International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 03, March-2021

End to End Route Guidance

Riya Vyas
Dept of Information Technology
V.C.E.T.
Vasai, India

Reeya Shah
Dept of Information Technology
V.C.E.T.
Vasai, India

Parshva Vora
Dept of Information Technology
V.C.E.T.
Vasai, India

Prof. Maryam Jawadwala
Asst. Prof Dept of Information Technology V.C.E.T.
Vasai, India

***Abstract*—It is usually difficult for travelers to travel from source to destination without any prior knowledge about the same, the fare prices, the amount of time required to reach there and other travelling factors. We have proposed a project to tackle these issues, all at once. Whenever a person wants to travel from a specific source to the destination, the person can input the data about source and destination so the web page will suggest them the possible shortest route based on traffic analysis. It will also suggest the traveler the most convenient transportation medium and the fare rates for the same. The travel distance factor is represented by the distance based fare strategy, which is an existing differential strategy. We also aim to provide the rainfall possibility because weather is also an important factor while traveling which sometimes aims to threaten the safety and time measures. So a reliable solution for rainfall prediction is also provided. This problem is formulated as a bi-level program, of which the upper level maximizes the social welfare and the lower level capturing traveler choice behavior is a variable-demand stochastic user equilibrium assignment model.**

Index Terms*—Java script, Deep learning, python CSS, HTML, dynamic traffic flow, optimum route identification.

I. INTRODUCTION

Optimum route identification is a very common part of the network analysis or best routing identification to minimize the time as well as the fuel consumption. The optimum route means the shortest route between two junctions i.e. the source and the destination with minimum time consumption. It also depends upon the traffic condition, width of the road, road type and the number of junction (Optimal Route Analysis). The optimization can be possible on the basis of the length and the travel time criteria. These problems are generally difficult to be encountered by a single person who is travelling. Each of the factor keeps on affecting the other. For e.g. the misleading of the route can affect the fare price and time estimation.

Identify applicable funding agency here. If none, delete this.

This project aims to provide a solution to the problems like fetching the route, time estimation, distance, and fare and weather conditions prior to the travelling.

II. PROBLEM STATEMENT

For a traveler to reach from source to destination, there arises several questions such as, how to reach? Where is this place exactly? By which route can I reach quickly? Which is the least traffic congested route? What are the available transportation media? These are the difficulties that a traveler who is new to the area. In case a travel is misguided and he takes a wrong rout, it will lead to his waste of time, fuel and extra fare prices. Many a times, there are cases when a cheaper mode of travelling is available yet the person is unaware of it and goes with the costlier one. This also leads to unnecessary waste of money. The purpose of this paper is to show real time route, which can be used to develop routing strategies that tend to improve both cost and service productivity measures. More specifically, motivated by situations where time-sensitive delivery is required, we examine the value of a real-time traffic information technology such as Google API key for arriving at an efficient vehicle routing method. We have considered the most reliable public transportation like buses, auto and cab to conclude the fare prices and compare all of them. The user can also check the fare prices in case he wants to use his own mode of transport by inputting the current value of diesel. Our approach uses a simple and computationally rapid method that is fully automated and reliable. We've used Deep learning (LSTM) to predict the rainfall possibilities based on the historical data by training a model.

III. RELATED WORK

Many researches have been done on different methods for finding optimized route and fare calculation.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 03, March-2021

A. A Real-Time Dynamic Route Control Approach on Google Maps using Integer Programming Methods

Integer programming approaches can be applied to TSP problems when small number of visiting nodes exist. In the integer programming filed, the most common solution methods for this problem are Exhaustive Search, Greedy Algorithm, Dynamic Programming and enhanced with Heuristic search in order to calculate the least cost of a route. TSP problem normally is accepted in the non-deterministic-polynomial (NP-hard) problem category since there is no a suited linear solution. Namely, if a problem cannot be solved

In a polynomial-time, it should be considered in NP-hard category. When the number of visiting nodes increase in the route, the complexity of the computational solution increases as well. Similarly, TSP problem can be solved in totally O (N!) time complexity by using recursion programming methods. As it is known, the big-O notation indicates the time complexity of the related algorithm. Recursion methods always give the best result of all possibilities when com- pared with optimization methods. When the number visiting points increase, calculation time increases exponentially as well. That's why optimization methods are preferred then linear approaches. According to the explanations of the methods, they have some different time complexities. In below table, there are time complexities in big-O notations and whether optimal solution is guaranteed by the method

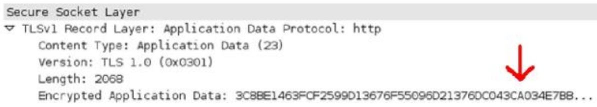
Or not. Obviously Bitmask Dynamic Programming style outperforms than the others in time complexities.

