SAVITRIBAI PHULE PUNE UNIVERSITY

A PROJECT REPORT ON

TRAVEL TIME PREDICTOR

By

Nidhi Rajurkar	Roll No.:3346
Gaurav Sawant	Roll No.:3357
Shivaji Sawant	Roll No.:3358
Rushabh Gandhi	Roll No.:3350

Under The Guidance of

Prof. S.N. Reddy

Class: TE 3

Pune Institute of Computer Technology

Dhankawadi, Pune-411043

Pune Institute of Computer Technology

DEPARTMENT OF COMPUTER ENGINEERING

CERTIFICATE

This is to certify that the Project Entitled

TRAVEL TIME PREDICTOR

Submitted by

Nidhi Rajurkar	Roll No.:3346
Gaurav Sawant	Roll No.:3357
Shivaji Sawant	Roll No.:3358
Rushabh Gandhi	Roll No.:3350

is a bonafide work carried out by Students under the supervision of Prof. S.N. Reddy and it is submitted towards the partial fulfillment of the requirement of Bachelor of Engineering (Computer Engineering)

Signature of Internal Examiner

Signature of External Examiner

PROJECT APPROVAL SHEET

TRAVEL TIME PREDICTOR

Is successfully completed by

Nidhi Rajurkar	Roll No.:3346
Gaurav Sawant	Roll No.:3357
Shivaji Sawant	Roll No.:3358
Rushabh Gandhi	Roll No.:3350

At

DEPARTMENT OF COMPUTER ENGINEERING

PUNE INSTITUTE OF COMPUTER TECHNOLOGY

SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE

ACADEMIC YEAR 2017-2018

Prof. S.N. Reddy	Dr. Rajesh Ingle
Internal Guide	H.O.D
Dept. of Computer Engg.	Dept. of Computer Engg.

Abstract

It may be countries economy or world's economy, nothing falls in place if there is time delay. Right from day to day chores to international deals, time delay affects everything. In India, main reason for time delay in travelling is traffic.

Travel Time Predictor is an initiative to improve Travel time and reduce the time delay by predicting the travelling time for the coming next week. The predictor can give the time, a particular vehicle (public/private) will take at a particular time of day well in advance.

Travel Time Predictor is a Standalone Application which can be used to predict the time user will take to travel from location A to location B, well in advance (prior to week). It also allows user to analyze the day so that user can plan the schedule accordingly.

This may help greatly in planning our activities according to the travel time specified by our predictor. The best time to travel throughout the day can also be estimated for a particular area.

The main source of our information is Google Maps. And this is a small step from side in increasing the country's economy if the Travel Time Predictor is implemented on large scale.

Acknowledgments

It gives us great pleasure in presenting the preliminary project report on `TE Mini PROJECT TRAVEL TIME PREDICTOR'.

I would like to take this opportunity to thank my internal guide Prof. S.N. Reddy for giving me all the help and guidance I needed. I am really grateful to them for their kind support. Their valuable suggestions were very helpful.

I am also grateful to Prof. Dr. Rajesh Ingle, Head of Computer Engineering Department, Pune Institute of Computer Technology for his indispensable support, suggestions.

In the end our special thanks to Other Person Name for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for Our Project.

Nidhi Rajurkar

Gaurav Sawant

Shivaji Sawant

Rushabh Gandhi

(T.E. Computer Engg.)

Table of contents

TRAVEL TIME PREDICTOR

It may be countries economy or world's economy, nothing falls in place if there is time delay. Right from day to day chores to international deals, time delay affects everything. In India, main reason for time delay in travelling is traffic.

TTP is an initiative to improve Travel time and reduce the time delay by predicting the travelling time for the coming next week. The predictor can give the time, a particular vehicle (public/private) will take at a particular time of day well in advance.

TTP is a Standalone Application which can be used to predict the time user will take to travel from location A to location B, well in advance (prior to week). It also allows user to analyze the day so that user can plan the schedule accordingly.

This may help greatly in planning our activities according to the travel time specified by our predictor. The best time to travel throughout the day can also be estimated for a particular area.

Keywords

TTP: Travel Time Predictor

1. INTRODUCTION

- Purpose

  The travel time predictor predicts the time taken to travel from some place A to other place B a week in advance from the travel date and travel time. The main purpose behind the project is to give people the opportunity to plan tasks well before time. This saves time, enables efficient resource use and increases a human's economic value.

- Scope

  The scope of this project is to develop a web application for Travel time prediction using Python GUI. This will help user to know the duration of their travel in advance.

  The application is designed to help user plan things beforehand. More specifically, the application will increase the user's efficiency choose the date and time of travel. The web application will provide output for every time slot between 0800 to 2100 IST.
  The web application is free of cost for the user. The Internet connection is mandatory.

