Analysis of Road Congestion based on Weather Conditions in Sri Lanka

Submitted by:

A.L.M.D. Silva

158776R

Supervised by: Mr. B.H. Sudantha

Faculty of Information Technology

University of Moratuwa

February 2019

Analysis of Road Congestion based on Weather Conditions in Sri Lanka

Submitted by:

A.L.M.D. Silva

158776R

Supervised by: Mr. B.H. Sudantha

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfillment of the requirements of the Master Degree of Science in Information Technology

February 2019

Declaration

I declare that this is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of reference is given.

Name of the student : A.L.M.D. Silva
Student Number : 158776R
Signature of the student : …….
Date : ……

Supervised by:
Name of the supervisor : Mr. B.H. Sudantha
Signature of the supervisor : …….
Date : …….

Dedication

This Dissertation is dedicated to my loving parents for being part of me and encouraging me always by being my side.

Acknowledgement

First I express my heartfelt gratitude to my supervisor Mr. B.H Sudantha for his most valued guidance, commitment and kind support to make this research a success.

Also I sincerely thank Prof. Asoka S Karunananda who taught us all the research related document preparation which was a great support to manage all work with busy schedules.

It's my pleasure to thank Mr. Chaman Wijesiriwardena and all other Senior Lecturers, Lecturers, Instructors and staff members who helped us in many ways to successfully complete this research.

And also I sincerely thank Mrs. Chamari De Silva who taught me about rapid miner tool which was a great support me to do the analysis part of the research.

Then I would like to thank all the batch mates of MSc. In Information Technology batch 9 for their companionship and various kinds of support given throughout the program.

In addition, I would like to thank my work mates and friends for encouraging me with their support and best wishes.

Last but not least, my sincere thank goes to my beloved parents and specially Lasith Udara, for helping me to conduct this work without much stress and encouraging me to complete this research.

Abstract

Sri Lanka incurs a huge economic loss of around Rs.4 billion annually due to the road traffic congestion and air pollution with too many vehicles on a limited road network. According to the statistics a large number of Sri Lankans spend more time on the roads paying more fuel, as the number of vehicles on the roads are rapidly increasing and it takes longer to reach one's destination. Road congestion is increasing due to many reasons such as increase of vehicle population, lack of proper vehicle parking system near urban area, lack of well-maintained road network, weather condition, etc. There are many researches which is done to monitor the road congestion based on the vehicular data. And the lack of monitoring the traffic based on the weather condition is available in Sri Lanka.

This research is mainly focused on the weather condition and the traffic data. To collect the information, google maps data, open weather Map data and police report data is supposed to use. To find the correlation between classification of the traffic congestion and weather, data is analyzed. To predict the most relevant weather factor for the road congestion, data is analyzing.

Aim of the research is provide a best route to travelers. System is to be analyze how the change of weather affects increase of road condition and in future analyze to be done to find how traffic congestion affects the increase of accidents.

Content

List of Figures

List of Tables

Chapter 1

Introduction to Analysis of Road Congestion based on Weather Conditions

1.1 Introduction

In order to gain the optimum utilization of available resources through the aggregate national production, the economy must fulfill some of the fundamental requirements. Out of that, supply of man and material to the production system at the right time, at the right place and at the right amount is very much significant. In this context, if there is no such a delays in between the location of men & material and the production place, then it can be defined as an availability of an effective transportation system[1]. But traffic growth has become a major problem for both developing and developed countries. Traffic congestion is a serious issue which leads to number of negative consequences including higher transit times, loss of productivity and environmental pollution. The concept of forecasting the road network in terms of traffic loading and flow, is an accepted approach world-wide.

Sri Lanka incurs a huge economic loss of around Rs.40 billion annually due to road congestion and air pollution with too many vehicles on a limited road network. According to the countrymeters[1], Sri Lanka with a population of 20 million in the year 2012 is projected to have a population 21 million in the beginning of 2020 and a projected annual growth is 0.33%. With the rapid growth of the population, it is expected that the demand of transportation is increasing accordingly.

The main reason for traffic growth is the available number of vehicles is bigger than the available road capacity. In addition to that there are multiple reasons for the traffic growth:

- Vehicle growth

[1]http://www.motortraffic.gov.lk/web/images/stories/document/new%20regist%20of%20vehicles%202015.pdf

- Road condition
- Driving behavior of humans[2]
- Mechanical failures of vehicles
- Bad weather
- Insufficient traffic system
- Poor city planning
- Inappropriate public transport facilities

According to the statistics of MotorTraffic[2], Sri Lanka with vehicle population of 326651 in the year 2013 and 429556 in the year 2014 and 669198 in the year 2015 and the projected annual traffic growth is nearly 31% in the period of 2013-2015. But in 2016 vehicle population is decreased by 85%. Shifting from private to public transport may tackle the problem of congestion due to the increase of vehicles. With this scenario, it is expected that most of the people are using the public transport and reduced the traffic. However, Traffic growth is increasing rapidly.

