

PREDICT THE TRAVEL TIME USING TIME SERIES MODELING AND CLASSIFIER

DR.S. Krsihnaveni¹,
Assistant professor,
Department of B.Com(Business Analytics),
PSGR Krishnammal College for women,Coimbatore,India.
krishnaveni@psgrkew.ac.in

R.V.Divya²,
UG Scholar,
Department of B.Com(Business Analytics),
PSGR Krishnammal College for women,Coimbatore,India.
rvdivya21@gmail.com

ABSTRACT

Uber Technologies Inc. is an American international transportation network company headquarter in San Francisco, California and founded as “UberCab” by Travis Kalanick and Garrett Camp in March 2009. It develops, markets and operates the uber mobile app, which allows consumers with smartphones to submit a trip request which is then routed to uber drivers who use their own cars. This paper explores the applications of machine learning for predicting the travel time in the ride sourcing networks using time series modelling in Uber movement dataset. To this end, in the first step, descriptive analytics is completed to include potential features(attributes) affecting the travel times of Uber (ride sourcing) services. Then forecasting techniques such as time series modelling are applied to identify key attributes(features) for the prediction of the average travel times. To predict the accuracy of the average travel times using Time series modelling and Decision tree classifier.

Keywords: Uber movement, Travel time prediction, Time series modelling, Decision tree classifier.

I.INTRODUCTION

Data-driven mobility modeling and prediction are important aspects of Ride-sourcing industry faces a booming development in recent years, occupying quite share of vehicle usage on daily travelling in big cities. Uber owned 54% of the United States ride-sourcing cities. Uber 2017, Uber launched “Uber Movement”, a website that employ Uber’s riding data to help urban planners in improvising urban and traffic decisions. The Uber movement website provides zone-to-zone travel time data (the arithmetic and geometric mean and standard deviations) of census Tracts and Traffic Analysis Zones to make decisions. With respect to the Uber movement data, this paper was motivated to predict the accuracy of average time taken of Uber services and identify major factors(attributes) affecting that by using time series modelling. In this time series modelling, the required steps are: data mining, data preparation, data preparation, data and predictive analytics. Decision tree classifier is one of the most popular data mining techniques. So here the objective was to show the comparative performance between Times series modelling and Decision tree classifier. Using the Python programming environment, a paper is presented to analyse the travel time of the ride-sourcing services from the Coimbatore to the given specific destinations by considering the distance and street density in different destination zones areas to predict the average time taken for each passenger using Time series modeling and Decision tree classifier in uber movement dataset.

II.OBJECTIVE

This objective is framed in the goal as to achieve and increase in the trips being booked in the day to day period. The datasets collected from Kaggle and modified according to the objective. The dataset contains attribute like Date, Driver_Id, Request_Id, place, status, ratings, no of passengers. The main objective is to increase the revenue of the organization. The machine learning algorithms are used to predict the accuracy of travel time according to distance and time.

III.RELATED WORKS

Travel time from one location to another has been widely used to measure transport accessibility. Conventional travel survey data usually cannot collect accurate travel times since it highly relies on participants' memories. Widely used in transport research, Geographic Information System (GIS) provides an approach to accurately measure travel time.[10]

Accurate time-series forecasting during high variance (e.g., holidays), is critical for anomaly detection, optimal resource allocation, budget planning and other related tasks. At Uber accurate prediction for completed trips during special events can lead to a more efficient driver allocation resulting in a decreased wait time for the rides.[4]

Accurate travel time prediction is an important problem since it enables planning of cost-effective vehicle route and departure times, with the aim of saving time and fuel while reducing pollution.[6]

To analyse the travel time of the ride-sourcing services from the central Washington D.C. to the given specific destinations by considering the distance, railway/subway and street density in different destination zones (areas) and also weather conditions.[9]

At Uber, probabilistic time series forecasting is used for robust prediction of number of trips during special events, driver incentive allocation, as well as real-time anomaly detection across millions of metrics. Classical time series models are often used in conjunction with a probabilistic formulation for uncertainty estimation.[2]

Estimating temporal patterns in travel times along road segments in urban settings is of central importance to traffic engineers and city planners. In this work, we propose a methodology to leverage coarse-grained and aggregated travel time data to estimate the street-level travel times of a given metropolitan area.[3]

Decision tree learning uses a decision tree as a predictive model which was observations about an item to conclusions about the item's target value. It is one of the predictive modeling approaches used in statistics, data mining and machine learning.[7]

Travel time estimation approach incorporates forecast of transportation nodes impedance and travel time on network links. Forecasting period is two hours and the estimation is based on historical data and real time data on traffic conditions.[1]

Travel time prediction (TTP) can be an important and useful tool for both travelers and traffic management. Accurate and reliable short-term travel time prediction can greatly help vehicle routing and congestion mitigation. One of the most challenging tasks in TTP is developing and selecting the most appropriate prediction algorithm using the available data.[5]