TSP Algorithms	Time Complexity	Data structure	Optimal Solution
Exhaustive Search	Exactly $O(N!)$	Not needed	Guaranteed
Greedy Search	$O(N)$	One dimensional Array	Not Guaranteed
A-Star (A* Heuristic) Search	Strongly better than $O(N!)$	Min Heap	Guaranteed
Branch & Bound Method	Better than $O(N!)$	Not needed	Guaranteed
BitMask Dynamic Programming	$O(N^2 * 2^N)$	Two dimensional Array	Guaranteed

Comparison of TSP algorithm

B. Traffic analysis with G-MAP trafficker

This proof-of-concept tool was developed in response to explaining SSL traffic analysis to several IO Active Customers. Although many papers have been published in the last decade on this topic—written by reputable research outlets including Microsoft Research—there hasn't been much to show in the realm of usable, real-world tools and/or source code. To meet this need and to more clearly illustrate the issue to IO Active's customers, this proof of concept tool was developed. To perform SSL traffic analysis, you must be able to identify corresponding TCP sessions and merge packets (either from an offline PCAP file or live capture sessions) into a correct data stream—this means correct packet reassembly, and Gaps-Trafficker uses libidos. Next, you need to parse the SSL structures—for the Actual traffic analysis we're not interested in the algorithm handshake at the beginning or at SSL alert messages, we're interested only in the actual Application Data containers of the SSL TCP stream.



Secure socket layout

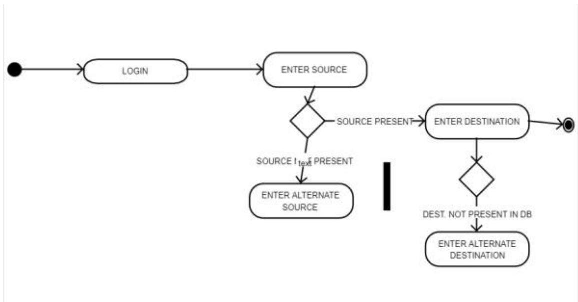
The best approach to fully understand the problem situation is to study a real-life traffic congestion scenario. This way we believe that we would be able to determine the most important varying parameters that could influence in the decision-making process. We studied the traffic system in the busy and well known Zarqa city in Jordan, where we collected data from its geographic information center (GIC) over a period of two weeks. Table 1 shows the data of average speed of a vehicle in each 15-minute interval Throughout a day based on distance and average congestion. Each intersection was examined and the distance between edges were calculated with the aid of a digitizing map and we used these information to represent the network as graph containing intersections (nodes) and roads (arcs).

	Loop number	0:00-0:15	0:15-0:30	0:30-0:45	0:45-1:00	.....	12--12:15
Day 1	1	73	72	72.9	73.82		74.2
	2	59.12	60.21	60.9	59.7		61
	..	..	..	..	..	..	..
Day 2	1	72.02	71.2	71,9	72.3		73.9
	2	60	61	62.2	61.9		61.82
	..	..	..	..	..	..	..