According to the statistics of RDA[3], National highways in Sri Lanka has divided into three types of road types such as class A, B and E. Class B covers more than 60% from the total length and road network is not well-maintained. Due to that reason road users cannot fulfill their work properly.

Recently most of the Sri Lankans got effected from the natural disasters. Such as flooding, land slide, etc. In such situations most of the routes cannot use properly. In Sri Lanka, Weather condition is different from one area to another. In some situations, Roads cannot use due to the unexpected behavior of the weather. Therefore, it is very import if there is a method which are available to measure those parameters and give the instructions to users about the best path.

[2]https://countrymeters.info/en/Sri_Lanka

[3]http://www.rda.gov.lk/source/rda_roads.htm

1.2 Background of the study

1.2.1 Types of congestion

According to the geotab[4], there are two types of traffic congestion: recurring and non-recurring. Recurring means which happens daily due to the lack of capacity of the road. Other type of traffic is happened due to unexpected disruptions. Such as bad weather and the vehicle collision.

Road congestion is caused when there are more vehicles than available space on the road. This happens when population grows faster than the city infrastructure. Figure 1[4](below) shows the population growth in 1959-2019.

Year	Population	Yearly % Change	Yearly Change	Migrants (net)	Median Age	Fertility Rate	Density (P/Km²)	Urban Pop %	Urban Population	Country's Share of World Pop	World Population	Sri Lanka Global Rank
2020	**21,084,042**	0.35 %	74,000	-90,000	34.1	2.03	336	19.9 %	4,194,261	0.27 %	7,795,482,309	58
2025	**21,349,942**	0.25 %	53,180	-75,000	35.7	1.96	340	21.0 %	4,488,307	0.26 %	8,185,613,757	63
2030	**21,474,701**	0.12 %	24,952	-75,000	37.1	1.91	342	22.7 %	4,867,530	0.25 %	8,551,198,644	66
2035	**21,491,895**	0.02 %	3,439	-75,000	38.4	1.86	343	24.9 %	5,352,803	0.24 %	8,892,701,940	68
2040	**21,397,716**	-0.09 %	-18,836	-75,000	39.6	1.83	341	27.8 %	5,943,636	0.23 %	9,210,337,004	71
2045	**21,165,080**	-0.22 %	-46,527	-75,000	41.0	1.81	338	31.0 %	6,564,628	0.22 %	9,504,209,572	75
2050	**20,792,352**	-0.35 %	-74,546	-75,000	42.6	1.80	332	34.6 %	7,190,065	0.21 %	9,771,822,753	76

Table 1: Sri Lanka Population Forecast

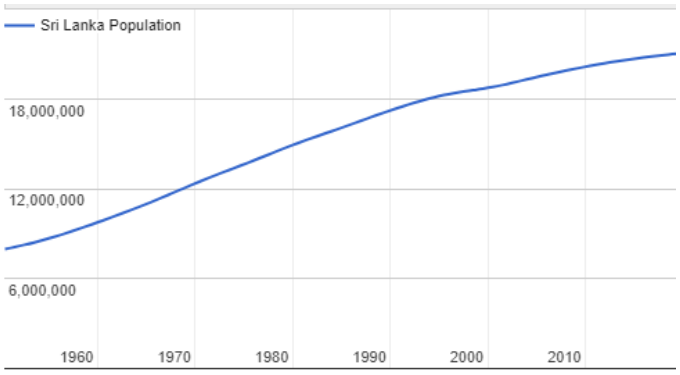


Figure 1: Sri Lanka Population

[4]http://www.worldometers.info/world-population/sri-lanka-population/

Therefore, it is expected that demand of transportation is accordingly increasing with the population. According to the motor traffic, vehicle population is increased in the period of 2008-2015 and it is decreased in the period of 2015-2016.