- Process

  The application allows user to input the starting location and destination by typing the addresses location. Along with the location, specify the travel date, time and day. The output will show user the time required for journey. User can find the travel time prior to a week and plan his schedule.

  Data over several weeks has been collected at different times and different days. This data is categorized in various classes depending on the time slots, days, weekday or holiday etc. The data processing is done using machine learning. And as a result the output travel duration is obtained from the learned data.

- Functional Requirement:

Sr.no.	Type	Requirement
1.	Services	• For each time slot and mode of transport separate data is analyzed. • And very specific output will be provided.
2.	Internet	• Fast Internet connection is not mandatory, but it would increase the performance of the application.

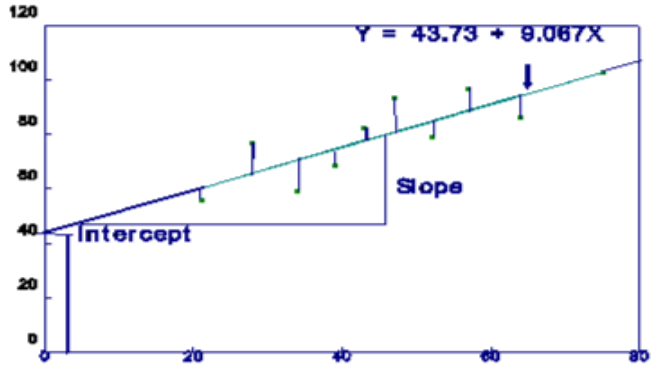
3.	Interface	• GUI using Python.
4.	Data and Integration	• Data is categorized depending on different slots and various other factors like working day or holiday.
5.	Performance	• The application is designed in a way that it will work under all circumstances.
6.	Operations	• Linear Regression for data processing using machine learning.

- Non- Functional requirements:

Sr. No.	Type	Requirement
1	Performance Requirement	• The data predicted by the application must be as accurate as possible.
2	Safety Requirement	• This application will not cause any harm to environment and humans.
3.	Security Requirement	• Security from external malwares, intrusion and worms. • Secured server. • Data interpretation error should not be caused due to external mediums.
4.	Software Quality Attributes	• The application will be available as many users possible and adaptable to the systems. • The application will be modifiable for new changes.
5.	Reliability	• The application will be reliable.

2. SURVEY OF MATHEMATICAL MODELS

- Linear Regression Model



Regression analysis is utilized to develop an accurate mathematical formulation of the regression analysis. The line of best fit is defined as a line for which minimizes the sum of squares of deviation of the various data points from the line. The regression line is also referred to as the least squares line.

In case of a multi-variable regression, the analysis is a sequence of multiple linear regression equations that are developed in a stepwise manner. At each step of the sequence, one variable is added to the regression equation.

The variable added is the one that makes the greatest reduction in the error sum of squares of the sample data. Equivalently it is the variable that when added, provides the greatest increase in the F value. Variables not having a significant correlation with the dependent variable, are those whose addition does not increase the F value and are not featured in the regression equation.

- Mathematical Computation of the Regression Coefficients

With One Independent Variable

The Mathematical Computation of the Regression Coefficients for the case of a single independent variable is given below:

The slope (regression coefficient) for the line of least squares is given by b, where

$$b = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

The intercept of the line is given by a, where a = y – bx.

The mathematical formula used for this computation is as follows:

$$b = \frac{\left(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} (x_i \sum_{i=1}^{n} (x_i y_i)\right)}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2}$$

THE RESIDUAL

The residual is defined as the difference between the actual and predicted values of the dependent variable. The standard error of the estimate is the standard deviation of the residuals. The standard error of the estimate can be calculated as follows:

$$s_x = \sqrt{\left(\frac{\left(\sum_{i=1}^{n} y_{i2} - a\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n}(x_i\, y_i)\right)}{n-2}\right)}$$

3. PROPOSED MATHEMATICAL MODEL

The proposed mathematical model uses linear regression algorithm for analyzing the data input and processing.