Average traffic flow - vehicle speed at loops in each 15-min interval. (Unit: mph)

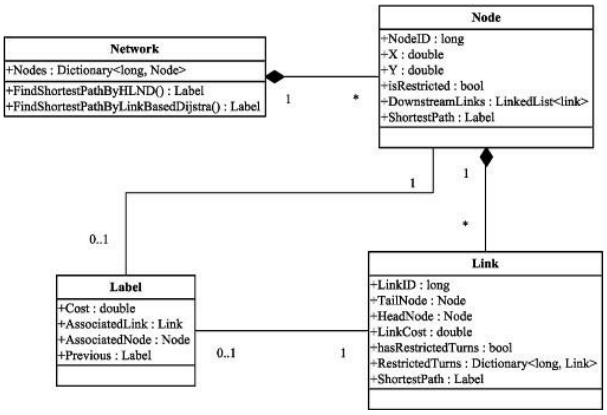
We then performed data analysis, which confirmed that the shortest path based on distance does not always guarantee fastest arrival and many times it was possible through other paths. This data collection and analysis helped us in arriving at the required parameters for the study. We observed that the total cost was directly proportional to distance and traffic congestion and inversely proportional to the vehicle speed

C. Optimum route analysis using road graphic plugin in QIS



Activity diagram

Mohammad Abousaeidi et. al has carried out GIS modelling approach for the determination of the fastest delivery route for the fresh vegetables. The work was carried out to Kuala Lumpur, Malaysia with the help of ArcGIS software.



Class diagram to find the optimal path

According to author, network analysis tool helps the decision-makers determine the best routes among all of the existing road networks for transportation and delivery services. M. Sureshkumar et.al argues that route optimization is one of the important requirement for proper traffic management in cities. Kanchipuram , Tamilnadu has been adopted for the study area for the determination of alternate routes for effective traffic management via Arc map 10.1 software.

J.R. Kinobe et.al has used GIS tool for Kampala city to optimize travel distances, trips and collection time, for the waste collection and transport system. Total number of trips and travel time has been decreased through GIS hence fuel consumption and vehicle emission also decreased. As per the author, the GIS based routing procedure is flexible and could be used in planning of waste collection policies and decision making mechanisms in waste management. – Study Area Surat city, Gujarat, India – Methodology includes the following steps: Step-1: Software and plugin installation Step 2: Creating shape file and adding base map Step 3 : Selection of locations for getting shortest route between them. For this step, the direction of travelling, number of trip, average speed between nodes, output result units can be changed as per our requirements

Trip	Type	Length	Average speed	Time required
Athwagate circle to Surat railway station	One way trip	4.2 km	20 km/h	≈12.5 minutes
	Two way trip	8.4 km	20 km/h	≈25 minutes
L.P.Savani circle to Udhna Darwaja	One way trip	6.4 km	20 km/h	≈19.2 minutes
	Two way trip	12.8 km	20 km/h	≈38.4 minutes

Table 1: Result table obtained via QGIS

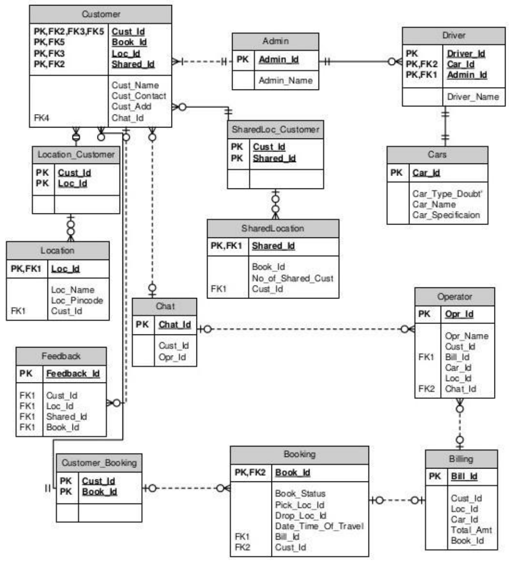
Trip	Type	Length	Average speed	Time required
Athwagate circle to Surat railway station	One way trip	4.9 km	18.37 km/h	≈16 minutes
	Two way trip	9.8 km	18.37 km/h	≈32 minutes
L.P.Savani circle to Udhna Darwaja	One way trip	6.6 km	30.55 km/h	≈13 minutes
	Two way trip	13.2 km	30.55 km/h	≈26 minutes

Table 2: Result table obtained via Google map

Result table

D. Fare estimation using deep networks

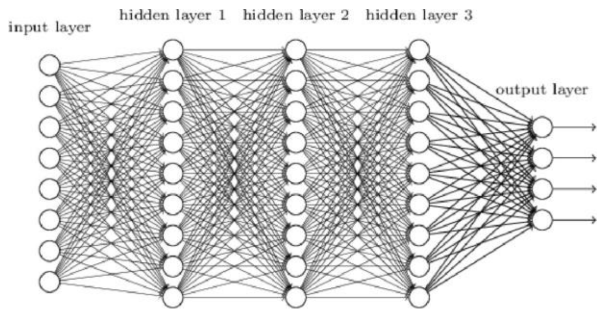
The aim of this competition is to predict the fare class using the features. The dataset consists of various codes or different formats, such as timestamp, longitude and latitude.