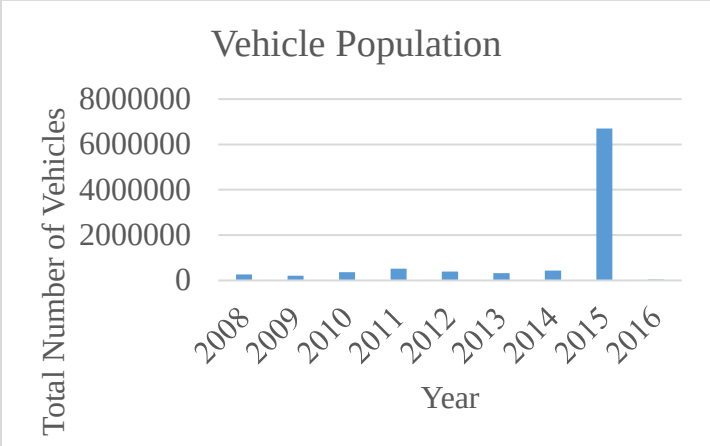


Figure 2 : Vehicle population forecast in Sri Lanka

With the rapid growth of the population, roadway saturation is the critical issue that lawmakers have sought to address for years.

Road accidents are essentially caused by improper interactions between vehicles and other users. This could be the result of a number of factors such as driver's characteristics, vehicle design, road user's behavior, geometric features, environmental aspects, traffic characteristics and pavement characteristics. Road accident may be blocked the whole road or side of the road[2]. Therefore, it disrupts the flow of the traffic.

Weather condition interrupts the free flow of traffic because it makes drivers more cautious. Because speed of the moving may be slow due to the rain, snow or ice in order to decrease the risk of crashing.

Traffic may be caused by the distracted driving. When drivers are distracted by their smartphones or other handheld devices, they might not drive at a constant speed. Thus, they might not see the traffic signals. According to the aaafoundation[5], smartphone distraction at traffic lights can negatively impact regular traffic flow for an average of 27 seconds after you've stopped texting. Therefore, distracted driving may increase the traffic and also risk for crashing.

[5]https://aaafoundation.org/objectively-derived-self-reported-measures-driving-exposure-patterns-older-adults-longroad/

1.2.2 Weather Condition

Recently most of the Sri Lankan people got effected from natural disasters. According to the situation updated by the Disaster Management Centre (DMC) [6] of Sri Lanka confirmed that a total of 1,537,122 people in districts have been affected and 20 people have died due heavy rains, strong winds, lightning and landslides in May 2018. Landslides warnings is given to effected areas and it is restricted to live in that areas. In that situation, traffic growth is increased.

Extreme weather conditions can change the way of normal driving. Because it effects to the ability of see which is in the adverse weather conditions such as rain, fog, dust. In that conditions drivers have to drive slowly[3].

1.2.3 Impact of traffic congestion

Sri Lanka loss 1.5% of the GDP (Gross Domestic Product) due to traffic congestion[7]. Because the road network in the Sri Lanka is not capable of handling the increasing of traffic flow. A large number of Sri Lankans spend more time on the roads paying more fuel as the number of vehicles on the road is rapidly increasing and it takes longer to reach one's destination.

According to the Figure 2, new registration of vehicles decreased during 2017. Due to the rising of prices of fuel (Figure 3), shifting from private transportation to public transportation has increased.

Figure 3: Average price of Crude oil(Brent) in the international market and the crude oil import price of the CPC

[6]http://www.dmc.gov.lk/index.php?lang=en

[7]https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/publications/annual_report/2017/en/7_Chapter_03.pdf