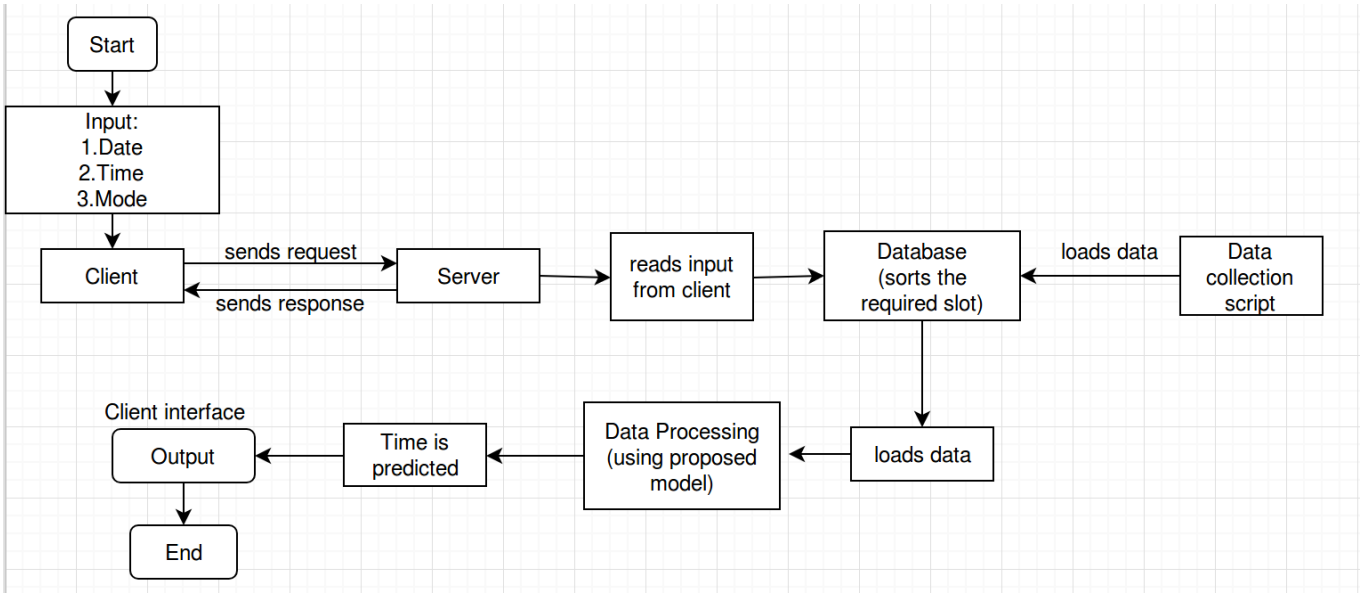
The implemented algorithm considers the time slot of the day against the duration of travel in minutes. The data collection is done at a frequency of 7mins from 0800 IST to 2100 IST. This is done for analyzing the data for cars and buses.

Based on the previous outputs and data collection, every time new data set is processed. The model approach does not take into account minute details like road conditions, weather conditions.

The residual or the standard error improves the learning of machine. The effectiveness of the project and the accuracy of the output increases subsequently after each processing. As the errors are recorded, the deviation of the output from the actual data is reduced.

When any request is sent for the data of any particular time slot, the y-intercept for the corresponding x-intercept is plotted. And the output is the predicted time taken for travel.

- **Class Diagram**



4. DESIGN AND ANALYSIS OF SYSTEM

- Assumptions, Constraints and Conditions

  Assumptions:

  All weekdays are working.
  Sundays are non-working.
  Festivals for a short period are neglected.
  The data provided by Google Maps is accurate and correct.

  Constraints:

  The data for all the routes throughout the city could not be considered due to lack of personalized database.
  Lack of source to cross check the data provided by Gmaps.

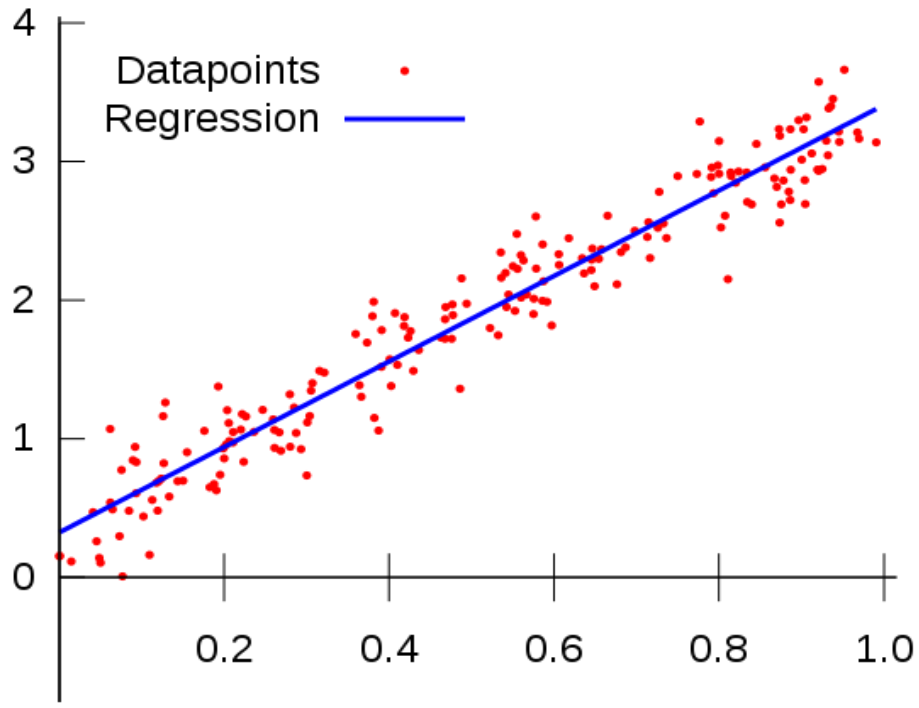
- Requirements

  Google maps API
  Operating System: Linux open source Ubuntu 16 OS

- Other Requirements

  As Google Map API is being used for the application, it is mandatory that we abide by the terms of use able by Google. Fast Internet connection not mandatory but it will increase application performance. Consistent Internet connection is required. The application uses GMaps API data to plot the graph.

