanishing gradient problem commonly encountered in traditional RNNs. LSTMs are widely used in sequential data tasks such as natural language processing, time series prediction, and speech recognition.

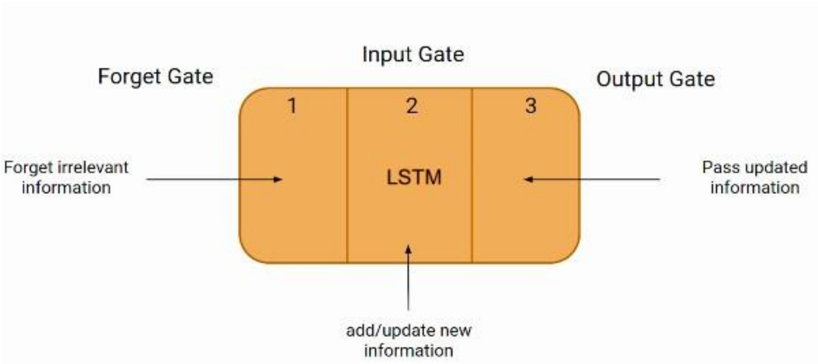

Figure 3.5.1: LSTM Model

LSTMs can capture and propagate long-term dependencies through time, making them effective for tasks that involve sequences with long-range dependencies. The gating mechanisms enable LSTMs to selectively remember, forget, and update information over time, which helps alleviate the vanishing gradient problem encountered in traditional RNNs.

A new intermediate network design called long short-term memory adapts its learning strategy to the data by using a gradient-based approach. With LSTM, backflow issues are resolved. Additionally, blunder discharge was the reason behind the creation of LSTM.

$$i_t = \sigma(w_t[h_t - 1, x_t] + b_i)$$
$$f_t = \sigma(w_f[h_t - 1, x_t] + b_f)$$
$$o_t = \sigma(w_o[h_t - 1, x_t] + b_o)$$

i_t = represents input gate.

f_t = represents forget gate.

o_t = represents output gate.

σ = represents sigmoid function.

w_x = weight for the respective gate(x) neurons.

h_t-1 = output of the previous lstm block(at timestamp t-1).

x_t = input at current timestamp.

b_x =biases for the respective gates(x).

3.6 RECURRENT NEURAL NETWORK

Recurrent Neural Networks (RNNs) are a class of neural networks designed to handle sequential data by maintaining internal memory. They are widely used in tasks involving sequential input or output, such as time series prediction, natural language processing, and speech recognition. The fundamental idea behind RNNs is to process sequences of inputs one element at a time while maintaining a hidden state that captures information about previous elements in the sequence.

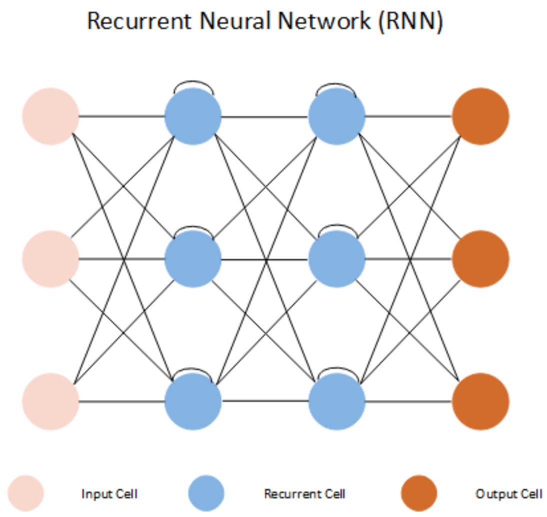

Figure 3.6.1: RNN Model

While standard RNNs have limitations in capturing long-term dependencies due to the vanishing gradient problem, they are still widely used for many sequential tasks, especially when short-term dependencies are sufficient or when computational resources are limited. Unlike feed forward neural networks, which process input data in a single direction, from input to output, RNNs have connections that form a directed cycle, allowing information to persist Mathematically, the hidden state h_t of an RNN at time step t is calculated as a function of the current input x_t and the previous hidden state h_t-1

$$h_t = f(w_h h_t - 1 + w_x x_t + b)$$

Where W_h and W_x are weight matrices, b is a bias vector, and f is an activation function (commonly tanh or ReLU)

The output y_t of the RNN at time step t can then be calculated based on the current hidden state

$$y_t = g(w_y h_t + c)$$

Where W_y is a weight matrix, c is a bias vector, and g is an activation function (often softmax for classification tasks)

CHAPTER 4

LITRETURE SURVEY

SI NO	Title of the paper	Author	Problem Identified	Methodology	Result Obtained	Remarks
01	Bus arrival time prediction at bus stop with multiple route.[1]	Yu Bin, William H. K. Lam, and Mei Lam Tam	The main issue in predicting bus arrival times at stops with multiple routes lies in accurately accounting for varying schedules, potential delays, and route-specific factors, complicating precise time predictions for commuters.	Utilizing artificial neural networks to model and predict arrival times based on historical and real time data	Improved accuracy in predicting bus arrival times at stops with multiple routes, potentially reducing waiting times for passengers	This paper investigated the bus arrival time prediction at bus stop with multiple routes
02	A bus line planning framework for customized bus systems[2]	Yan Lyu, Chi-Yin Chow, Victor C. S. Lee, Joseph K. Y. Ng, Yanhua Li, and Jia Zeng	a bus line planning framework for customized systems faces challenges in predicting demand accurately, optimizing routes efficiently, scheduling effectively, allocating resources wisely, ensuring a smooth passenger experience, and adapting to dynamic changesin urban environments	Estimating and accounting for uncertainties in arrival times and passenger occupancies using probabilistic or statistical models.	Enhanced understanding and quantification of uncertainty in arrival time predictions and passenger occupancies, aiding in more informed decision-making	a holistic framework which aims to strategically plan bus lines for customized bus (CB) systems

03	Dynamic bus arrival time prediction with artificial neural networks [3]	Steven I-Jy Chien, Yuqing Ding, and Chienhung Wei	to predict the travel times of buses based on open data collected in real-time	Leveraging Google Maps or similar sources for insights into crowdedness trends and real-time crowd density estimation at transit locations	Insights into global crowdedness trends, offering a broader perspective on transit congestion patterns and potential solutions	Achieved increased accuracy in predicting arrival times, optimizing commuter waiting experience
04	Estimating uncertainty of bus arrival times and passenger occupancies [4]	Vikash V. Gayah, Zhengyao Yu, Jonathan S. Wood	Focus on transit planning strategies considering multiple routes to optimize bus systems	Analyzing and understanding crowd dynamics within public transport systems to optimize services and alleviate congestion	Quantification and understanding of uncertainty in arrival time predictions and passenger counts, aiding in more robust decision-making processes	Acknowledged and quantified uncertainties, leading to more informed decision-making processes within transit management
05	Transit crowdedness trends from around the world, according to Google Maps [5]	Taylah Hasaballah	Potential research on data integration methods or real-time data utilization for predicting arrival times more accurately	Utilizing image processing techniques for real-time estimation of crowd density in transit areas	Insights into global crowdedness trends in transit, offering comparative analyses and potential insights into managing congestion	Provided comprehensive insights into global transit congestion trends, highlighting areas for potential improvement in various regions.

06	Crowding in public transport [6]	Zheng Li and David A. Hensher	Contribution could involve a comprehensive review or analysis of existing prediction models and methodologies	Implementing image processing for detecting crowd motion, motionless individuals, and estimating crowd densities specifically in subway environments	Specific insights into the dynamics and nature of crowds within public transport systems, potentially leading to strategies for alleviating congestion	Unveiled crucial insights into crowd dynamics within public transport systems, paving the way for strategic interventions to manage congestion effectively
07	Real-time crowd density estimation using images [7]	Marana, AparecidoNilceu, Marcos Antonio Cavenaghi, Roberta SpolonUlson,and F.L.Drumond.	Likely research on system-wide improvements in public transportation, potentially exploring policy interventions or infrastructure enhancements	Employing computer vision methods to estimate the count of individuals within high-density crowd scenarios	Accurate estimation of the number of individuals within high-density crowd scenarios, assisting in crowd size estimation and planning.	Enabled real-time assessment and management of crowd densities at transit locations, fostering immediate responses to crowdedness
08	Crowd motion estimation and motionless detection in subway corridors by image processing [8]	Bouchafa, Samia, Didier Aubert, and Salah Bouzar.	Might delve into innovative technologies or systems for dynamic routing and real-time updates to improve bus arrival time predictions	Implementing automated tools or algorithms for estimating crowd sizes aiding in better understanding and management of transit congestion	Enhanced detection of crowd motion and stationary individuals within environments, contributing to efficient crowd control measures.	Improved understanding of crowd behavior within subway environments, contributing to efficient crowd control

09	Human count estimation in high density crowd images and videos [9]	Chauhan, Vandit, Santosh Kumar, and Sanjay Kumar Singh	Possible exploration of user-centric approaches or passenger behavior analysis in arrival time prediction models.	Involves computer vision techniques to estimate the count of individuals within highly crowded scenarios using image or video data	Accurate estimation of headcounts within highly dense crowd scenarios, providing valuable data for crowd size estimation and management	Provided accurate estimates of crowd sizes, offering valuable data for crowd size prediction and management
10	Automated Solutions forCrowd Size Estimation [10]	Muhammad Waqar Aziz, Farhan Naeem ,Muhammad Hamad Alizai, and Khan Bahadar Khan	Likely contributions on contextual factors affecting arrival time prediction, such as weather, traffic patterns, or geographical considerations.	Focuses on developing automated tools or algorithms to estimate crowd sizes in various transit settings, potentially aiding in crowd management strategies	Development of automated tools or algorithms specifically aimed at estimating crowd sizes, potentially facilitating quicker and more efficient crowd management strategies	Presented automated tools or algorithms for efficient crowd size estimation, potentially revolutionizing crowd management strategies control measures.