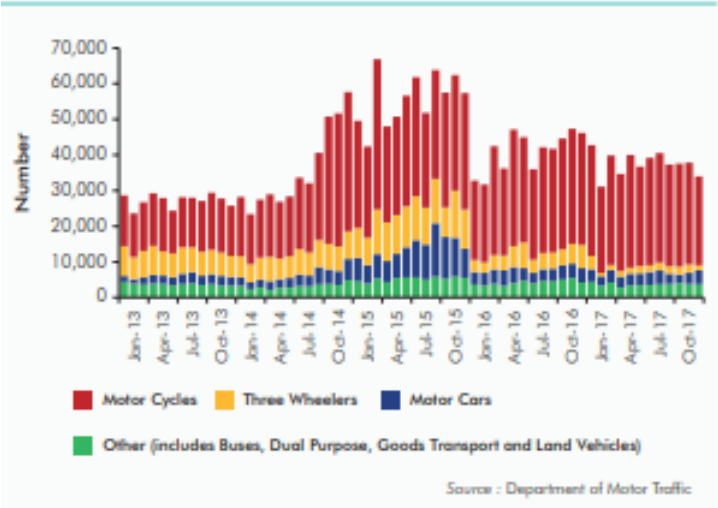


Figure 4: New Registration of Motor Vehicles

The total operated kilometers of the Sri Lanka Transport Board (SLTB) decreased marginally by 0.8% to 448.1 million km while passenger kilometers also decreased by 1.8 percent to 15.8 billion km in 2017, in comparison to 16.1 billion km in 2016. During 2017, operated kilometers of private buses decreased by 2.1% to 1 billion km and passenger kilometers also declined by 2.4% to 50.8 billion km, in comparison to the previous year. The reduction in terms of bus operations for both public and private operators was mainly due to the impact of floods in many parts of the island during May 2017[7]. Therefore different types of adverse weather has a different impact of traffic parameters[4].

The effect of traffic is growing. Without implementing new strategies to mitigate congestion, congestion will continue to grow with negative impact of our lives.

1.3 Motivation

In the past, low enforcement for traffic couldn't address the problem. Without implementing new strategies cannot find the solution for traffic congestion.

Most of the researches, the number of vehicles which are passing through a section are used to determine the flow of traffic. In Vietnam, A traffic monitoring system is used to monitor the road traffic by using a CCD camera which is mounted to view the rear of moving vehicles and gaps between vehicles are used to determine the density of vehicles appearing in the camera span[5].

Philippines have presented a violation checking and recording system by using an Accurate Name Plate Recognition to handle the traffic system[8]. Overall procedure of the system consists of following steps:

- The user inputs a desired destination on the app
- The GPS feature fetches the driver's location
- En route to the destination, the app continuously checks both terrain and weather condition

If ever a terrain hazard/ weather condition is nearby, displays warning and suggested speed reduction, the system is helped to minimize the traffic and road accidents.

[6] proposed a method which utilizes mobile devices (Smart phones) to detect passing vehicles and identify their class (personal cars, semi-trucks and trucks). Bluetooth beacons which are placed on opposite sides of road are received radio signals from mobile devices and then it analyzes the strength of the radio signal to monitor traffic.

[7] Proposed a model which uses infrared proximity sensors and a centrally placed microcontroller and uses vehicular length along a length to implement intelligent traffic monitoring system.

Traffic congestion is creating economic and social issues such as economic loss, wasting time of drivers and passengers, wasting fuel, reduced the health of drivers and passengers and increase air pollution. As a developing country, spending more money to change the road infrastructure is impossible. We can collect data from available sources and find the path to minimize the effect of traffic congestion in Sri Lanka.

1.4 Problem in brief

All above research described different approaches used to minimize the traffic congestion but we won't find any framework which are designed for Sri Lanka. I would like to suggest a model or framework which can be used with weather and traffic data to provide awareness of actual situation to drivers.

To collect information, we can use google map and openweather app because they are freely available. Analyzing the data collectively we can attempt to an affordable solution.

1.5 Aim and Objectives

Aim: Analyze the road congestion based on weather conditions and provide a model to use the analyzed data in decision making

Objectives of the research are:

1. Identify the techniques and tools can be used to extract data in road traffic and weather
2. Develop a customized application to download the data
3. Identify methods to be followed to clean extracted data
4. Discover features of data mining such as classification and correlation.
5. Evaluate features
6. Develop a model

1.6 Proposed solution

An automated system for analyze the road congestion based on weather requires information of the origin and destination as inputs. Then it analyzes the information through the given route in past few minutes and provide awareness of the actual situation as the output.

1.7 Summary

In this chapter it is identified the importance of analyzing traffic congestion data and weather data collectively and how we can use those data to derive meaningful information about the situation. Further what others have done in this traffic congestion area is presented in this chapter in brief and in detail information will be given in the next chapter.

Chapter 2

Literature review on analysis of traffic congestion

2.1 Introduction

In chapter 1, we describe what traffic congestion is, what are the factors effect for traffic congestion, how the traffic flow is changing with the weather conditions including motivation for the study as well as research objectives. Further there was a brief introduction to problem of research and importance of analyzing traffic data and weather data collectively.

This chapter critically reviews the problem of road congestion. In this sense, first we identify about the factors which are affects to road congestion and related work done by other researches in the discipline. Subsequently, special applications which are used to analyze the traffic related data will be discussed.

Having knowledge of the flow of traffic congestion improve our understanding of the performance of an entire country or region's transportation network. During the planning, design, construction and maintenance period of the road network, traffic information will be important in decision making.