Class diagram for fare estimation

We also attempted to introduce derived features to increase the efficiency of our system. Derived features using the primary features, we derived selected features listed below:

- Time (morning, afternoon and evening) • Weekends • Holiday • Distance from airport • Distance from city centre
- Distance from one of the tourist spots (Church of Agios Dimitrios)



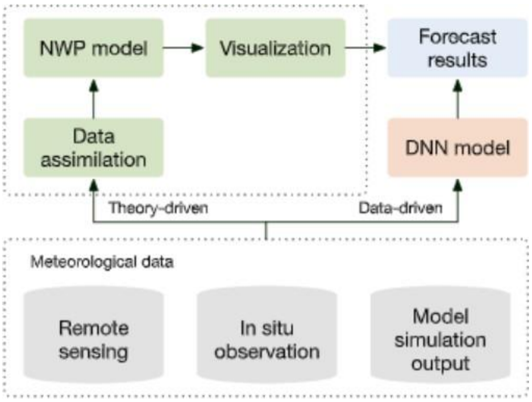
Dnn clustering

E. Deep learning based weather prediction

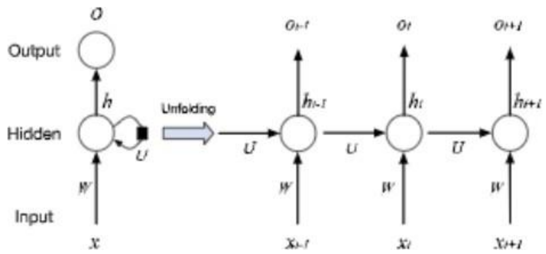
The aim is to use Deep learning techniques to predict the rainfall possibility of the next day at any particular hour, based on the weather data of the current day of this city and a couple of its surrounding cities.

Different types of data are used to achieve the goal. It includes multi-dimensional data, satellite data and longtime sequence data.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 03, March-2021



The result achieved from multi-dimensional data are in multiple factors. The temperature of a region at a particular time is represented as array. Satellite image data are used for extreme conditions. They are used to predict the weather conditions for extreme times. It gave a highly accurate view in to the result achieved. Convolutional neural network (CNN) which is a type of DNN is applied through images. Long time sequence data is used for the cases where the constraints of data is very high and overs hundreds of years-of data. In such cases there are different elements of the data that comes into existence. RNN are used in such cases where a large data set cannot be shown at once. So, the data is processed step by step using RNN. The input, Forget and output gates are used by the LSTM model to redirect the data and get the best result.



IV. COMBINED STUDY

Deep learning consists of assigning weights across multiple hidden layers. Outlined in their reviews of deep learning methods and architectures, types of deep learning models, and purposes. Several types of deep neural network architectures have been designed. For this task, we used Recurring neural networks (RNN), which consist of multiple learning layers. RNN have various parameters that control the output of the systems. Reasons to choose RNN method:

•RNN helps to remember a long amount of data for a period of time. It has huge data storage capacity.

• Learning rate: The neural network updates the coefficients for each iteration as it corrects for error. A large learning rate ensures that the network learns quickly but might result in over-fitting. A small learning rate slows the learning process but could lead to under-fitting.

• Activation function: This parameter refers to functions that determine the threshold at each node above which a signal is passed through the node and below which the signal is stopped.

• Update function: This parameter describes the manner in which a neural net minimizes error and loss or finds the least error as it adjusts its coefficients in a step-by-step manner in each mini-batch.