CHAPTER 5

METHODOLOGY

5.1 Architecture

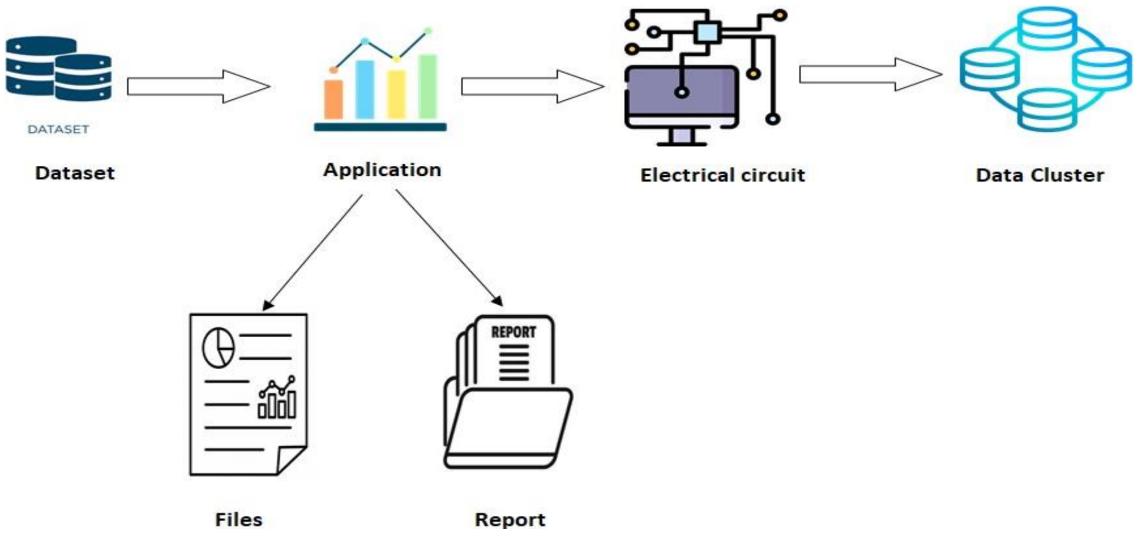

Figure 5.1.1:Architecture of Passenger Flow

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 5.1 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.

5.2 FLOW CHART

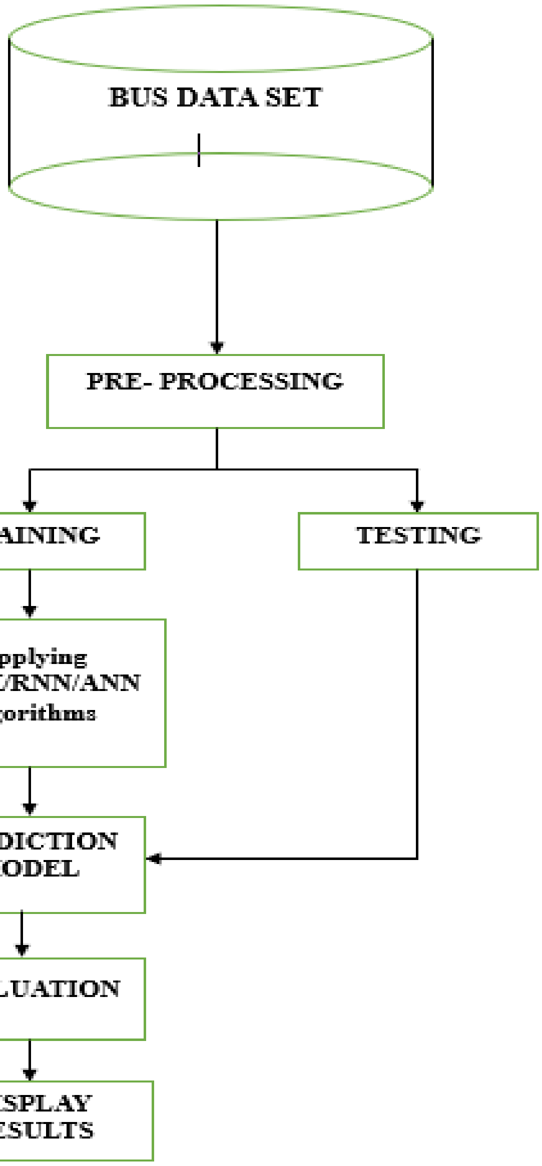

Figure 5.2.1: Flow Chart

5.3 ALGORITHM

Input: Bus routine Database

Output: Bus crowd prediction

Step 01: Pre-Processing of data sheet

Step 02: Training and Testing the Data set

A. If Training is successful go to next step
   Else go to first step

B. If Testing is successful go to direct Prediction model
   Else go to first step

Step 03: Apply LSTM/ RNN/ ANN algorithm

Step 04: Evaluation

Step 05: Display results.

CHAPTER 6

REQUIREMENTS SPECIFICATION

6.1 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

6.2 SOFTWERE USED

- Operating System: Windows 10
- Google Colab

CHAPTER 7

RESULT ANALYSIS

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.