Providing transportation system operators and travelers with accurate travel time information allows them to make more informed decisions, yielding benefits for individual travelers and for the entire transportation system.[8]

IV.METHODOLOGY

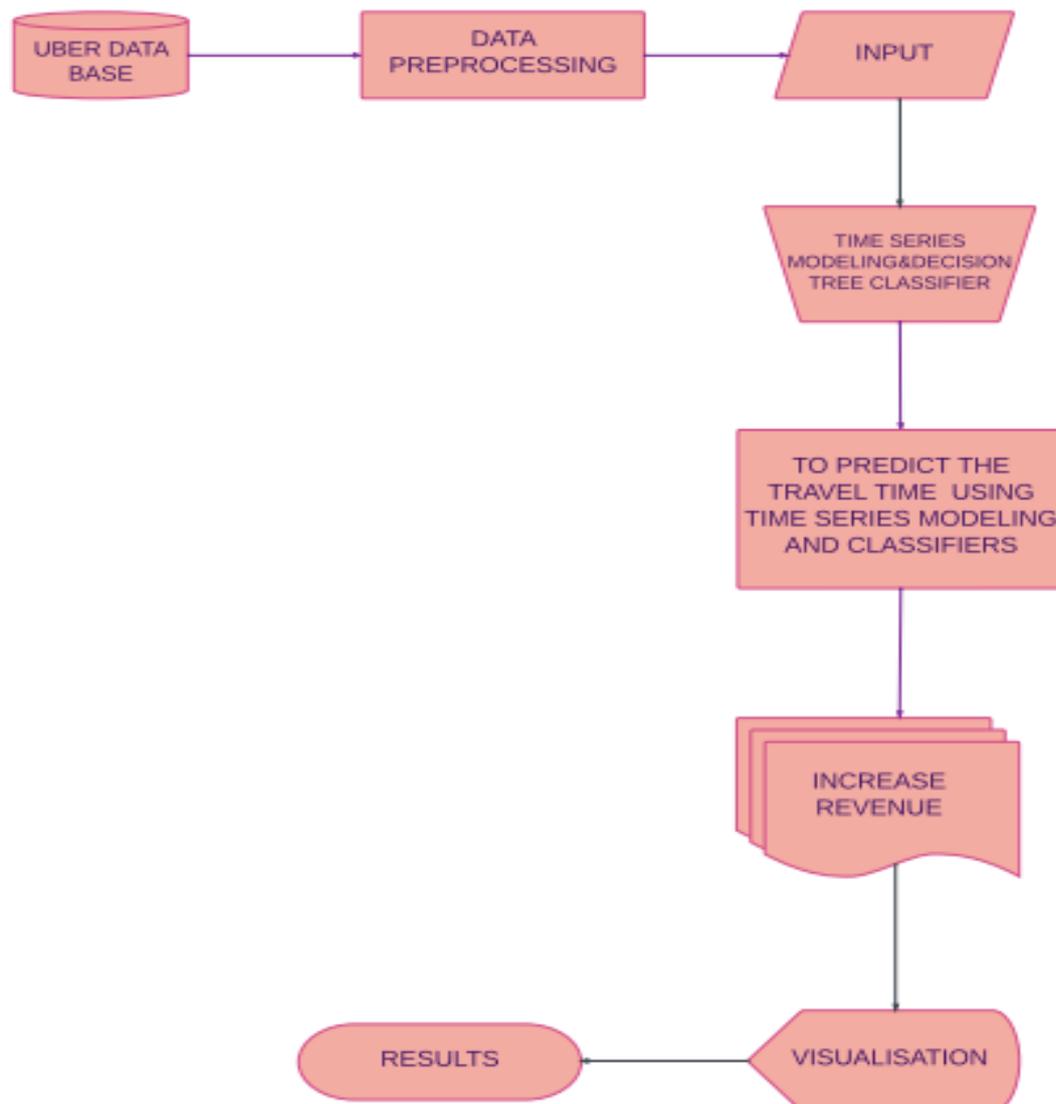
A. TIME SERIES MODELING

In descriptive statistics, a time series is defined as a set of random variables ordered with respect to time. Time series are studied both to interpret a phenomenon, identifying the components of a trend, cyclicity, seasonality and to predict its future values. Time series models are used to forecast future events based on previous events that have been observed through data collected at regular time intervals and it is a useful business forecasting techniques. The time variable is the independent variable and supports the target variable to predict the results. Using time series modeling to predict the travel time with the help of rolling mean and standard deviation.

B. DECISION TREE CLASSIFIER

Decision tree is a type of supervised learning algorithm this means that they use prelabelled data in order to train an algorithm that can be used for both regression and classification problems and it is an intuitive supervised machine learning algorithm that allows you to classify data with high degrees of accuracy. Using Decision tree classifier to predict the accuracy of travel time in cab.