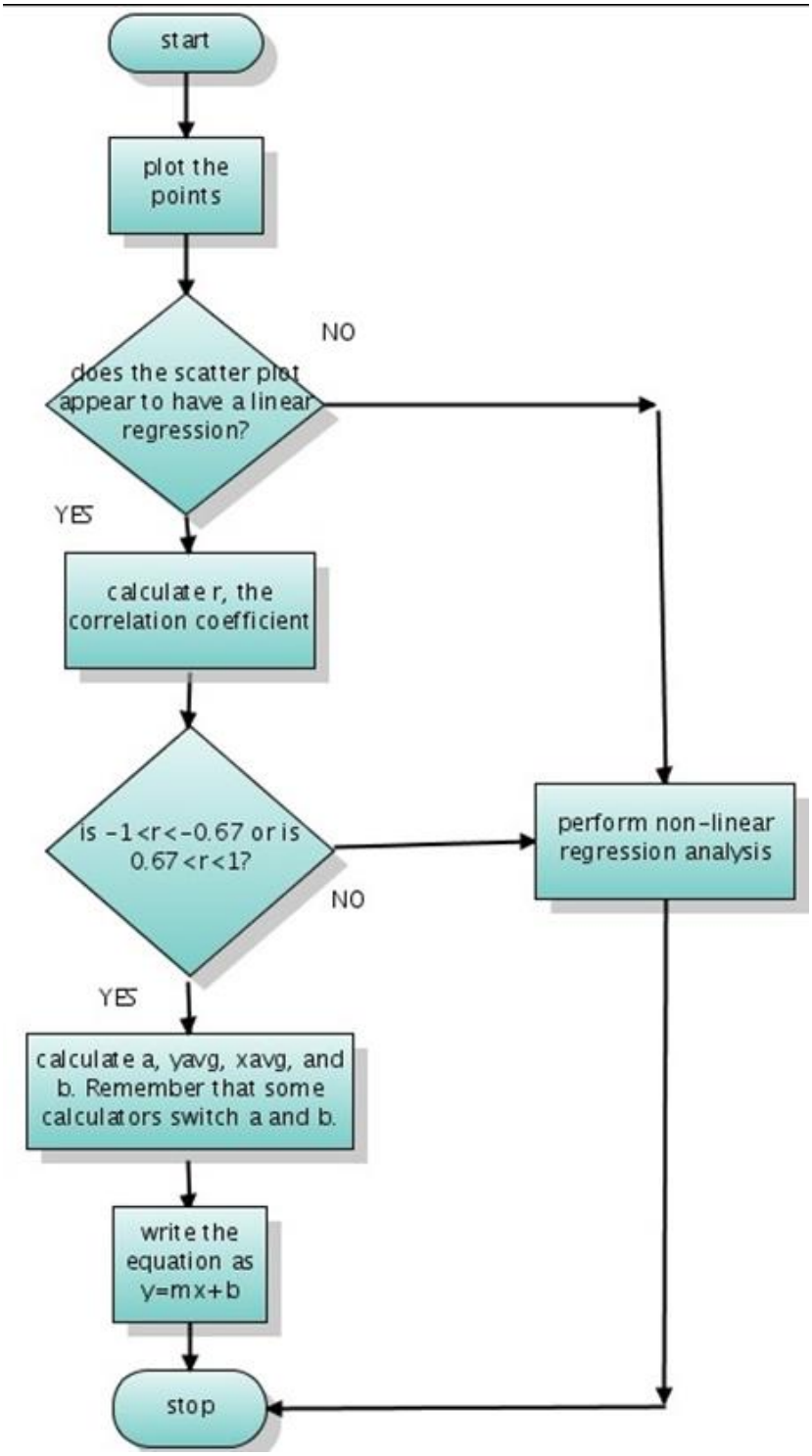
- Machine Learning: Linear Regression:



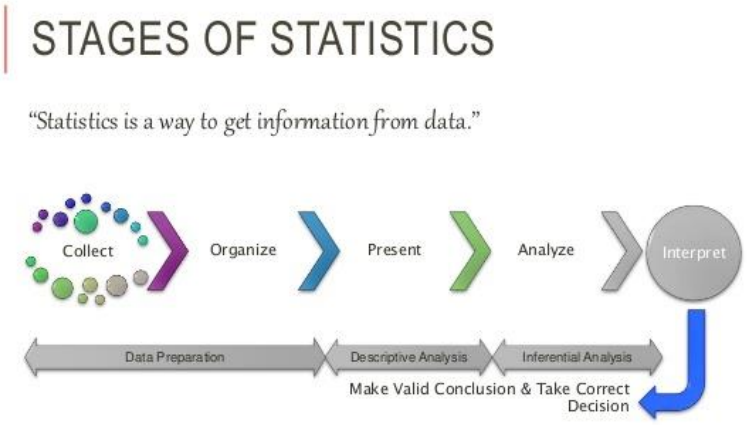
Form of Linear Regression

- System Architecture

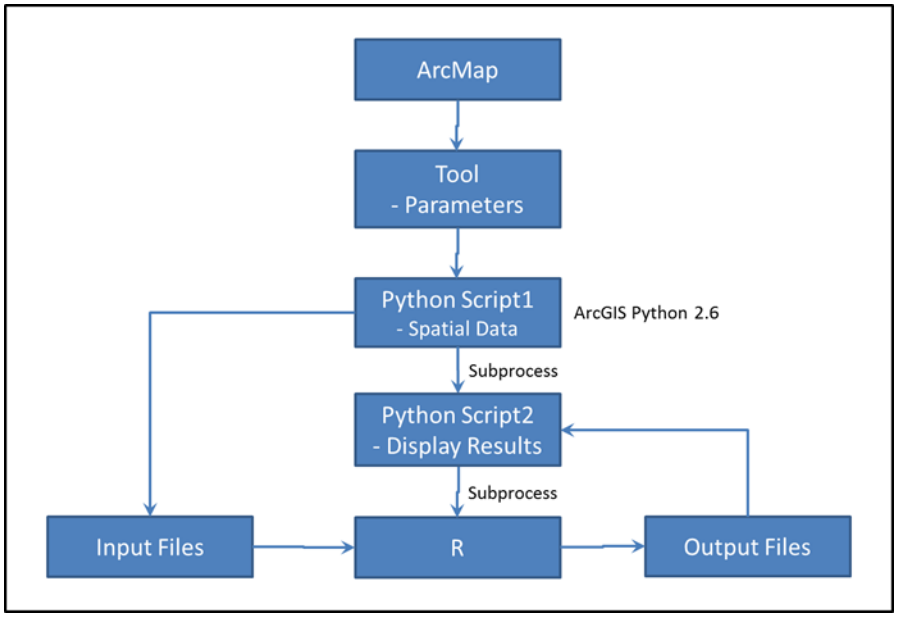
  Data Flow Diagram:



- Statistic Data Analysis

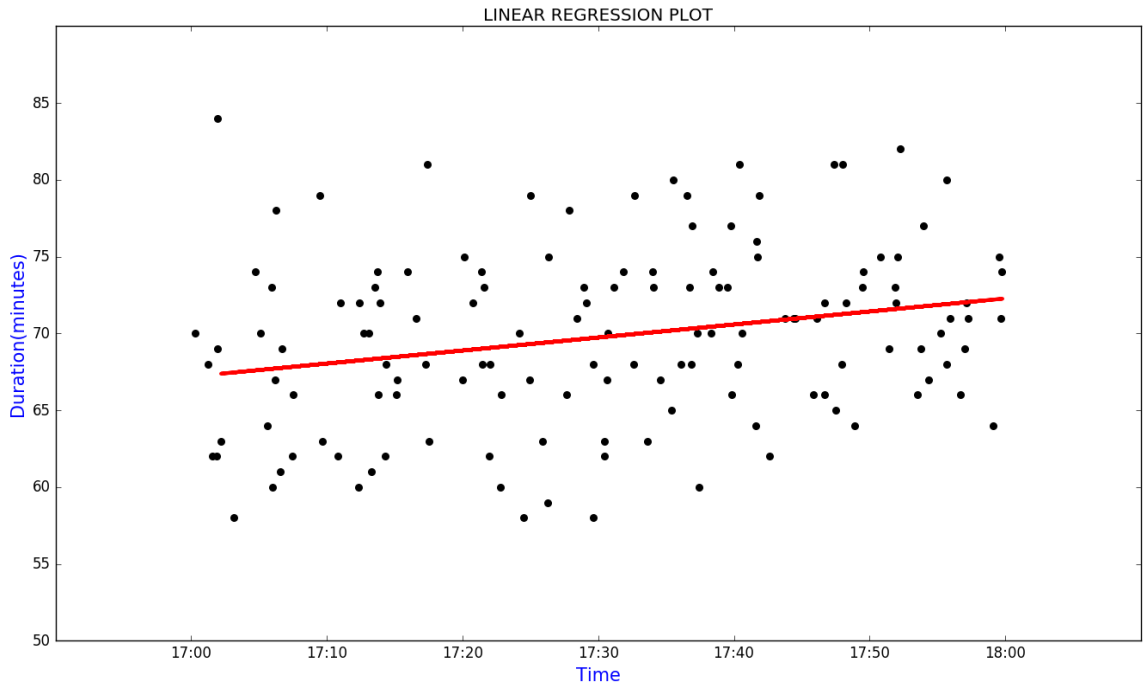


- UI Flowchart

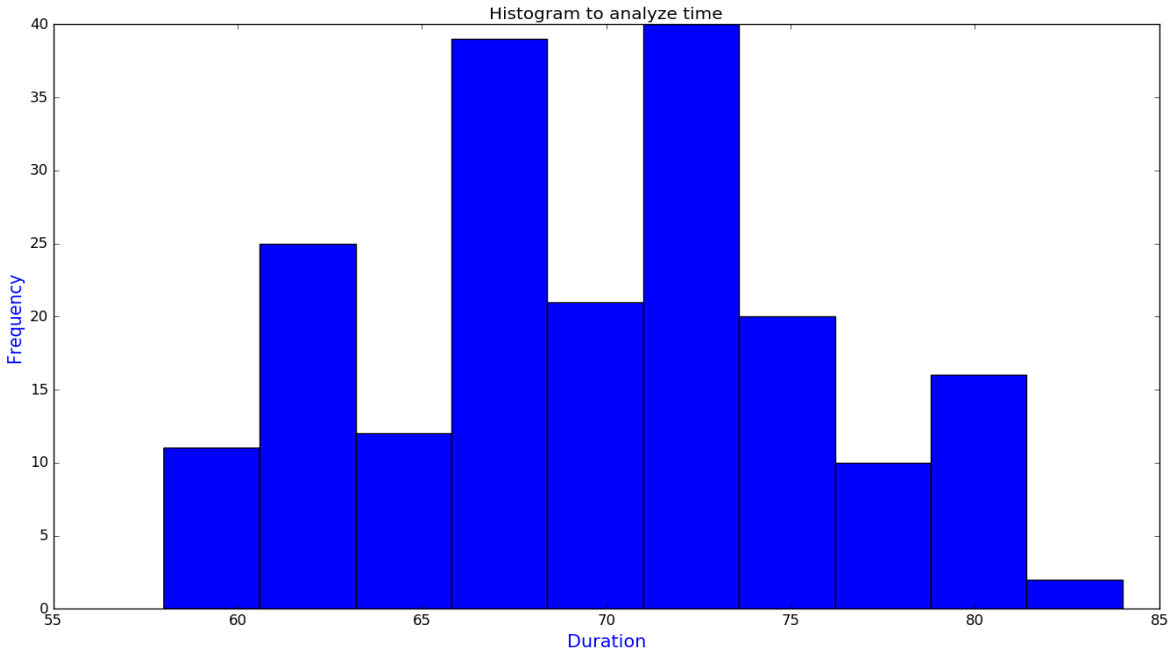


5. DISCUSSION ON IMPLEMENTATION RESULTS

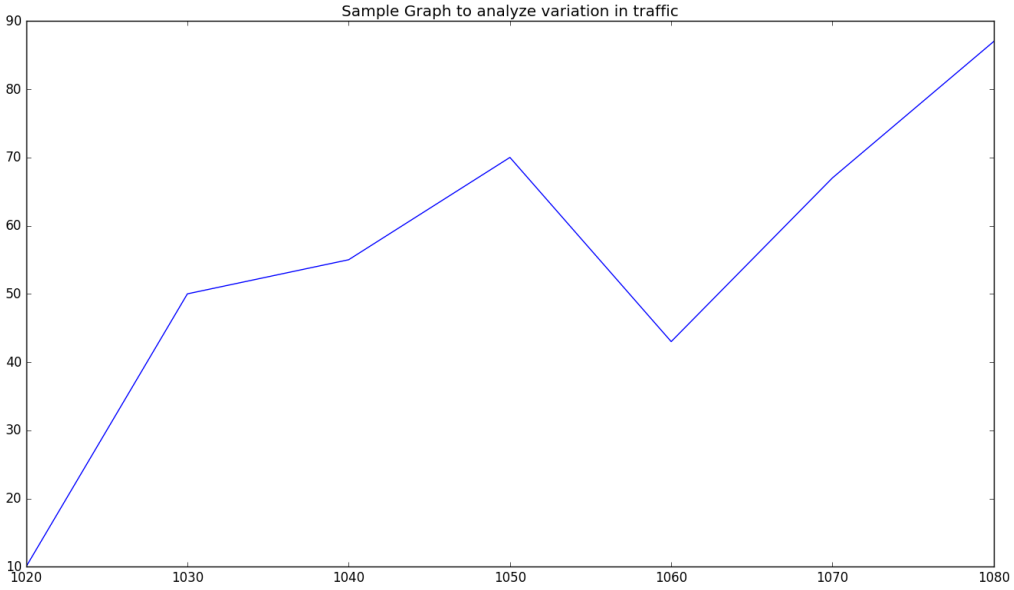
The following is linear regression graph which shows the predictive model for travel time predictor. This graph has been plotted for the time slot 5:00 PM to 6:00 PM. The black dots in the graph indicates the learning dataset. Based on the dataset, the regression line has been modelled and plotted which predicts the output.