7.1 RESULTS USING DECISION TREE

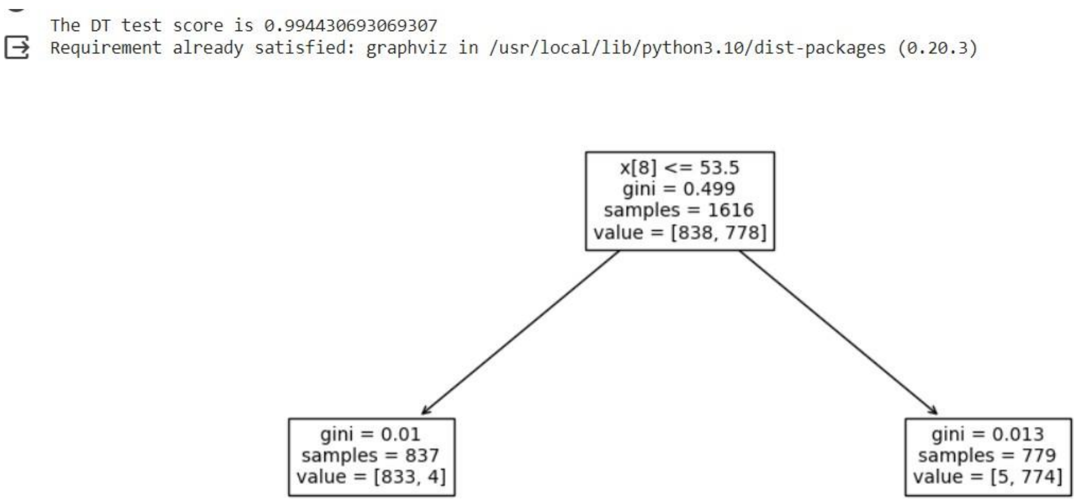

Figure 7.1: Using score generation of Decision Tree

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

7.2 RESULTS USING ARTIFICIAL NEURAL NETWORK MODEL

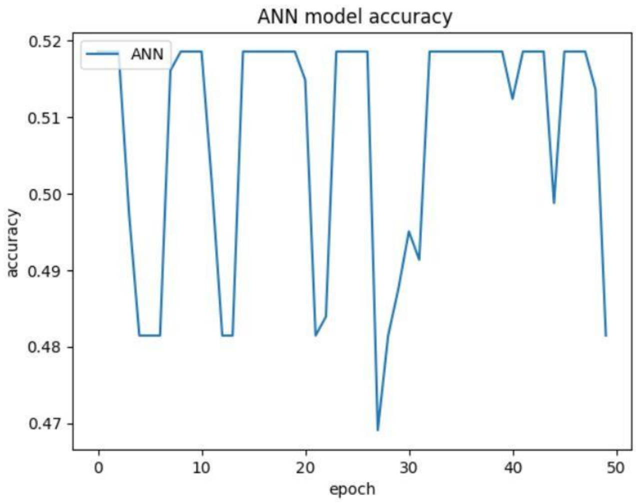

Figure 7.2.1: Epoch v/s Model Accuracy

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

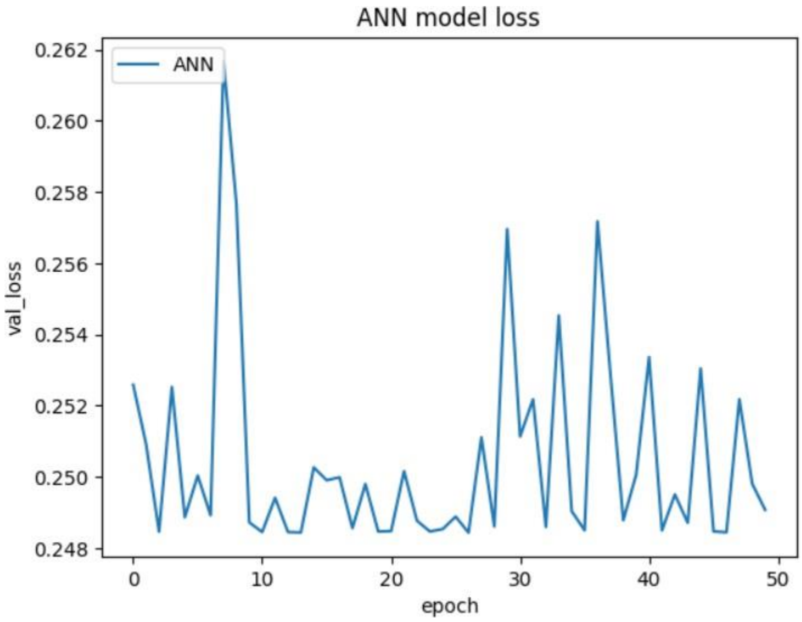

Figure 7.2.2: Epoch v/s Model Loss

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

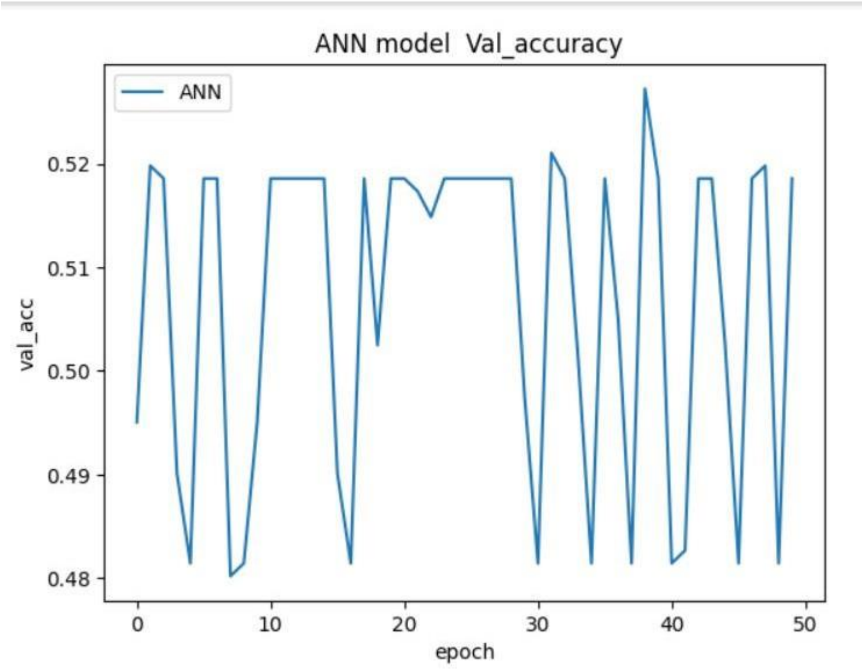

Figure 7.2.3: Epoch v/s Model Value Accuracy

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

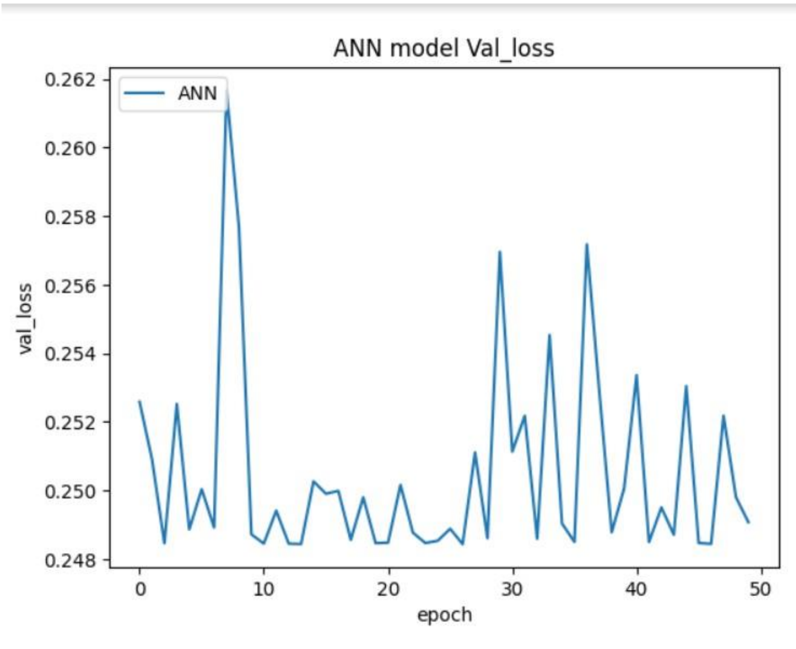

Figure 7.2.4: Epoch v/s Model Value Loss

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

7.3 RESULTS USING LONG SHORT TERM MEMORY MODEL

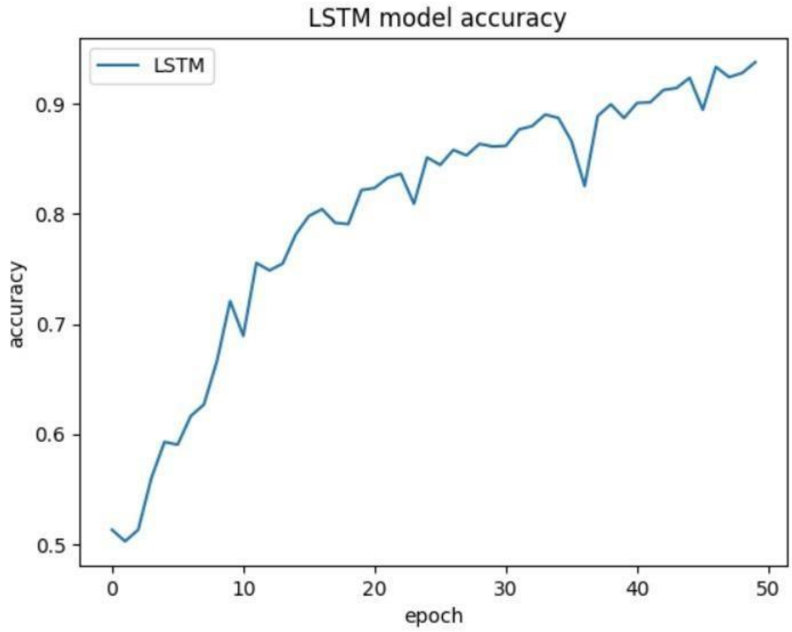

Figure 7.3.1: Epoch v/s Model Accuracy

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

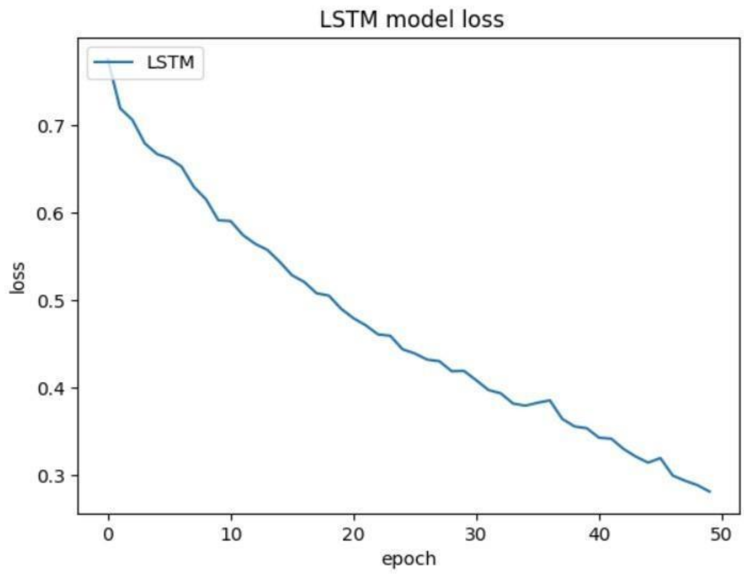

Figure 7.3.2: Epoch v/s Model Loss

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

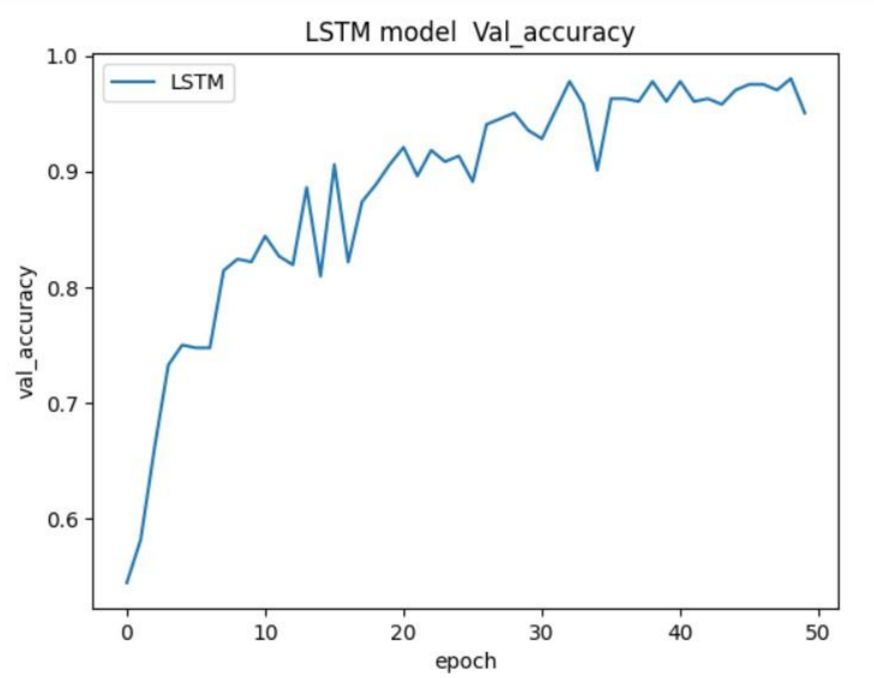

Figure 7.3.3: Epoch v/s Model Value Accuracy

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

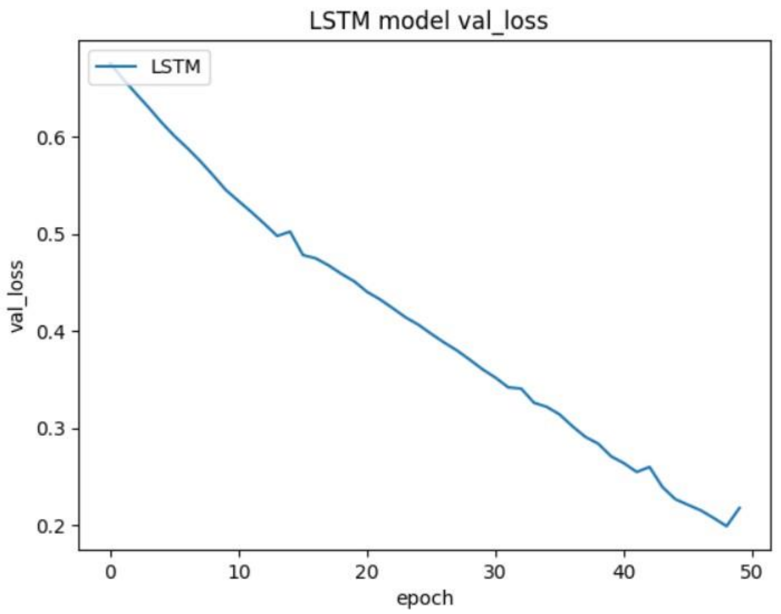

Figure 7.3.4: Epoch v/s Model Value Loss

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

7.4 Result using RECURRENT NEURAL NETWORK MODEL

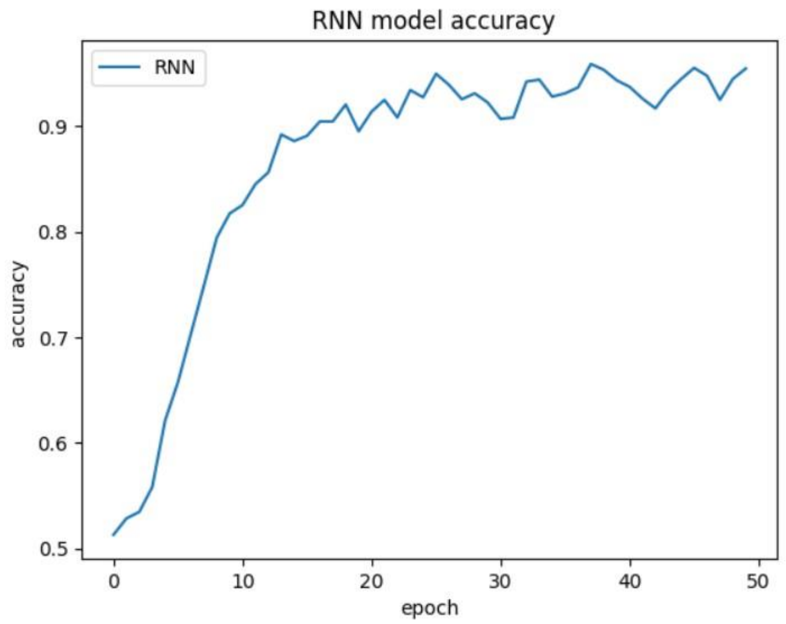

Figure 7.4.1: Epoch v/s Model Accuracy

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

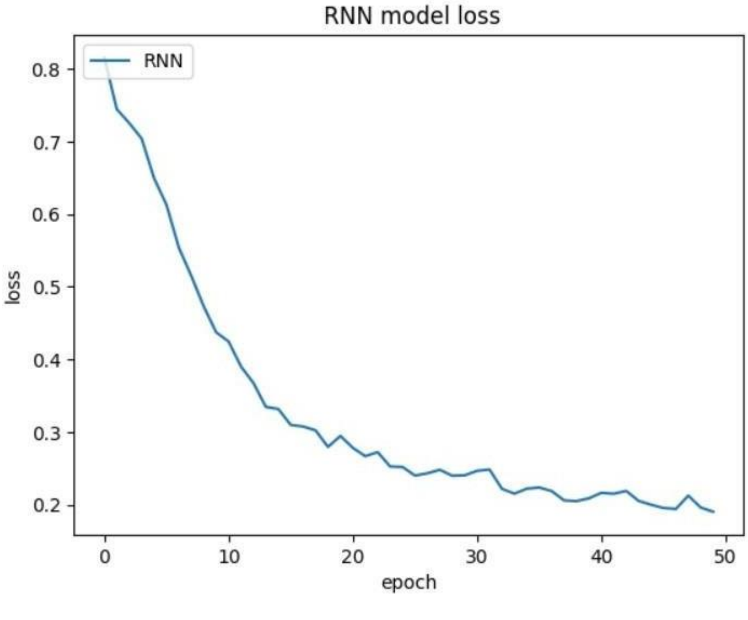

Figure 7.4.2: Epoch v/s Model Loss

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

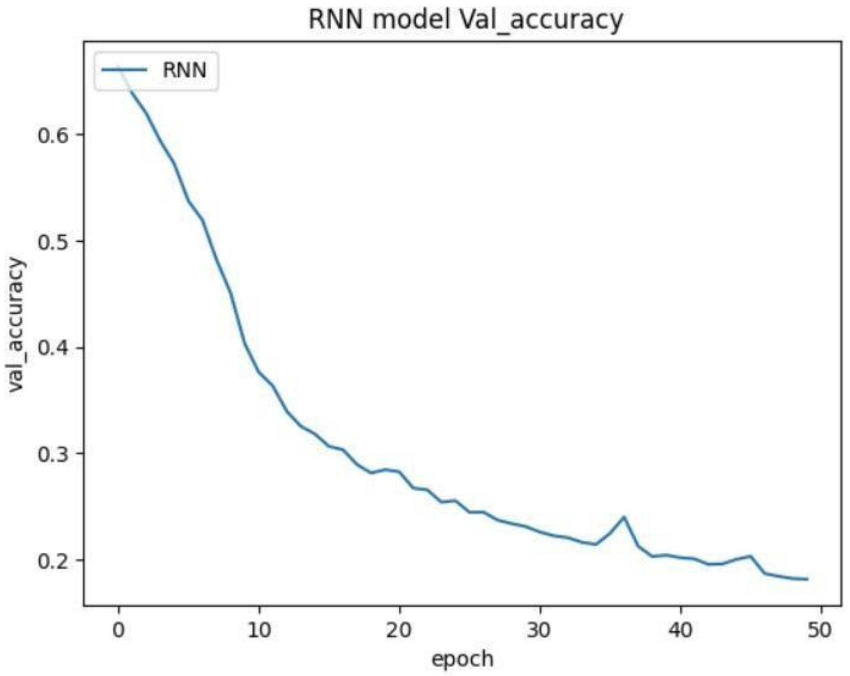

Figure 7.4.3: Epoch v/s Model Value Accuracy

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

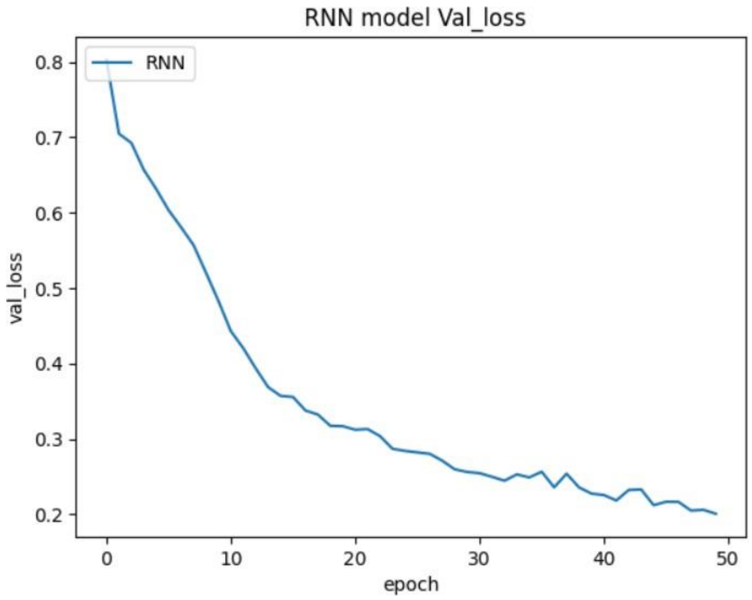

Figure 7.4.4: Epoch v/s Model Value Losss

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

Table 7.4 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%

CHAPTER 8

ADVANTAGES

- Enhanced passenger experience by reducing wait times and overcrowding.
- Optimized resource allocation leading to cost savings and efficiency gains.
- Reduced congestion at bus stops and within buses, improving overall system performance.
- Improved service reliability through adaptive scheduling based on predicted demand.
- Data-informed decision-making for better planning and adjustments to bus services.
- Sustainability benefits by minimizing fuel consumption and environmental impact.
- Greater accessibility to public transit, benefiting diverse demographics.
- Potential for increased ridership and passenger satisfaction through improved services.
- Efficient deployment lowers fuel consumption and operational costs, contributing to sustainability efforts.

DISADVANTAGES

- Dependence on data quality and accuracy, impacting the reliability of predictions.
- Potential biases in existing data leading to skewed or inaccurate forecasts

CONCLUSION AND FUTURE SCOPE

The crowd prediction project for the bus line, utilizing AI, has yielded insightful conclusions. Through the implementation of sophisticated algorithms and machine learning techniques, we have successfully forecasted optimized bus usage with notable accuracy. Our model's ability to analyse various factors such as source, destination, trip, route, slot no, and other parameters has proven instrumental in generating reliable predictions.