In addition to the mentioned classifier, stacking classifier systems have proven to be highly competitive for classification tasks, and thus we decided to include a stacking classifier in our submission. However, each classifier produces its own output, and therefore, we need a combination mechanism to combine the results. This task can be accomplished by voting on each classifier, weighting the voting (certain classifiers have more authority than the others), and averaging the results, etc. In stacking, the combining mechanism is given as follows. The outputs of the classifiers (Level 0 classifiers) are used as training data for another classifier (level 1 classifier) to approximate the same target function, as shown in the figure below. In general, the combining mechanism is used in the level 1 classifier. The first step in this type of architecture selects the set of base classifiers that form this classifier. For this purpose, we used the support vector machine and multi-layer perceptron. The evaluation metric, i.e., quadratic weighted Kappa, is described and used to assess the results of our system.

Cohen's Kappa coefficient [5] is widely used to quantify agreement between two raters on a nominal scale [6] by correcting the observed percentage of agreements between raters for the effect of chance. A value of 0 implies no agreement, whereas 1 corresponds to a perfect agreement between the two raters. Situations exist in which disagreements between raters or users might not all be equally important. To account for these inequalities, Cohen [7] introduced weights in the formulation of the agreement index, leading to the weighted Kappa coefficient. Consider two raters who classify a sample of n subjects into Z categories of an ordinal scale, where nij is the number of items classified into category i by rater 1 and category j by rater 2, ni. is the number of subjects classified into category i by rater 1, and n.j is the number of subjects classified into category j by rater 2, as given in Table 4 below

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 03, March-2021

	Rater 2			
Rater 1	**1**	**j**	**Z**	**Total**
1	n_{11}	n_{1j}	n_{1Z}	$n_{1.}$
i	n_{i1}	n_{ij}	n_{iZ}	$n_{i.}$
Z	n_{Z1}	n_{Zj}	n_{ZZ}	$n_{Z.}$
Total	$n_{.1}$	$n_{.j}$	$n_{.Z}$	n

The weighted kappa coefficient can be defined in terms of agreement weights by the following:

$$Z_w = \frac{p_o - p_e}{-p_e} \quad \textbf{(1)}$$

where $p_o = \sum\sum w_{ij} p_{ij}$ and $p_e = \sum\sum w_{ij} p_{i.} p_{.j}$ with $\leq w_{ij} \leq$, $w_{jj} =$, whereas Fleiss and Cohen [9] used the quadratic weights given below:

$$w_{ij} = {1 - (i-j)^2}/{(Z-1)^2} \quad \textbf{(2)}$$

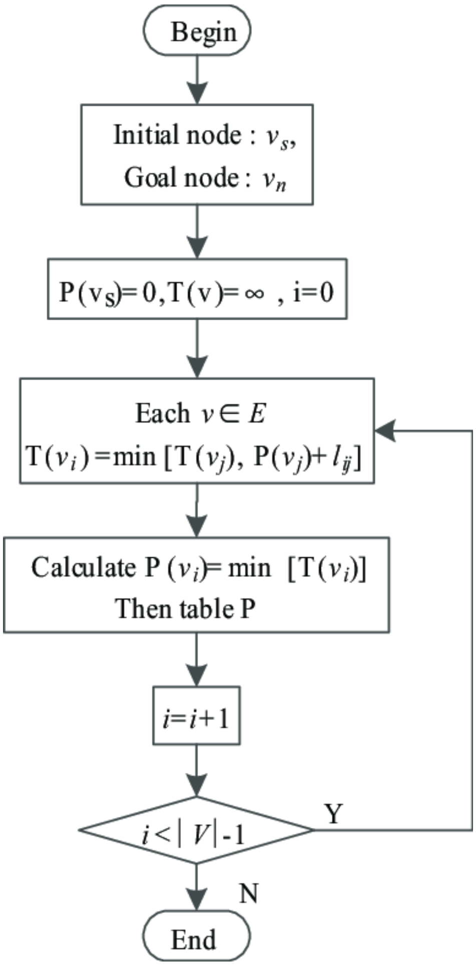
Kappa coefficient

Because RNN deep learning requires a large amount of data for proper learning and classification, the competition test dataset is particularly good for this purpose. Therefore, in this paper, we compare our models using two larger datasets: a validation dataset composed of 434987 tuples and a test dataset composed of 467767 tuples. The validation set is randomly chosen using approximately 30for the dataset, whereas the test dataset is used to compare our score with those of various other datasets.

