

PREDICT THE CANCELLATION OF TRIPS USING CLASSIFIERS AND TIME SERIES MODELING

Dr.S.Krishnaveni¹

Assistant Professor

Department of B.Com Business Analytics

PSGR KrishnammalCollege for Women, Coimbatore, India.

krishnavenis@psgrkcw.ac.in

M.Jaya Tharunigha²

UG Scholar,

Department of B.Com (Business Analytics),

PSGR Krishnammal College for Women, Coimbatore, India.

javatharunigha.m@gmail.com

ABSTRACT

Data analytics has helped companies optimize their growth and their performance for generation. Data analytics and visualization has helped us with several benefits for studying and analyzing the trips based on the patterns in data. Analysis in depth and cherry on top are the insights that draw from the patterns. This project is all about understanding one such data set of Uber from Coimbatore and is very component to understand the use of data analytics and visualization. It is generated with the help of python language using classifiers and prediction. In this projects we can learn through various machine learning algorithm. ML algorithm is the scientific study of algorithm and statistical model that computer use to perform a specific task. Using this we can predict the cancellation of trips.

KEYWORDS: Time series modeling, cancellation of trips, decision tree classifier.

I.INTRODUCTION

Uber is a transportation company with an app that allows passengers to praise a ride and drivers to change fares and get paid. Uber was started with an idea of luscious transport but in the following years it has been changed into something to offer a new service to choose an application while detailing pickup location. It shares their rides that hire independent contractors as drivers as they pay them a certain percentage of fair. Uber offers rides under demand pricing model for both drivers and passenger and they provide location sharing service from point to point destination for the customers. It aims to help the user to find nearby car according to both the passenger and the driver. This helps the driver to reach out to the point soon. Besides people can easily travel from one place to another through Uber service and their willingness to own a private car will decrease. It increases the users willingness to use a car for their convenience to take them from one point to another point. Uber charges a base fare, the distance and the time spent is calculated according to the driving regardless of whether the vehicle is moving or not. Prices vary from place to place but each one has a fixed minimum fair to be paid. It uses an algorithm that increases the price according to the demand faced by them. During peak hours from Friday to Saturday nights, after work rush hour and big events and festivals are some of the main reason where drivers cancel their trips due to lack of time and this is why customers suffer from trip cancellation. The dataset of Uber from Kaggle is used. By using time series modeling and decision tree classifier are used to predict the cancelling of trips and the solution is that if the number of cars are increased it will be able to meet the high demand and it will reduce the cancellations.

II.OBJECTIVE

This Objective is framed in the goal as to achieve to reduction of trips being cancelled. The datasets collected from kaggle and contains attributes modified according to the objective. The dataset contains attributes like DATE, DRIVER ID, STATUS, REASON, and NO OF PASSENGERS. The main motto is to reduce the cancellations of trips by using machine learning algorithm. By using this we can find out the availability of cars and suggest Uber to avail more number of cars.

III. RELATED WORKS

Uber is a leading Cab Booking service that offer services like cab booking and ride-sharing by using an app. To increase demand from customers and expand their outlook, these technology that involved in consistent competition which gain them their new users and retain existing ones. This paper gives you a complete overlook of considerable study of the idea of cab booking in India.[2] This objective is to understand the nature and the perception of Indian users who avail these cab services through mobile apps, especially users of Uber . In doing so, the motive is to empathise with Indian cab users, both drivers and traveler's, by understanding their problems and frustrations. This is achieved by using the ethnographic approach of Contextual Inquiry and work models are developed based on the insights found therein.[4]

The central idea of getting passengers from point to point with different customers can share the same ride and get it for half fair amount. Some times this concept does not work for all the kind of passengers. they implemented the door-to-door model and this caused trips to get unnecessary bend and even go back just to pick up another passenger, this causes a negative felling about the service provided. By 2016 a new service was provided to increase the revenue [1]

The impact to expand routes and weather conditions on pick-up waiting time, trip duration, and ride fare with a focus on the ride-sourcing mode in the cities. This shows that the trips organised from the city side have higher fares compared to the trips head toward the city center [3]

A potentially useful approach to this prediction task is the use of machine learning algorithms. Trip cancellation shows certain paratransit system operations. This has been a major drawback of overbooking cabs and this is the major drawback of cancellation of trips. [6] The judgment concerning customer's cancellation is based on the delay of cabs due to service of insufficient capacity. This shows that the good predictive model develops the aid in improving service. [5]

The concept of shared travel, making trips with other users via a common vehicle, is far from story. A change in technological climate has laid the way for cancellation of trips. [7]

IV. METHODOLOGY

A. TIME SERIES MODELLING

Time series modeling is used to predict future events based on preceding events that have been observed at regular time intervals. Time series analysis is a useful business forecasting technique. Time series is a series of dependent data points that are indexed in time order, usually taken at successive and equally spaced points in time. Time series models are widely used to predict the seasonal variability of a target variable over time, where past values are used as the input variables for the model. We are using time series modeling and decision tree classifier to predict the cancellation of trips.