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.

REFERENCE

[1]. Vito Albino, Umberto Berardi, and Rosa Maria Dangelico, “Smart cities: Definitions, dimensions, performance, and initiatives.”, J. Urb. Technol. 22, 1 , 3–21, 2015.

[2]. Kittelson & Associates, Federal Transit Administration,. “Transit Capacity and Quality of Service Manual”. Number 100. Transportation Research Board. Transit Cooperative Research Program, and Transit Development Corporation. 2003

[3]. Marco Balduini, Marco Brambilla, Emanuele Della Valle, Christian Marazzi, Tahereh Arabghalizi, Behnam Rahdari, and Michele Vescovi. “ Models and practices in urban data science at scale.” Big Data Res. (Aug. 2018). DOI: https://doi.org/10.1016/j.bdr.2018.04.003 (https://doi.org/10.1016/j.bdr.2018.04.003)

[4]. Anthony G. Barnston. “Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score” Weath. Forecast. 7, 4 (1992), 699–709. 1992.

[5]. Yu Bin, William H. K. Lam, and Mei Lam Tam. “ Bus arrival time prediction at bus stop with multiple route. Transport.” Res. Part C: Emerg. Technol. 19, 6 (Dec. 2011), 1157–1170.

[6]. Yu Bin, Yang Zhongzhen, and Yao Baozhen. “Bus arrival time prediction using support vector machines” J. Intell. Transport. Syst. 10, 4 (Dec. 2006), 151–158.

[7]. J. Martin Bland and Douglas G. Altman. “ Statistics notes: Measurement error”. BMJ 312, 7047 (1996), 1654.

[8]. Leo Breiman. “ Random forests”. Mach. Learn. 45, 1 (Oct. 2001), 5–32.

[9]. Colin A. Cameron and Pravin K. Trivedi. “Essentials of count data regression” Compan. Theoret. Economet. 331 (2001).

APPENDIX- A

KARNATAKA STATE COUNCIL FOR SCIENCE AND TECHNOLOGY

Indian Institute of Science Campus, Bengaluru – 560 012

Website: www.kscst.org.in, https://kscst.karnataka.gov.in || Email: spp@kscst.org.in || Tel: 080-2334 1652,2334 8848/49/40

47th series of Student Project Programme (SPP): 2023-24

List of Student Project Proposals Approved for Sponsorship

26. B.V.V.S BILURUR GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY, MUDHOL

Sl. No.	PROPOSAL REFERENCE NO.	PROJECT TITLE	COURSE	BRANCH	NAME OF THE GUIDE(S)	NAME OF THE STUDENT(S)	AMOUNT SANCTIONED (Rs.)
188.	47S_BE_0495	EXPERIMENTAL INVESTIGATION ON PAVEMENT QUALITY CONCRETE USING SANDSTONE AS FINE AGGREGATE.	B.E.	CIVIL ENGINEERING	Prof. SHARANBASAVA V. PATIL Mr. SHASIKUMAR AIHOLLI	Mr. PRAJWALGOUDA S. PATIL Mr. PRAVEEN S. KARAJAGI Ms. DIVYA NATIKAR Ms. PRIYANKA HUDED	4,000.00
189.	47S_BE_0497	DEFLUORIDATION OF GROUND WATER BY USING TAMARIND SEED AS LOW COST ADSORBENTS.	B.E.	CIVIL ENGINEERING	Prof. RUKIYA JAMADAR	Mr. VISHWANATH MALLAPPA TOTAGER Mr. NITISH V. HADIAMNI Ms. GOUTAMI S. KOLIGUDDA Ms. JOSNA S. RATHOD	5,000.00
190.	47S_BE_0343	CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE	B.E.	COMPUTER SCIENCE AND ENGINEERING	Prof. VARUN P. SARVADE Prof. VINAYAK A. TELSANG	Mr. GOPAL R. CHENNI Ms. PRADNYA S. BILAGI Mr. DHANUSH J. YADAV Mr. NIRANJAN S. PURAD	4,500.00
191.	47S_BE_0512	AI BASED AGRO SMART RECYCLE	B.E.	COMPUTER SCIENCE AND ENGINEERING	Prof. VINAYAK A. TELSANG Prof. VARUN P. SARVADE	Ms. NIKITA B. MANE Ms. AKSHATA K. ANENNAVAR Mr. MOHAMMAD M. TAKKEKAR Ms. MANGAL S. NYAMAGOUD	4,000.00