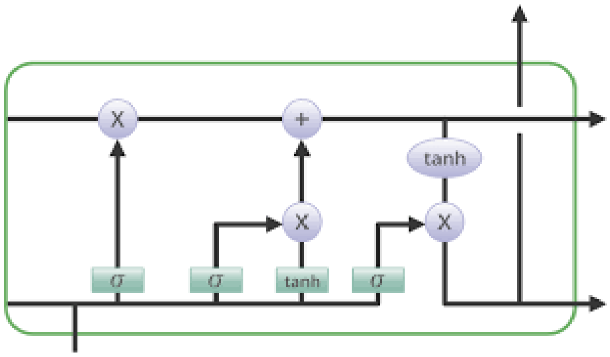
System	Validation	Private Leaderboard	Public Leaderboard
DNN	0.312	0.23376	0.23416
Stacking Classifier	0.288	0.2264	0.23262

Dataset

V. Project Implementation



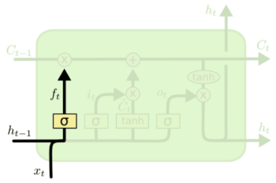
The algorithm to find shortest route



The algorithm for weather forecast

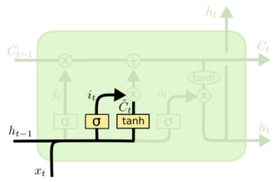
The very first step of out LSTM model is to decide what information it is fed with. Once the data is fed, the descisions are predicted by sigmoid layer. It is called as forget gate. It comprises of either 0 or 1. 0 stands for get rid of this and 1 stands for completely keep this.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 03, March-2021



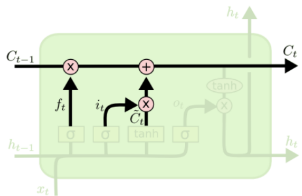
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

Next, we will decide what informaion we want to feed. This will be segregated into two steps. The first one is input gate which takes the input and the next step is tanh state which updates the value.



$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

Next, we need to decide which information is more relevant enough for our next step and which one we don't require anymore. The information which we don't require anymore will be drop out by our model.

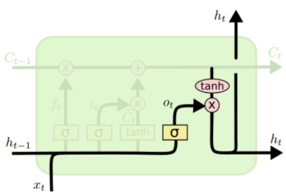


$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

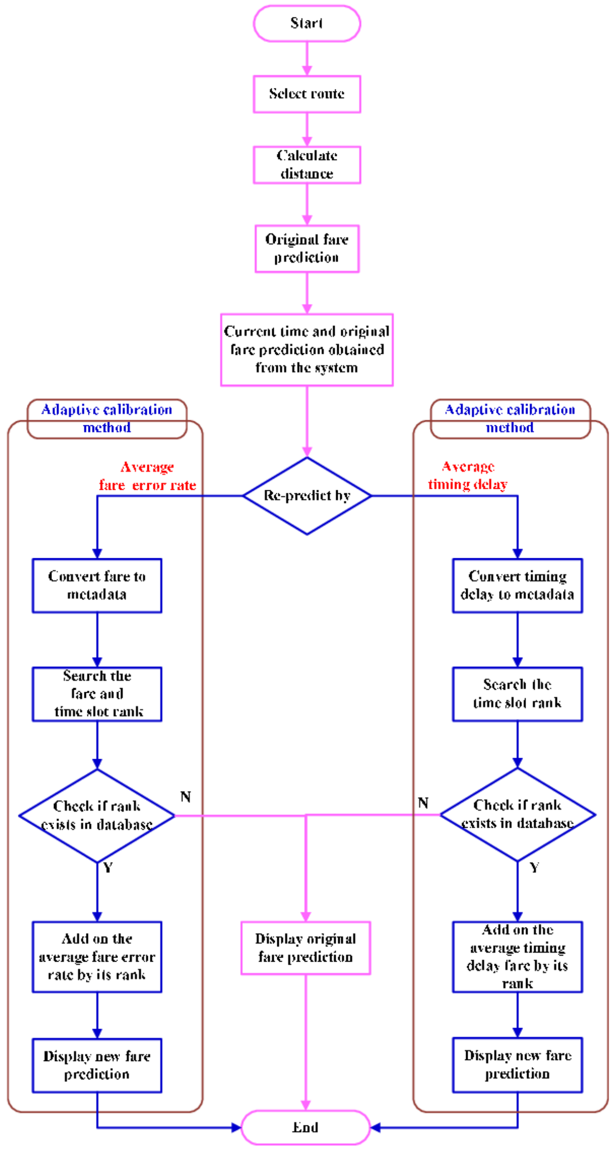
Now, we need to update the old cell. Our new cell already knows what to do. The information about the old state will be dropped completely now and we will only focus on new states now. We will multiply the current values will ft. In the next part we decide about the output. to know that, we run our sigmoid layer. It helps us to filter out out desired output. Then we put the cell through sigmoid and multiply it by the output of the sigmoid gate. this will help us to get our desired output.