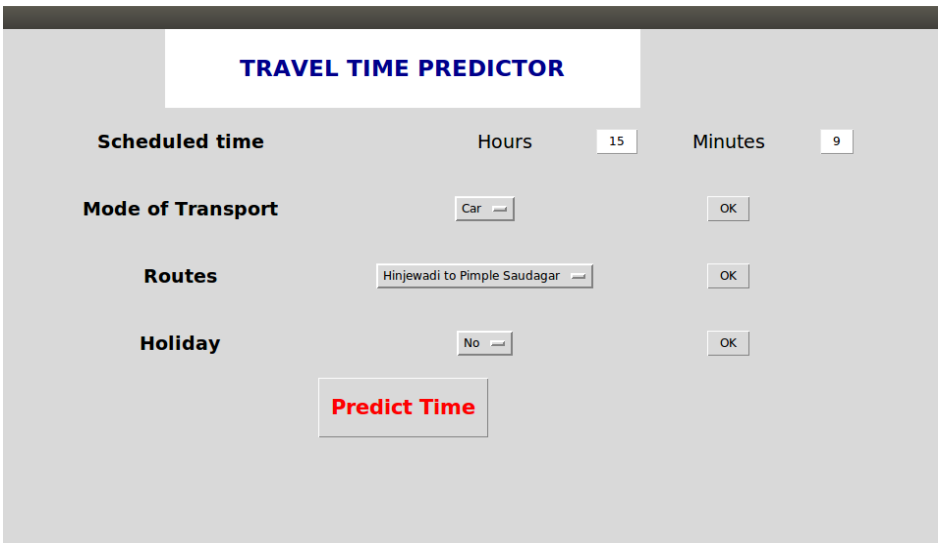
The following is the histogram depicting the frequency of the time required by car for the same time slot as given above. We observe that the frequency of time taken is high around 65 to 75 minutes which means in future the time taken would be around 65 to 75 minutes.