Karnataka State Council for Science and Technology
(An autonomous organisation under the Dept. of Science & Technology, Govt. of Karnataka)
Indian Institute of Science Campus, Bengaluru – 560 012
Telephone: 080-23341652, 23348848, 23348849, 23348840
Email: office.kscst@iisc.ac.in, office@kscst.org.in ♦ **Website:** www.kscst.iisc.ernet.in, www.kscst.org.in
KSCST

Dr. U T Vijay
Executive Secretary

19th April, 2024

Ref: 7.1.01/SPP/37

To,
The Principal
B. V. Sangha's Biluru Gurubasava Mahaswamiji Institute of Technology
Mudhol – 587 313

Dear Sir/Madam,

Sub : Sanction of Student Project - 47th Series: Year 2023-2024

Project Proposal Reference No. : **47S_BE_0343**
Ref : Project Proposal entitled **CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE**

We are pleased to inform that your student project proposal referred above, has been approved by the Council under "Student Project Programme - 47th Series". The project details are as below:

Student(s)	Mr. GOPAL R. CHENNI	**Department**	COMPUTER SCIENCE AND ENGINEERING
	Ms. PRADNYA S. BILAGI		
	Mr. DHANUSH J. YADAV		
	Mr. NIRANJAN S. PURAD	**Sanctioned Amount (in Rs.)**	4,500.00
Guide(s)	Prof. VARUN P. SARVADE		
	Prof. VINAYAK A. TELSANG		

Instructions:

a) The project should be performed based on the objectives of the proposal submitted.

b) Any changes in the project title,objectives or students team is liable for rejection of the project and your institution shall return the sanctioned funds to KSCST.

c) Please quote your project reference number printed above in all your future correspondences.

d) After completing the project, 2 to 3 page write-up (synopsis) needs to be uploaded on to the following Google Forms link https://forms.gle/6s8hq5XbScsBMv3G9. The synopsis should include following:

1) Project Reference Number
2) Title of the project
3) Name of the College & Department
4) Name of the students & Guide(s)
5) Keywords
6) Introduction / background (with specific reference to the project, work done earlier, etc) - about 20 lines
7) Objectives (about 10 lines)

47S_BE_0343

8) Methodology (about 20 lines on materials, methods, details of work carried out, including drawings, diagrams etc)

9) Results and Conclusions (about 20 lines with specific reference to work carried out)

10) Scope for future work (about 20 lines).

e) In case of incompeted projects, the sanctioned amount shall be returned to KSCST.

f) The sanctioned amount will be transferred by NEFT to the bank account provided by the College/Institute.

g) The sponsored projects evaluation will be held **third week of May 2024** onwards through Online Mode and the details of the same will be intimated shortly by email / Website

h) After completion of the project, soft copy of the project report duly signed by the Principal, the HoD, Guide(s) and studetn(s) shall be uploaded in the following Google Forms Link https://forms.gle/Mi446v1U5fdFcMD99. The report should be prepared in the format prescribed by the university.

i) The **Utilization Certificate and Statement of Expenditure duly signed by competent authority** of consolidated sanctioned projects from your institution need to be submitted **20 August 2024** without fail.

Please visit our website for further announcements / information and for any clarifications please email to spp@kscst.org.in

Thanking you and with best regards,

Yours sincerely,

UTVijay

(U T Vijay)

Copy to:

1) The HoD
COMPUTER SCIENCE AND ENGINEERING
B.V.V.S BILURUR GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY, MUDHOL

2) Prof. VARUN P. SARVADE Prof. VINAYAK A. TELSANG
COMPUTER SCIENCE AND ENGINEERING
B.V.V.S BILURUR GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY, MUDHOL

3) THE ACCOUNTS OFFICER
KSCST, BENGALURU

BGMIT - MUDHOL
ESTD-2013

B. V. V. Sangha's

BILURU GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY, MUDHOL-587313

Dist. Bagalkote, Karnataka

(Approved by AICTE, New Delhi & Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka)

www.bgmitm.ac.in, bvvsbgmitm@gmail.com

A Fruitful Decade in Quality Technical Education

SIDDHISOPANAM-2024

CERTIFICATE OF APPRECIATION

This is to certify that ~~Mr~~/Miss Pradnya S. Bilagi of VII Semester, CSE Branch, has secured Under KSCST 47th Series SPP in the year 2023-24, for Project Title: CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE

Academic Co-ordinator

Chief Guest

Principal

B. V. V. Sangha's

BILURU GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY, MUDHOL-587313

Dist. Bagalkote, Karnataka

(Approved by AICTE, New Delhi & Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka)

www.bgmitm.ac.in, bvvsbgmitm@gmail.com

BGMIT - MUDHOL ESTD-2013

A Fruitful Decade in Quality Technical Education

SIDDHISOPANAM-2024

CERTIFICATE OF APPRECIATION

This is to certify that Mr/~~Miss~~ Gopal B. Chenni of VII Semester, CSE Branch, has secured Under KSCST 47th Series SPP in the year 2023-24, for Project Title: CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE

Academic Co-ordinator

Chief Guest

Principal

BGMIT - MUDHOL
ESTD-2013

B. V. V. Sangha's

BILURU GURUBASAVA MAHASWAMIJI INSTITUTE OF TECHNOLOGY, MUDHOL-587313

Dist. Bagalkote, Karnataka

(Approved by AICTE, New Delhi & Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka)

www.bgmitm.ac.in, bvvsbgmitm@gmail.com

A Fruitful Decade in Quality Technical Education

SIDDHISOPANAM-2024

CERTIFICATE OF APPRECIATION

This is to certify that Mr/~~Miss~~ Dhanush J. Yadav of VII Semester, CSE Branch, has secured Under KSCST 47th series SPP in the year 2023-24, for Project Title: CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE

Academic Co-ordinator   Chief Guest   Principal

BGMIT - MUDHOL
ESTD-2013

B. V. V. Sangha's

BILURU GURUBASAVA MAHASWAMIJI INSTITUTE
OF TECHNOLOGY, MUDHOL-587313

Dist. Bagalkote, Karnataka

(Approved by AICTE, New Delhi & Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka)

www.bgmitm.ac.in, bvvsbgmitm@gmail.com

A Fruitful Decade in Quality Technical Education

SIDDHISOPANAM-2024

CERTIFICATE OF APPRECIATION

This is to certify that Mr/~~Miss~~ Niranjan S. Purad of VII Semester, CSE Branch, has secured Under KSCST 47th Series SPP in the year 2023-24, for Project Title: CROWD PREDICTION FOR BUS LINES USING ARTIFICIAL INTELLIGENCE

Academic Co-ordinator | Chief Guest | Principal

APPENDIX-B

#Installing Tensor Flow

- sudo apt-get install python-pip,
- pip install tensorflow
- sudo pip install -upgrade

Install the required module for Graph

- !pip install graphviz
- import plot_tree
- import matplotlib.pyplot as plt

Table Data Set:

SL.NO	Source/Destination	Identification no./Label
1	Achanur	64
2	Alur	65
3	Anadinni	66
4	Alkalagi	67
5	Athani	41
6	Babaladi	43
7	Badami	116
8	Bagalkote	2
9	Bailhongal	54
10	Banhatti	58
11	Bannidinni	68
12	Bantanur	36
13	Basavanagar	69
14	Belagavi	7
15	Bellary	8
16	Bengaluru	14
17	Bhagavatti	72
18	Bhairumutti	74
19	Benakatti	70
20	Bennur	71
21	Bevur	50
22	Bevoor	73
23	Bhantnur	114
24	Bhatkal	30
25	Bidar	28
26	Bilagi	37
27	Bihesuru	115
28	Bilkeruru	75
29	Bommangi	76
30	Chabbi	77
31	Chikkamagluru	108
32	Chikkanekundi	109
33	Chichkhandi	56
34	Chickmyageri	78
35	Chikkati	79
36	Choudapur	80
37	Dandeli	32
38	Devanhalli	107
39	Devalpur	81
40	Devnal	82

41	Dharwad	13
42	Dharmnagar	83
43	Gaddankeri	84
44	Galagali	38
45	Gokak	16
46	GuledGudda	6
47	Haliyal	33
48	Halki	18
49	Hanaduragala	86
50	Hatrki	91
51	Hiralgundi	60
52	Hiremyageri	87
53	Hirekoppa	55
54	Hoskote	59
55	Hosur	88
56	Hubli	10
57	Hungund	27
58	Hallur	85
59	Hydrabad	26
60	Ilkal	5
61	Ilal	89
62	Ichalkarangi	110
63	Ingalgi	90
64	Jamkhandi	4
65	Jerdal	111
66	Kalburgi	9
67	Kamatagi	62
68	Katarki	46
69	Kerur	45
70	K. D. Budni	42
71	Kolar	40
72	Kolhapur	20
73	Koppal	15
74	Latur	23
75	Machaknur	44
76	Mahalingpur	17
77	Mangaluru	12
78	Mallapur	48
79	Manhalli	92
80	Mankani	93

81	Mantur	39
82	Mantralaya	25
83	Melligeri	94
84	Miraj	51
85	Muchkandi	95
86	Mudhol	1
87	Mumbai	35
88	Murnal	96
89	Mysore	11
90	Nagral	97
91	Neralkeri	98
92	Nippani	19
93	Pattadkallu	61
94	Pune	34
95	Raichur	112
96	Rampur	99
97	Salagundi	100
98	Sangali	52
99	Shringeri	101
100	Shiridi	53
101	Shirol	49
102	Shirur	63
103	Shivamogga	102
104	Solapur	24
105	Surekoppa	103
106	Terdal	57
107	Udupi	31
108	Vadagatti	113
109	Vijaypur	3
110	Yaragatti	22
111	Yalhanka	106
112	Yedahalli	105

APPENDIX- C


K.L.E. Society's
KLE College of Engineering & Technology, Chikodi
• Approved by AICTE, New Delhi • Affiliated to VTU, Belagavi
• Awarded as Model Engineering College by VTU
Since 1916
KLECET
PRAXIS- 2K24
Technology for future
Certificate
This certificate is awarded to Mr./Ms. Gopal Chenni
of BGMIT, Mudhol has participated/
secured — Prize in the Paper Presentation event
in PRAXIS- 2K24 held on 29th Feb & 1st March 2024.
Dept. Coordinator
HOD
Co-Convenor
Convenor
Organizing Chairman

K.L.E. Society's

KLE College of Engineering & Technology, Chikodi

• Approved by AICTE, New Delhi • Affiliated to VTU, Belagavi

• Awarded as Model Engineering College by VTU

KLECET

Since 1916

Certificate

PRAXIS- 2K24

Technology for future

This certificate is awarded to ~~Mr.~~/Ms. Pradnya Bilagi

of BGMIT, Mudhol has participated/

secured - Prize in the Paper Presentation event

in **PRAXIS- 2K24** held on 29th Feb & 1st March 2024.