Over the past 50 years, a lots of research have been done in this area[8]. Traditional methods and intelligent systems have been modeled and introduced to monitor traffic flow and analyze factors which influence for traffic congestion. Therefore the researchers conducted in this area of study can be categorized as:

- Development of technology to monitor and analyze traffic flow
- Effect of adverse weather on traffic flow and driving behavior
- Issues related in available technologies
- Highlights of literature

Therefore, the chapter 2 describes in detail review of literature related to above categories and finally the research problem will be highlighted with limitations in related work and importance of having a proposed system to assist decision making in Sri Lanka.

2.2 Development of technology to monitor and analyze traffic flow

A good, well maintained road network helps economic development of a country as well as the well-being of its citizens. Therefore, accurate description and classification of traffic congestion involves to improve our standing of flow of traffic and performance of the road network.

Traditional regional congestion measures were mostly based on travel survey data, including travel delay and travel time, like National Household Travel Survey (NHTS) in the US and Person Travel (PT) survey in Tokyo metropolitan area. In recent years, with the popularity of "Big Data," more accurate regional congestion measure became possible. INRIX provided detailed traffic speed data every 800 feet (250 meters) across 4 million miles of roads in 40 countries. Powered by INRIX's data, the Texas Transportation Institute annually developed a measure of traffic congestion for many urbanized areas. However, for most developing countries like Sri Lanka, real-time traffic data is either not published due to restrictions or not collectable owing to lack of devices[8].

Generally, there are three ways to analyze factors which effect to traffic congestion.

- Build a model and solved mathematically
- Statistical analysis
- Case study

[9]Presents a Dynamic Traffic Assignment Model which is based on two assumptions such as the time spent by a vehicle on a link may be decomposed into a fixed travel time plus a waiting time. The fixed travel time corresponds to the free or uncongested travel time over the link. Then the vehicle is put in an exit queue until it becomes possible to enter a forward link (it is based on the link costs and their capacities. This model shows an expansion of the network and solution can be obtained using an algorithm.

[10] Describes Intelligent Transport System based on computer vision techniques. System uses video cameras for the detection and classification of vehicles using classical visual surveillance techniques such as background estimation and motion tracking for some time.

But evaluation under weather conditions is challenging. [11] Presents a system for detecting vehicles by means of image analysis and rule-based reasoning. Difference between this system with the above-mentioned system is a formal separation between daytime vehicle tracking (low-level image processing module) and night time vehicle tracking (high level module is designed as a forward chained production rule system, working on symbolic data and exploiting a set of heuristic roles tuned to urban traffic conditions). Image processing modules are used to detect moving vehicles in day time and morphological analysis of headlight pairs are used to detect moving vehicles in night times.

Global Positioning Systems(GPS) are being used for a wide variety of applications [12]. Google maps is a GPS based vehicle navigation system. Google Transport Tracker is a set of applications designed to track a range of moving assets (such as vehicles) and visualize them on a live map. The applications use a mixture of technologies - Android, Firebase, Google Maps, GTFS (General Transit Feed Specification), and more. However, Google Transport Tracer does not provide information based on the size of the vehicle. Sometimes we are not able to use the routes which they provide because of the road surface condition and the natural disasters such as flooding, landslide.

Most developing countries do not have well maintained road networks there by hindering the development of those countries. The lack of funds is certainly a factor that contributes to the deteriorated road networks, but the lack of proper monitoring and reporting systems is also a major contributor. To address this issue BusNet architecture is proposed. The design of the BusNet is based on the observation that environmental monitoring systems need a large number of sensor nodes spread over a large area to be effective [13].

Nowadays, there is an increasing interest in using Virtual Reality Geographical Information System (VRGIS), which can obtain the landscape geospatial data dynamically, and perform rich visual 3D analysis, calculations, managements based on Geographical Information System (GIS) data [14]

Web based systems[9] are available to monitor traffic congestion such as OpenStreet map, Google maps and Bing maps. Open StreetMaps and Google Maps provide the updated and

current road configuration. Bind Maps is older than others. But it doesn't support some areas. However, Google Maps shows a complete shifting north-wards of approximately 10 meters on an average of all line features in the given area.

2.3 Effect of adverse weather on traffic flow and driving behavior

Few number of researchers have compared the effect of reduced visibility due to fog and rain on the traffic parameters using field data since the vehicle based traffic data and weather data especially the reduced visibility data under fog condition are hard to collect[4]. Most researchers have conducted driving simulator-based studies in order to identify the effect of adverse weather including fog and rain separately.