B. DECISION TREE CLASSIFIER

The decision tree classifier is used to create a grouping model by constructing a decision tree. Each node specifies a test on an attribute, each branch has a descending from that node corresponds to one of the possible values for the attribute. It tree is a type of direct learning algorithm that can be used for both regression and classification problems. This algorithm helps us to predict cancellation of trips.

STEP 1: Imported the dataset from kaggle and modified the dataset and saved in excel .csv format.

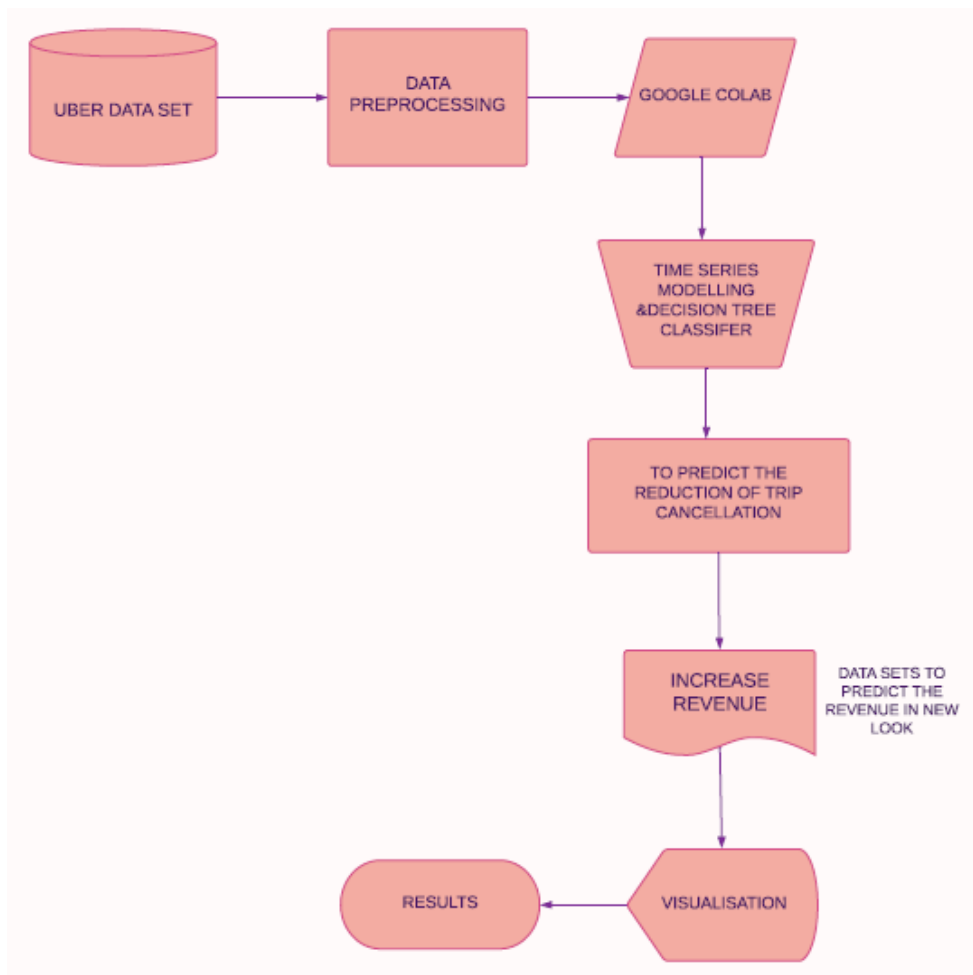
STEP 2: Used Google colab for executing python coding.

STEP 3: Preprocessing is done by removing all unwanted data from data set.

STEP 4: Time series modeling is used to predict cancellation of trips

STEP 5: Visualization are made in google colab for better understanding of dataset.

STEP 6: Decision tree classifier is used to predict the output positive and accuracy is calculated.