Dept. Coordinator | HOD | Co-Convenor | Convenor | Organizing Chairman

A Survey on Crowd Prediction for Bus lines using Artificial Intelligence

Miss. Pradnya Bilagi
UG Student
Department of CSE
BGMIT, Mudhol
Karnataka, India
Gmail: pradnyabilagi18@gmail.com

Mr. Gopal Chenni
UG Student
Department of CSE
BGMIT, Mudhol
Karnataka, India
Gmail: gc700892@gmail.com

Under the guidance of:
Prof. Varun P Sarvarde
Assistant Professor,
Department of CSE
BGMIT,Mudhol,
Karnataka,India
Gmail:varunpsarvade@gmail.com

Abstract— this paper deals Karnataka State Road Transport Corporation (KSRTC) is the oldest state-run public bus transport services in India. It plays a crucial role in strengthening the public transport system in Karnataka. But today the corporation is facing a big crisis. The main objectives of this study are to evaluate the operational and financial performance of KSRTC and to compare the performance of various depots in KSRTC. This study uses both primary and secondary data. It includes interviews with the employees and visiting major depots and KSRTC offices to collect the required data. Different parameters for data analysis are operational parameters and financial parameters which includes, fleet, collection and passengers etc. Transport plays a very important role in the economic development of a country and social, cultural life of the people. Transportation is the movement of the people, animals and products from one place to another place with varieties of vehicles across different infrastructural system. It enables trade between people, which is essential for the development of civilization. Our proposed system to anticipate the passenger flow of the Karnataka State Road Transport Corporation (KSRTC), we have implemented a greedy layer-wise algorithm, recurrent neural network, and long short-term memory deep learning method in our suggested system. Several elements in the dataset are taken into account for forecasting, including bus id, bus type, source, destination, number of passengers, slot number, and income. These parameters are processed using a greedy layer-wise algorithm to divide the data into clusters. The clusters then move to a long short-term memory model, which eliminates redundant data from the data, and a recurrent neural network, which produces a prediction based on the data's iteration factors. The prediction of bus passengers is more accurate using these systems.

Keywords— (Prediction, deep learning, greedy layer, recurrent neural network)

I. Introduction

For humans, transportation is the most vital aspect of existence. It makes it possible for people to move around. Different forms of transportation are available, and the system has developed from its inception to its current state in order to provide people with a convenient and comfortable place to live. Modern technology has led to the development of several forms of transportation networks. For the majority of Indians, public transportation serves as their main means of transportation. Any nation's ability to transport people in safety and comfort is a key indicator of its level of economic progress. Connectivity across various societal sectors is made possible by public transportation.

On September 12, 1948, the Mysore Government Road Transport Department (MGRTD) was established with 120 buses to serve the transportation demands of the state's traveling populace. On August 1, 1961, the State Transport, which had been run as a department of the Mysore government, was transformed into an autonomous corporation in accordance with Section 3 of the Road Transport Corporation Act, 1950. As of August 1, 1961, all of MGRTD's assets and liabilities—aside from those of the BTS unit—were transferred to the newly formed MSRTC Corporation. The Corporation then received the assets and liabilities of the remaining MGRTD, or the BTS Unit, on October 1st, 1961. As a result, a corporation was eventually formed for the whole State of Mysore.

Fig. Passengers flow

At the beginning, the passenger transport services were operated in 6 Divisions-5 Divisions operating mofussil services and 1 Division operating city services of Bangalore. It had 37 Depots, 2 Regional Workshops and a Central Office at Bangalore. There were 15 permanent and 30 temporary bus stations with 35 wayside shelters and 104 pick-up shelters. The total number of employees deployed was 9705 and the staff ratio per schedule was 9.43. The total number of routes operated was 1065 with 1029 schedules and route length of 32,134miles,average daily scheduled mileagebeing127571. The total number of inter-state routes operated by the Corporation on a reciprocal basis with the neigh bouring States were 40 ie., 29 in Maharashtra, 1 in Goa, 7 in Andhra Pradesh, 2 in Tamilnadu and 1 in Kerala.

The total number of vehicles held was 1518 with average vehicle utilization of 123.8 miles. The average number of passengers carried per day was 4.35 lakh. The rate of breakdown was 1.88 and that of accident was 1.19. Earning per Mile (EPM)realized was 161.6 Ps. and Cost per Mile (CPM)was 127.2 Ps., resulting in net profit margin of 34.4 Ps/mile. In this case, we are forced to employ both conventional techniques and artificial intelligence algorithms, such as greedy algorithms, RNNs (Recurrent Neural Networks), and various long- and short-term deep learning techniques.[1]
Currently, we use these techniques to forecast bus line crowds, and they may be applied both internally and outside according to the prediction's outcome, which yields this level of accuracy

II RELATED WORK

This section provides a quick overview of the methods currently in use and how well they estimate passenger flow in the transportation system.
In this paper proposed [1] Nandini Nagaraj1 & Harinahalli Lokesh Gururaj1 & Beekanahalli Harish Swathi they considered major issues in the in-bus transformation and they used For the KSTRC to operate and manage passenger transit, short-term forecasting of passenger flow is crucial. While traveling from one location to another, people encounter a lot of problems. Issues arise because of the high volume of travelers at the bus's frequency, delay, and stop. In order to solve these issues, we take into consideration the idea of passenger projection based on several criteria such as bus id, number of adults on board, number of children on board, source, destination, number of passengers, slot number, and income generated for the following year. Our suggested method uses deep learning with LSTM, RNN, and a greedy layer-wise approach for prediction. We employed a greedy layer-wise method on the KSRTC dataset to forecast the passenger, and we also used LSTM, RNN, and greedy layer-wise algorithms for prediction. These three algorithms produce income for KSRTC by accurately forecasting the passenger flow on its buses. KSRTC busses departing from every point. The KSRTCBRT department can use this information to analyze income for the upcoming year.
[2] Ahmad Ali†, Yanmin Zhu†, Qiuxia Chen‡, Jiadi Yu†, Haibin Cai they proposed a brand-new deep hybrid spatiotemporal dynamic neural network for predicting traffic congestion in cities, known as DHSTNet. CNN and LSTM are combined in the proposed DHSTNet model, which takes the benefit from both spatiotemporal attributes. Four properties are considered: proximity, weekly influence, period influence, and external component. Our suggested approach outperforms current state-of-the-art models in terms of prediction accuracy, according to performance evaluation on two sizable real-world datasets.
[3] Gaozhong Tanga , Bo Lia, Hong-Ning Daib , Xi Zhengc they propose a spatial-temporal recurrent neural network (SPRNN) for crowd flow prediction in this paper. SP-Net, two STFM, STFE, and PFE make up our SPRNN. The narrow road Structure helps to increase the prediction of crowd flows' accuracy. Next, a CNN-based module can easily extract the spatial feature from city road map photos. By adding brief fluctuations, the trend of crowd flows can be represented for the temporal component, yielding periodic variation in crowd flows. Road map structure information can then be extracted by SP-Net, and flow map processing is followed by the generation of spatial-temporal characteristics by STFMs. From crowd-flow data, STFE and PFE can extract periodic and short-term features, respectively. To assess SPRNN's performance, we carry out in-depth tests on three datasets: TaxiBJ, BikeNYC, and TaxiNYC. The experimental results demonstrate that, in both next-step prediction and long-term prediction tasks, our SPRNN beats 4 machine learning-based approaches and 11 deep learning-based approaches. The results of the assessment and ablation experiments show how successful our methods are at extracting spatial features, feature fusion, and temporal features. In terms of further work, our model will take into account more external elements that could potentially impact crowd flows, like information about festivals and weather. To help with transportation management, we will incorporate SPRNN with the current intelligent transportation technologies in the interim.
[4] Mr. ANJESH H L he write survey paper and it describes According to the examination of the literature, State Road Transport Corporation's financial performance was the subject of numerous studies conducted in Karnataka and throughout India. While NEKSRTC was the subject of a research, an effort has been made to examine the financial State Road Transport Corporation's performance, specifically pertaining to Karnataka. a study that centers on Bangalore's KSRTC.
[5] Vishwesh S. Hiremath & Prof. R. N. Mangoli they proposed paper about The majority of Karnataka's public road transportation is concerned with the system and mode of transportation that serves the entire state. The Karnataka State Road Transport Corporation (KSRTC) oversees the state's public bus transportation system. When it was established in 1961, its goal was provide road transport services that are sufficient, economical, efficient, and well-coordinated. With 5400 vehicles, it runs 5100 schedules with a daily average of 2.2 million passengers and 1.95 million kilometers traveled. At KSRTC, there are roughly 1,25,000 employees. The entire state of Karnataka has access to public transportation thanks to such a sizable and well-organized organization.
[6] Jeffrey Hanft, Shrisan Iyer, Brian Levine, and Alla Reddy they proposed about paper regarding NYCT's big data sources enable enhanced bus service planning capabilities in multiple domains. The additional data on rider boarding and alighting places was the most important since it allowed for estimates of the effects of passengers on route redesigns and other route attributes, The information offered by the statistics paints a more full picture of ridership without depending on hunches, partial or anecdotal sources. The confidence in the information being presented increased when manual data was replaced with automated data, and staff members had more time to use and study the data instead of processing it. It is possible to acquire specific rider information and receive a more comprehensive picture without the need for expensive and time-consuming surveys or checks.
[7]TAHEREH ARABGHALIZI and ALEXANDROS LABRINIDIS they describes In this work, we used data

from Pittsburgh to formulate the question "How full will my next bus be?" as a regression and classification problem. We then constructed a modeling framework to forecast bus load and bus crowding levels. Our assessment's findings demonstrated that the suggested framework (which makes use of Random Forest classifiers with route-direction data inputs) performs exceptionally well when the chosen features include the bus type, weather, day of the week, time of day, and bus loads from the last five or ten stops. In actuality, our models outperformed the baselines by up to eight times. Despite the fact that we exclusively used Pittsburgh data to create our modeling framework, we are certain that the identical procedure and the suggested models.