For this prediction model we have taken a data set which helps us consider the data of last 90 days to predict the rainfall of next five days. We have used sequential function to train the model layer one by one.

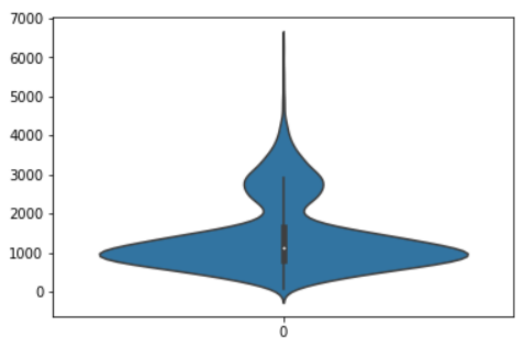
We split the data in 90:10 ratio for training and testing data set. this ratio is selected specifically to obtain the best accurate result as per the size of the data set.



$$o_t = \sigma\left(W_o \; [h_{t-1}, x_t] \; + \; b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

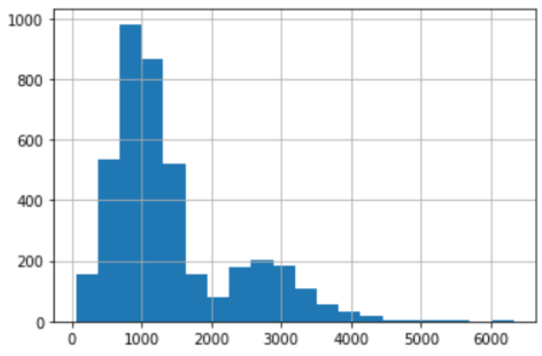


Algorithm for fare estimation



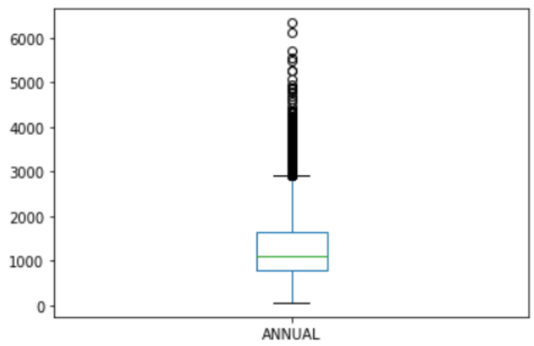
Violin graph

The annual rainfall is display using violin graph methodology. The most occurred frequency of rainfall lies around 1000 mm.



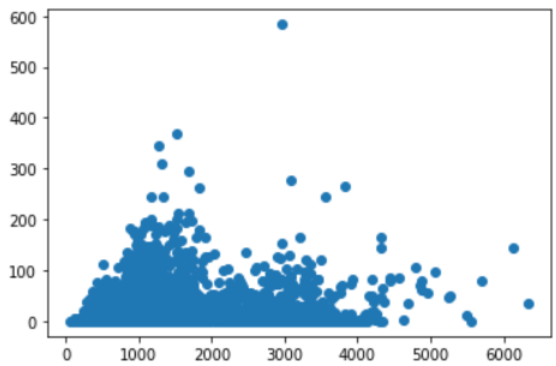
Histogram

Histogram is used to display the annual rainfall. Maximum frequency is observed between 500-2000 mm.