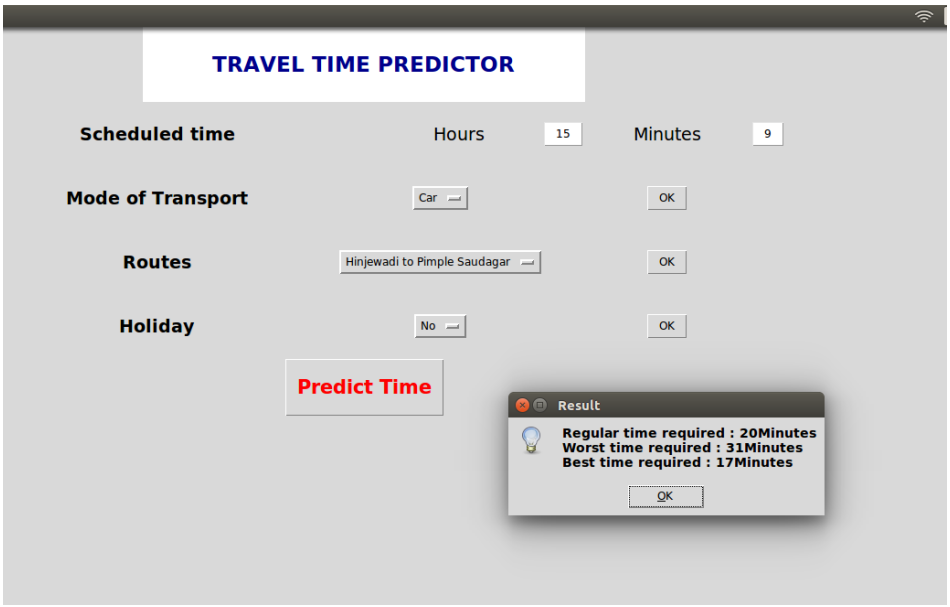


This is the sample plot showing variation in the traffic

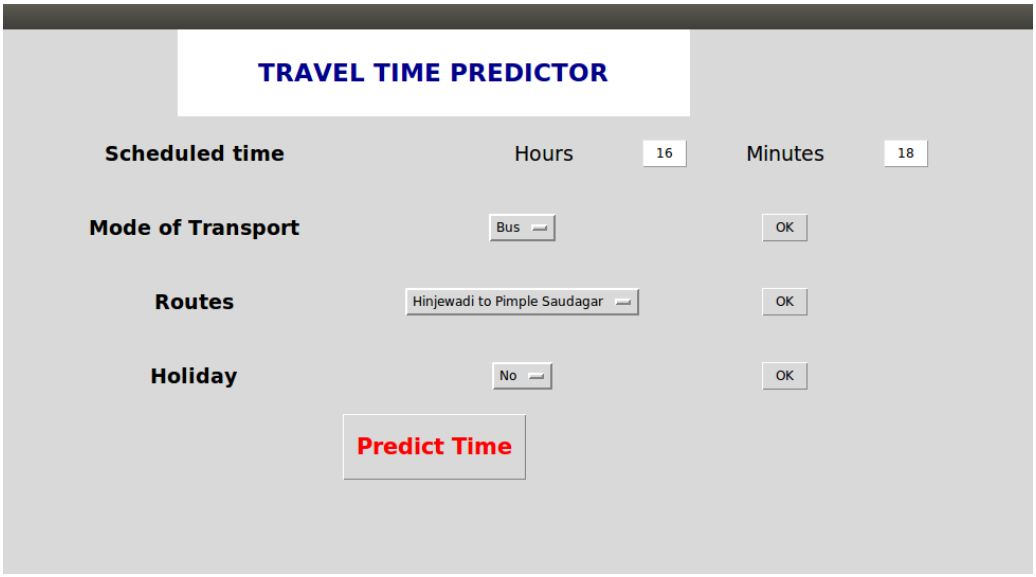


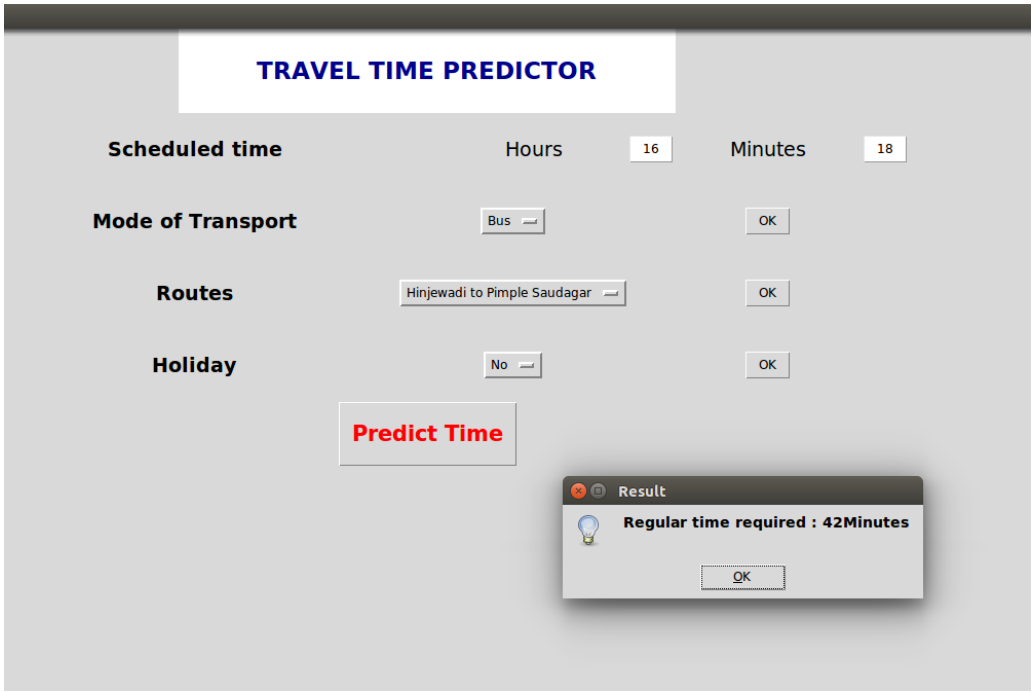
- Screenshots
  - For Car





- For Bus


TRAVEL TIME PREDICTOR
Scheduled time
Hours
16
Minutes
18
Mode of Transport
Bus
OK
Routes
Hinjewadi to Pimple Saudagar
OK
Holiday
No
OK
Predict Time


TRAVEL TIME PREDICTOR
Scheduled time
Hours
16
Minutes
18
Mode of Transport
Bus
OK
Routes
Hinjewadi to Pimple Saudagar
OK
Holiday
No
OK
Predict Time
Result
Regular time required : 42Minutes
OK

6. CONCLUSION AND FUTURE ENHANCEMENT

Prediction of your future travel time is crucial if we have to plan a trip or have important meetings and appointments in future. We have successfully predicted the Travel Time for the predefined routes. The travel time has been predicted for car as well as by bus. During holidays, the time taken for the respective travel is less as compared to that taken during working days. Analysis for the entire day is also represented through a histogram.

7. REFERENCES

http://scikit-learn.org
https://www.python-course.eu/python_tkinter.php
https://www.datacamp.com