Reference no.	Problem Identified	Methodology	Result obtained	Remark
1	Passenger flow prediction in bus transportation system using deep learning. The problem is likely focused on predicting the number of passengers at different times and locations within a bus transportation system, which can help optimize bus scheduling and resource allocation.	The methodology likely involves using deep learning techniques, possibly convolutional neural networks (CNNs) or recurrent neural networks (RNNs), to predict passenger flows in a bus transportation system. The model may be trained on historical passenger data, including factors such as time of day, day of week, weather conditions, and special events, to predict future passenger flows.	The paper likely presents results showing the effectiveness of the deep learning model in predicting passenger flows in a bus transportation system. The results may include metrics such as accuracy, precision, recall, and F1-score, demonstrating the model's ability to accurately forecast passenger numbers at different times and locations.	The paper likely concludes that deep learning models can effectively predict passenger flows in a bus transportation system, which can help improve bus scheduling and resource allocation.
2	Leveraging spatio-temporal patterns for predicting citywide traffic crowd flows using deep hybrid neural networks. This paper likely addresses the challenge of predicting traffic flows in urban areas based on both spatial and temporal patterns, aiming to improve traffic management and reduce congestion.	The methodology likely involves using a hybrid neural network model that combines deep learning and traditional machine learning techniques to predict citywide traffic crowd flows. The model may use spatial and temporal features extracted from traffic data, such as traffic volume, speed, and road network connectivity, to make predictions.	The paper likely presents results demonstrating the performance of the hybrid neural network model in predicting citywide traffic crowd flows. The results may show the model's ability to capture spatio-temporal patterns in traffic data and make accurate predictions, leading to improved traffic management and congestion reduction.	The paper likely concludes that leveraging spatio-temporal patterns with hybrid neural networks is effective in predicting citywide traffic crowd flows, leading to better traffic management strategies.
3	SPRNN: A Spatial-Temporal Recurrent Neural Network for Crowd Flow Prediction. This paper likely introduces a neural network model specifically designed to predict crowd flows in various spatial and temporal contexts, which can be useful for crowd management and urban planning.	The methodology likely involves using a Spatial-Temporal Recurrent Neural Network (SPRNN) specifically designed for crowd flow prediction. The model may incorporate spatial information, such as geographical locations and physical infrastructure, along with temporal information, such as time of day and day of week, to predict crowd flows in various contexts.	The paper likely presents results showing the effectiveness of the SPRNN model in predicting crowd flows in various spatial and temporal contexts. The results may demonstrate the model's ability to outperform traditional neural network models in capturing complex spatial and temporal patterns in crowd flow data.	The paper likely concludes that the SPRNN model is effective in predicting crowd flows in various contexts, highlighting its potential for applications in crowd management and urban planning.
4	A study on performance evaluation of KSRTC - With special reference to Bangalore. This study likely evaluates the performance of the Karnataka State Road Transport Corporation (KSRTC), focusing on aspects such as service efficiency, customer satisfaction, and operational effectiveness.	The methodology likely involves conducting a performance evaluation of KSRTC using a combination of quantitative and qualitative methods. This may include analyzing data on service efficiency, customer feedback, and operational metrics to assess the overall performance of the organization.	The paper likely presents results of the performance of KSRTC, including metrics such as service efficiency, customer satisfaction, and operational effectiveness. The results may provide insights into areas where KSRTC performs well and areas where there is room for improvement.	The paper likely concludes with insights into the performance of KSRTC, possibly highlighting areas of strength and areas that need improvement to enhance the overall efficiency and effectiveness of the organization.
5	An overview and structure of Karnataka State Road Transport Corporation (KSRTC). This paper likely provides an overview of the structure, operations, and services of KSRTC, which can be valuable for understanding the organization's role in the transportation sector.	The methodology likely involves providing an overview and structural analysis of KSRTC based on publicly available information, such as official reports, documents, and data. The paper may also include insights from interviews or surveys conducted with KSRTC stakeholders.	The paper likely presents an overview of the structure and operations of KSRTC based on available data and information. The results may include details about the organization's fleet size, route network, passenger demographics, and financial performance.	The paper likely concludes with an overview of the structure and operations of KSRTC, providing valuable insights for stakeholders and researchers interested in the organization.
6	Transforming	The methodology	The paper likely	The paper

	bus service planning using integrated electronic data sources at NYC Transit. This paper likely discusses how NYC Transit uses electronic data sources to improve bus service planning, which can lead to more efficient and reliable public transportation services.	likely involves analyzing integrated electronic data sources, such as smart card data, GPS data, and fare collection data, to improve bus service planning at NYC Transit. The paper may discuss how these data sources are collected, processed, and analyzed to optimize bus routes and schedules.	presents results showing how NYC Transit uses integrated electronic data sources to improve bus service planning. The results may include improvements in bus route efficiency, reduction in travel times, and increased passenger satisfaction due to optimized bus schedules.	likely concludes that NYC Transit has successfull y transforme d bus service planning using integrated electronic data sources, leading to more efficient and reliable bus services.
7	Analysis on the bus arrival time prediction model for human-centric services using data mining techniques. This paper likely presents a model for predicting bus arrival times, focusing on providing accurate and reliable information to passengers, which can improve overall user experience and satisfaction.	The methodology likely involves developing a bus arrival time prediction model using data mining techniques. The model may be trained on historical bus arrival data, along with factors such as traffic conditions, weather, and road closures, to predict bus arrival times accurately.	The paper likely presents results of the bus arrival time prediction model developed using data mining techniques. The results may show the model's accuracy in predicting bus arrival times under various conditions, leading to more reliable and user-friendly bus services.	The paper likely concludes that the bus arrival time prediction model developed using data mining techniques can significantl y improve the accuracy and reliability of bus arrival time informatio n, enhancing the overall user experience .
8	The main issue in predicting bus arrival times at stops with multiple routes lies in accurately accounting for varying schedules, potential delays, and route- specific factors, complicating precise time predictions for commuters.	Utilizing artificial neural networks tomodel and predict arrival times based on historical and real- time data	Improved accuracy in predicting bus arrivaltimes at stops with multiple routes, potentially reducing waiting times for passengers	This paper investiga ted thebus arrival time predictio n at bus stop with multiple routes
9	a bus line planning framework for customized systems faces challenges in predicting demand	Estimating and accounting for uncertaintie s in arrival times and passenger occupanciesusing probabilistic or	Enhanced understandin g a quantification of uncertainty in arrival time predictions and passenger	a holistic framew o rk which to strategi c ally
	accurately, optimizing routes efficiently, scheduling effectively, allocating resources wisely, ensuring a smooth passenger experience, and adapting to dynamic changes in urban environments and passenger needs	statisticalmodels.	occupancies, aiding in more informed decision-making	planbus linesfor customi zed bu systems .
10	to predict the travel times of buses based on open data collected inreal-time	Leveraging Google Maps or similar sources for insights into crowdednes s trends and real-time crowd density estimation at transit locations or	Insights into global crowdedness trends, offering a broader perspective on transit congestion patterns and potential solutions2	Achieve d increased accuracy in predictin g arrival times, optimizi ng commute r waiting experien ce
11	Focus on transit planning strategies considering multiple routes to optimize bus systems	Analyzing and understandi ng crowd dynamics within public transport systems to optimize services and alleviate congestion	Quantificatio n and understandin g of uncertainty in arrival time predictions and passenger counts, aiding in more robust decision makin g processes	Acknowl edged and quantifie d uncertain ties, leading to more informed decision making processe s within transit manage ment
12	Potential research on data integration methods or real time data utilization for predicting arrival times more accurately	Utilizing image processing techniques for real time estimation of crowd density in transit areas	Insights into global crowdedness trends in transit, offering comparative analyses and potential insights into managing congestio	Provided compreh ensive insights into global transit congesti on trends, highlight ing areas for potential improve ment in various regions.
13	Contribution could involve a comprehensiv e review or analysis of existing prediction models and methodologies	Implementi ng image processing for detecting crowd motion, motionless individuals, and estimating crowd densities specifically in subway environmen ts	Specific insights into the dynamics and nature of crowds within public transport systems, potentially leading to strategies for alleviating congestion	Unveiled crucial insights into crowd dynamic s within public transport systems, paving the way for strategic intervent ions to

				manage congestion effectively
14	Likely research on system-wide improvements in public transportation, potentially exploring policy interventions or infrastructure enhancements	Employing computer vision methods to estimate the count of individuals within high density crowd scenarios	Accurate estimation of the number of individuals within high density crowd scenarios, assisting in crowd size estimation and planning	Enabled real-time assessment and management of crowd densities at transit locations , fostering immediate response s to crowdedness
15	Might delve into innovative technologies or systems for dynamic routing and real-time updates to improve bus arrival time predictions	Implementing automated tools or algorithms for estimating crowd sizes, aiding in better understanding and management of transit congestion	Enhanced detection of crowd motion and stationary individuals within subway environments , contributing to efficient crowd control measures.	Improved understanding of crowd behavior within subway environments, facilitating better crowd control measures
16	Possible exploration of user-centric approaches or passenger behavior analysis in arrival time prediction models.	Involves computer vision techniques to estimate the count of individuals within highly crowded scenarios using image or video data	Accurate estimation of headcounts within highly dense crowd scenarios, providing valuable data for crowd size estimation and management	Provided accurate estimates of crowd sizes, offering valuable data for crowd size prediction and management
17	Likely contributions on contextual factors affecting arrival time prediction, such as weather, traffic patterns, or geographical considerations.	Focuses on developing automated tools or algorithms to estimate crowd sizes in various transit settings, potentially aiding in crowd management strategies	Development of automated tools or algorithms specifically aimed at estimating crowd sizes, potentially facilitating quicker and more efficient crowd management strategies	Presented automated tools or algorithms for efficient crowd size estimation, potentially revolutionizing crowd management strategies

III. PROPOSED WORK

A real-time application that helps identify passenger movement at a separate place is the suggested method. A prediction of the KSRTC department's revenue and resource planning are also included in this plan. Fig. 1 displays the architectural diagram of the passenger movement. The information is provided in CSV (comma-separated values) format. The data is uploaded into a domestic neural network pattern recognition application, which uses a greedy approach to establish a cluster after synthesizing the data in long short-term memory. The recurrent neural network then wraps the synthesized data to produce visual analysis and statistical modelling of passenger forecasting.