WORK FLOW**IMPLEMENTATION**

Time series is used to predict the cancellation of trips in Uber . By using time series modeling it has been predicted only the difference in plots and the mean and standard deviation .So, by using the decision tree classifier algorithm. To predict the cancellation of Uber trips and the accuracy comes out to be positive. By comparing these to algorithm we can predict the count of trip cancellation and increase the cab availability.

V.RESULTS

I.TIME SERIES MODELLING

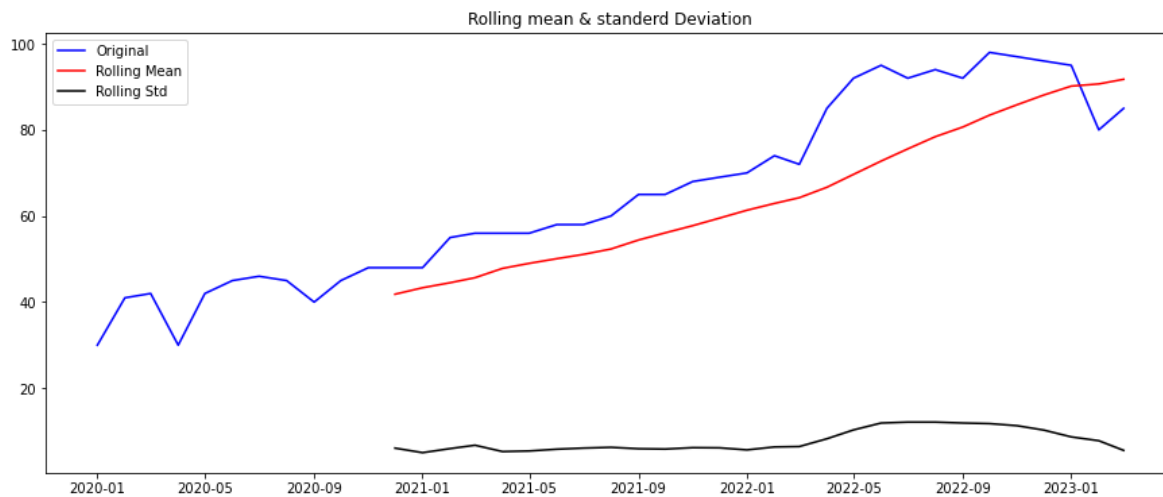


Fig 5.1 Visualization in Time series modeling

The fig5.1 helps to understand just the visualization that the Average passengers count lies between 40 to 100 for each month, but there is no accurate prediction. So, to predict this decision tree classifier is used.

II.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

In the above Fig 5.2 the accurate prediction value is being predicted. By comparing the two algorithm used here the decision tree classifier predict the correct value and provides a positive answer

VI.CONCLUSION AND FURTHER WORK

In this paper, importing data set, executing coding and visualization are done in Google colab. Time series modelling is used to predict the cancellation of trips. The cancellation of trips is high due to unavailability of cabs. To avoid the cancellation of trips the organization should increase the availability of cabs to reduce the cancellation. So that the revenue can be increased. The decision tree classifier is used to predict the accuracy.

REFERENCES

1. Acosta Pérez, Fernando & Ortiz, Gabriel & Muniz, Everson & Sacarello, Fernando & Kang, Jee & Rodriguez Roman, Daniel. (2020). Predicting Trip Cancellations and No-Shows in Paratransit Operations. *Transportation Research Record: Journal of the Transportation Research*
2. Avalos, Mariana & Perez, Daniel. (2021). Uber Express: A Case of Study. 10.13140/RG.2.2.13170.22721.
3. Carby-Hall, Joseph. (2021). The Uber Case. *Revue de droit comparé du travail et de la sécurité sociale*. 254-259.
4. Ch. Farronato, A. MacCormack, and S. Mehta. "Innovation at Uber: The Launch of Express POOL". In: *Harvard Business School* 619.003 (2020), pp. 1–21
5. Huang, H. C., A. Y. Chang, and C. C. Ho. Using Artificial Neural Networks to Establish a Customer-Cancellation Prediction Model. *Przeglad Elektrotechniczny*, Vol. 89, No. 1b, 2013, pp. 178–180.
6. Recanati, Francois. (2022). Understanding force cancellation. *Inquiry*. 1-19. 10.1080/0020174X.2022.2075920.
7. L.K. Poulsen, D. Dekkers, N. Wagenaar, W. Snijders, B. Lewinsky, R.R. Mukkamala, et al., "Green Cabs vs. Uber in New York City", 2016 IEEE International Congress on Big Data (BigData Congress), pp. 222-229, June 2016