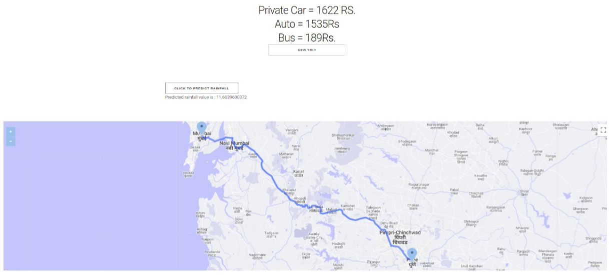
Box Plot

We have used our scale as Annual. This indicated most of the times, the rainfall lies between 800 mm to 2000 mm. In which our median holds at 1000 mm. Outliers are beyond 3000 because of heavy tailed distributions.



Scattered plot

The maximum frequency in the graph plotted for annual vs monthly is observed at 1000 mm to 2000 mm.



VI. Conclusion

Based on the above approach, we can reliably identify a complete zoom and get a bunch of coordinates back for the rectangle. We can then convert these coordinates to latitude longitude pairs. Route optimization can be effectively done via Google API, Here the obtained optimum length for two different locations closely matched with the google map result. We also obtained estimated time to reach to the destination point along with the shortest route. The route optimization through API can be effectively used for the road planning and network designing to save the time and cost of transportation. We presented a technology to utilize deep learning techniques to provide rainfall prediction. Deep learning technology can provide intelligent models, which are much simpler than traditional physical models. They are less resource-hungry and can easily be run on almost any computer including mobile devices which further helps the travelers to travel with a prior notice of rainfall possibilities. Our evaluation results show that these machine learning models can predict weather features accurately enough to compete with traditional models. We also utilize the historical data from surrounding areas to predict weather of a particular area. We show that it is more effective than considering only the area for which weather forecasting is done.

VII. Future scope

We plan to further optimize this project by providing cab booking facilities directly from our website. In future, we have plans to utilize low-cost Internet of Things (IoT) devices, such as temperature and humidity sensors, in collecting weather data from different parts of a city. The use of different sensors could increase the number of local features in the training dataset. This data, along with the weather station data, will further improve the performance of our prediction.

References

[1] Ammar Alazab1, Sitalakshmi Venkatraman, Jemal Abawajy and Mamoun Alazab, An Optimal Transportation Routing Approach using GIS-based Dynamic Traffic Flows, School of Information Technology Deakin University, Australia, 2011

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 03, March-2021

[2] Pranav Shah, Prof. Payal Zaveri, M. Tech Student 2Assistant Professor, Optimum Route Identification by using Road Graph Plugin in QGIS, Department of Environmental Engineering 2Department of Civil Engineering. Emerging Research and Innovations in Civil Engineering, 2019

[3] Wang Shu-Xi, The Improved Dijkstra's Shortest Path Algorithm andIts Application, School of Information Technology, The University ofInternational Business and Economics, Bei Jing, China, 2012

[4] 2018. Weather forecast using the sensor data from your IoT hub in Azure Machine Learning. https://docs.microsoft.com/en-us/azure/iot-hub/ iot-hub-weather-forecast-machine-learning.

[5] Y Radhika and M Shashi. 2009. Atmospheric temperature prediction using support vector machines. International journal of computer theory and engineering 1, 1

[6] Mark Holmstrom, Dylan Liu, and Christopher Vo. 2016. Machine Learning Applied to Weather F

[7] Rohit Kumar Yadav and Ravi Khatri. 2016. A Weather Forecasting Model using the Data Mining Technique. International Journal of Computer Applications 139,

[8] Kairanbay Magzhan, Hajar Mat Jani, A Review And Evaluations OfShortest Path Algorithms, International journal of scientific technologyresearch volume 2, issue 6, 2013

[9] Jonathan A. Weyn, Dale R. Durran, Rich Caruana "Improving Data-driven global weather prediction using deep convolutional neural networks on a cubed sphere" University of Washington, March 2020.