3.1 METHODOLOGY

Deep learning algorithms are classified as unsupervised algorithms since they use neural networks to produce their remarkable results. Deep learning is applied to voice restructuring, image rearrangement, and prediction. Consequently, our proposed method leverages three distinct algorithms for deep learning: greedy layer-wise algorithm, recurrent neural network, and long short-term memory.

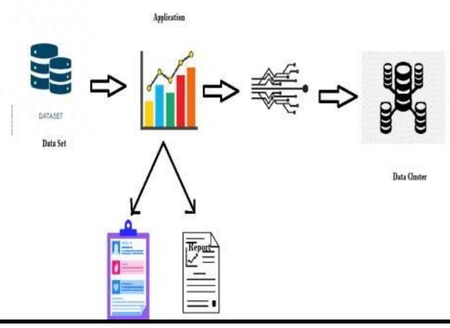

Fig: Architecture of Passenger flow

3.2 GREEDY LAYER ALGORITHM

The greedy layer-wise algorithm constantly selects the stages that result in immediate gains and operates bit by bit. In contrast, the greedy approach's decision makes no consideration for information or decisions that may arise in the future. The greatest option for an instant answer is to use greedy algorithms. outcome. The method is called "layer-wise" because the model is trained one layer at a time. Due to the layer-wise method used to handle the more challenging problem in the deep network, this can be described as "greedy." In this case, the algorithm gathers all datasets and extracts data according to regions. Following the formation of a cluster from all region-specific data using numerical data as the primary parameters, the data is then sent to an LSTM prediction.

Greedy layer algorithm

Step 1: Get all data

Step 2: Identification the region using the algorithm code

Step 3: Crop the data from real data

Step 4: Delete the column which not contain a numerical value

Step 5: send the data to prediction

3.3 LONG SHORT-TERM MEMORY (LSTM)

A new intermediate network design called long short-term memory adapts its learning strategy to the data by using a gradient-based approach. With LSTM, backflow issues are resolved. Additionally, blunder discharge was the reason behind the creation of LSTM.

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i)$$

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o)$$

$i_t \rightarrow \textit{represents input gate.}$

$f_t \rightarrow \textit{represents forget gate.}$

$o_t \rightarrow \textit{represents output gate.}$

$\sigma \rightarrow \textit{represents sigmoid function.}$

$w_x \rightarrow \textit{weight for the respective gate}(x) \textit{ neurons.}$

$h_{t-1} \rightarrow \textit{output of the previous lstm block}(\textit{at timestamp } t-1).$

$x_t \rightarrow \textit{input at current timestamp.}$

$b_x \rightarrow \textit{biases for the respective gates}(x).$

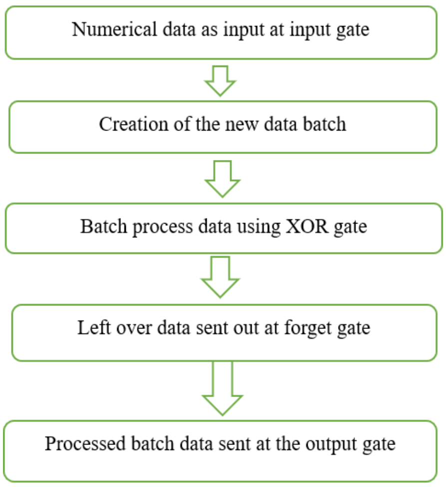

Fig 2: Processing Long short-term memory(LSTM)

Figure 2 illustrate the flow of LSTM in which processing of the long short-term memory algorithm for passenger flow prediction in bus transportation. In this model, redundant data and duplicate entries are stored in the long short-term memory after the primary parameters are taken into consideration in the greedy method. Data, and identical values for the index are eliminated. The resultant dataset is split into batches using the XOR gate. Following this process, the forget gate receives the remaining input, analyzes it, and forwards it to the neural network for prediction.

Long Short-term Memory Algorithm

Step 1: Fetch legacy bus transit dataset.

Step 2: Extraction day wise and region wise data

Step 3: To specify the output

3.4 RNN ALGORITHM

Among the several types of neural networks is the recurrent neural network. This involves using the earlier step's output as the current step's input. The main benefit of a recurrent neural network is that it has a hidden stage that is utilized to store sentence-related data. RNNs feature a memory, which is used to store information about the estimates that have been created thus far.

As was covered in the first section, both the prior and present inputs have an impact on the output. Let I1 be the bus transit's initial input, with a dimension of n × 1, where n is the column's length. S0 is the initial RNN cell with four neurons in its hidden state. For every cell, the input hidden state ought to be the previous one. Since no prior state is visible, initialize S0 in the first cell with zeros or a random integer. U is the dimension d × n dimensional matrix, where n is the number of data input columns and d is the number of neurons in the first RNN cell. W is an additional matrix within the KSRTC data frame, measuring d × d in dimension.

Mathematically outputs from the first RNN cell are as below

$$S1=AI1+WS0+v \qquad (2)$$

$$O1=BS1+c \qquad (3)$$

Parameters in the RNN are A,B,v,c we are shared among all the RNN cells, Parameters are learnable and are responsible for the training the model. At each step the loss is computed & is backpropagated through the gradient descent algorithm. Gradient represents the slope of tangent and points in the direction of the greatest rate of increase of function. From the loss, it means cost function or error. The move is made opposite to the direction of the gradient of the loss concerning V. Mathematically new value of V can be calculated according to the following equation.

$$Vnew= Vold=n\ d(L)/d(V) \qquad (4)$$

Where d (L)/d (V) is the sum of all losses obtained from time steps

Figure 3 shows the processing of recurrent neural networks for passenger prediction in a bus transportation system. Here the passenger datasets collect from the KSRTC department. Hereafter removing the similar index in LSTM then data will move to recurrent neural network here it consider as iteration based factors such as region-wise prediction, passenger count, revenue generated, year-wise prediction. After collecting the dataset, extract the dataset based on features like a holiday, weekdays, passenger travel from different areas, processing the dataset for prediction of passenger flow in the transportation system.

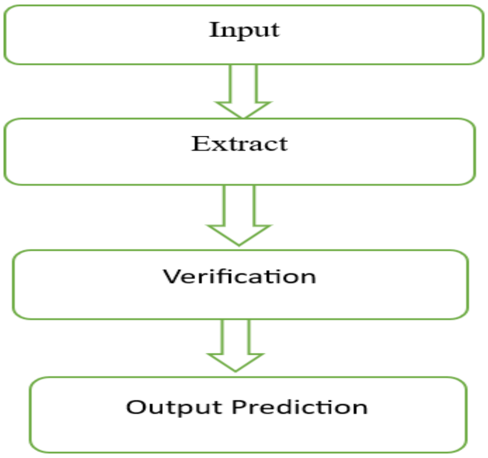

Fig 3: Processing of the recurrent neural network algorithm

Recurrent Neural Network Algorithm
Step 1: Selection of sample random numbers
Step 2: Create a decision tree by defining each sample's condition.
Step 3: Calculate cluster centroids
Step 4: Re-assign each point to the closest cluster centroid
Step 5: select the most voted prediction result as the final prediction result

V. CONCLUSION

For the KSTRC to operate and manage passenger transit, short-term forecasting of passenger flow is crucial. While traveling from one location to another, people encounter a lot of problems. Issues arise because of the high volume of travellers at the bus's frequency, delay, and stop. In order to solve these issues, we take into consideration the idea of passenger projection based on several criteria such as bus id, number of adults on board, number of children on board, source, destination, number of passengers, slot number, and income generated for the following year. The prediction in our suggested system is built on deep learning using LSTM, RNN, and a greedy layer-wise approach. These three algorithms produce income for KSRTC buses from every site and are capable of accurately forecasting the passenger flow in KSRTC buses. This could be beneficial to the KSRTCBRT division. to examine the revenue for the upcoming year.

VI. REFERENCES

1. Nandini Nagaraj1 & Harinahalli Lokesh Gururaj1 & Beekanahalli Harish Swathi1 & Yu-Chen Hu2 Passenger flow prediction in bus transportation system using deep learning exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022.

2. Ahmad Ali†, Yanmin Zhu†,∗, Qiuxia Chen‡,∗, Jiadi Yu†, Haibin Cai †Shanghai Jiao Tong University; ‡Shenzhen Polytechnic; East China Normal University Leveraging Spatio-Temporal Patterns for Predicting Citywide Traffic Crowd Flows Using Deep Hybrid Neural Networks IEE Internation Conference 2019.

3. Gaozhong Tanga , Bo Lia,∗ , Hong-Ning Daib , Xi Zhengc SPRNN: A Spatial-Temporal Recurrent Neural Network for Crowd Flow Prediction July 3, 2023.

4. Mr. ANJESH H L A Study on Performance Evaluation of KSRTC - With Special Reference To bangalore UGC Care Journal March 2020.

5. Vishwesh S. Hiremath & Prof. R. N. Mangoli An Overview and Structure of Karnataka State Road Transport Corporation (KSRTC) IJCRT July 2021.

6. Jeffrey Hanft, Shrisan Iyer, Brian Levine, and Alla Reddy New York City Transit Transforming Bus Service Planning Using Integrated Electronic Data Sources at NYC Transit Journal of Public Transportation, Vol. 19, No. 2, 2016.

7. N. Shanthi,1 Sathishkumar V E , 2 K. Upendra Babu, 3 P. Karthikeyan, 4 Sukumar Rajendran, and Shaikh Muhammad Allayear Analysis on the Bus Arrival Time Prediction Model for Human-Centric Services Using Data Mining Techniques Hindawi Computational Intelligence and Neuroscience Published 26 September 2020.