WORK FLOW



IMPLEMENTATION

Time series modeling is used to predict the travel time accuracy of the uber cab being booked but it gives a negative accuracy. So, by using the decision tree classifier algorithm we are able to predict the accuracy of travel time and it gives a positive accuracy. By comparing these two methods Decision tree classifier gives the positive accuracy of time series. This helps to contribute the increasing revenue of organization.

V.RESULTS

TIME SERIES MODELING

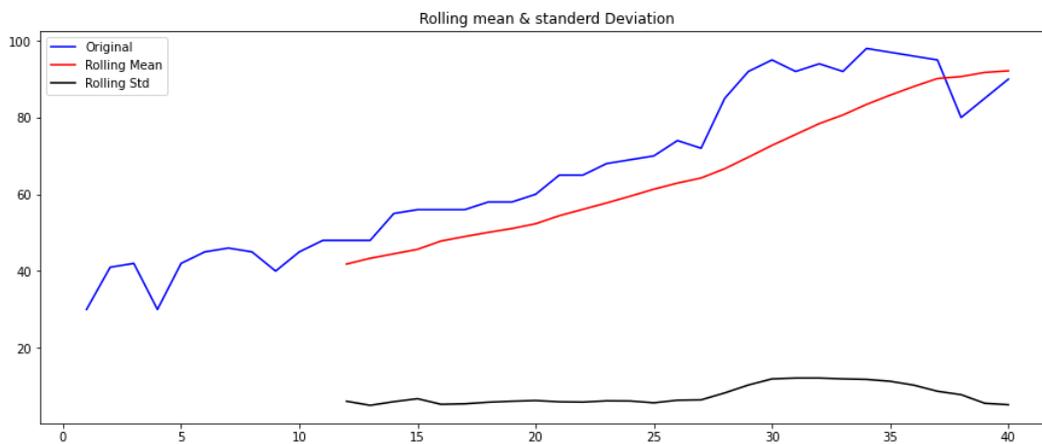


Fig 5.1 Visualization in Time series modeling

The Time series modeling predicts the travel time with the help of rolling mean and standard deviation.

DECISION TREE CLASSIFIER

```
# Measuring the accuracy of our model
from sklearn.metrics import accuracy_score
print(accuracy_score(y_test, y_pred))
```

0.968

Fig 5.2 Decision Tree Classifier prediction value

The Decision tree classifier predicts the accuracy of travel time using distance and time.

VI.CONCLUSION AND FURTHER WORK

In this paper, python google colab is used to predict the accuracy of travel time in uber cab. To find accurate travel time prediction is an important problem since it enables planning of cost-effective vehicle route and departure times, with the aim of saving time and fuel while reducing pollution. The further work can be proceeded with the different machine learning algorithms and also predict the travel time accuracy of different cities.

REFERENCES

1.Chien, S.I.-J.; Kuchipudi, C.M. Dynamic Travel Time Prediction with Real-Time and Historic Data. *J. Transp. Eng.* 2003, *129*, 608–616.

- 2.L. Zhu and N. Laptev, "Deep and Confident Prediction for Time Series at Uber," 2017 IEEE International Conference on Data Mining Workshops (ICDMW), 2017, pp. 103-110, Doi: 10.1109/ICDMW.2017.19.
- 3.Maass, Kelsey & Athanor, Arun & Rallo, Robert. (2020). Street-level Travel-time Estimation via Aggregated Uber Data.
- 4.N. Laptev, J. Kosinski, E. Li and S. Smyl, "Time-series extreme event forecasting with neural networks at uber", *International Conference on Machine Learning*, 2017.
- 5.Qiu, B.; Fan, W. Machine Learning Based Short-Term Travel Time Prediction: Numerical Results and Comparative Analyses. *Sustainability* 2021, 13, 7454.
- 6.Rupnik, J.; Davies, J.; Fortuna, B.; Duke, A.; Clarke, S.S.; Jan, R.; John, D.; Bla, F.; Alistair, D.; Stincic, C.S. Travel time prediction on highways. In *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology*, Liverpool, UK, 26–28 October 2015; pp. 1435–1442.
- 7.Sharma, Himani & Kumar, Sunil. (2016). A Survey on Decision Tree Algorithms of Classification in Data Mining. *International Journal of Science and Research (IJSR)* VL-5.
- 8.Shen, Luou. (2008). Freeway travel time estimation and prediction using dynamic neural networks. ProQuest ETD Collection for FIU.
- 9.Shokoohyar, Sine & Sobhani, Ahmad & Malhotra, Rashmi & Liang, Weimin. (2020). Travel Time Prediction in Ride-Sourcing Networks: A Case Study for Machine Learning Applications. 26. 1.
- 10.Sun, Yeran & Ren, Yanking & Sun, Xuan. (2020). Uber Movement Data: A Proxy for Average One-way Commuting Times by Car. *ISPRS International Journal of Geo-Information*. 9. 184. 10.3390/ijgi9030184.