[15] Describes how the weather is affected to many aspects of transportation such as traffic demand, traffic safe and traffic operation and flow. it was found that severe winter storms bring a higher risk of being involved in a crash by as much as 25 times–-much higher than the increased risk brought by behaviors that state governments already have placed sanctions against, such as speeding or drunk driving. Finally, it presents new estimates of capacity and speed reduction in extreme weather conditions.

[16]Describes visibility of drivers is changed in day, night and rain using driving simulator and shows low visibility in rain and night caused to the risk of crash. [17] Explains how the foggy conditions influence to the speed. A hierarchical method is proposed to examine fog influence on speed behaviors. And it highlights following factors.

- Drivers could not response timely to the impending changes in road geometry in fog.
- Driver's speed compensation could not sufficiently reduce the crash involved risk.

Some researchers have found that rain is effect to the road congestion. [4] Describes how the rain and fog effect to the speed and visibility. The research has done using ANOVA method.

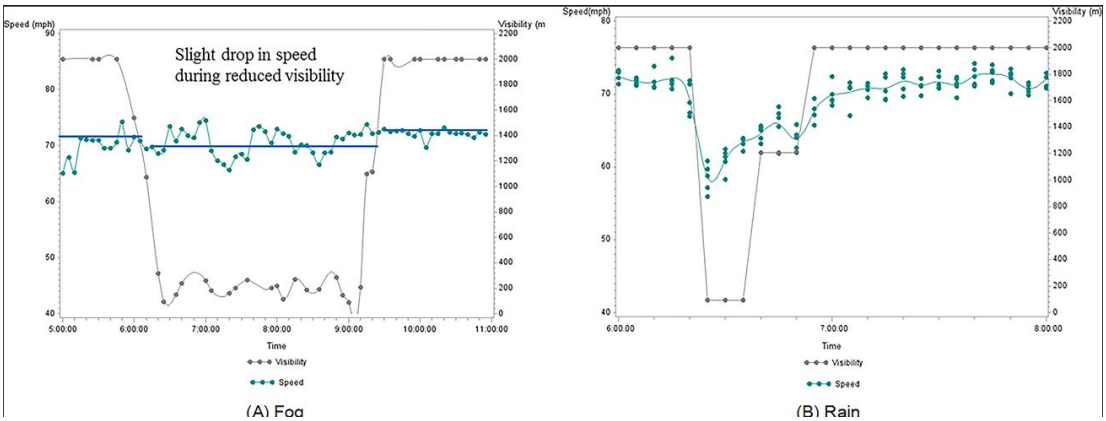


Figure 5 : Comparison of speed and visibility during fog and rain

[18]Describes heavy rain can significantly increase the variance of lane position. But some researchers defined that it is because of the low visibility of the drivers. [19] Describes surface condition of the road is effect to the road congestion.

All of the research have been done in various countries in different weather conditions. But weather culture of the Sri Lanka is different from other countries. Therefore, there is a need of further investigation that will be useful to propose a solution for the traffic congestion based on extreme weather conditions in Sri Lanka.

2.4 Issues related in available technologies

Due to lack of knowledge of using smart devices and intelligent traffic systems, adopting a new technology may take some period. Changing the infrastructure also impossible in Sri Lankan context. By defining new policies traffic congestion and related accidents cannot be control. Considering the popularity, availability and maintainability web based platform can be used to analyze the traffic in Sri Lanka.

[20] Shows that the number of reported accidents to the police has increase from 26,196 in 1989 to 52,444 in 2005. It mentioned that poor development of the road and the increase of number of vehicles on the road mainly effect to accidents. Further paper emphasis that during the extreme weather conditions accident rate are growing.

2.6 Problem definition

Road network of Sri Lanka is divided into 4 classes. Such as E, A, B and C.

Class	Average Speed	Number of lanes
E	100kmph	4 or 6
A	70kmph	4
B	60kmph	2
C	50kmph	

Table 2: Classification of road network in Sri Lanka

Based on the average speed and the number of lanes traffic congestion is varying in each road. But the traffic congestion in A and B roads are maximized. This may be due to the lack of vehicle parking facility.

According to the statistics number of vehicles on the road is rapidly increasing (Table 3). With the limited road network this may leads to the traffic.

NEW REGISTRATION OF MOTOR VEHICLES

CLASS OF VEHICLE	2008	2009	2010	2011	2012	2013	2014	2015	2016
MOTOR CARS	20,237	5,762	23,072	57,886	31,546	28,380	38,780	105,628	3,480
MOTOR TRICYCLES	44,804	37,364	85,648	138,426	98,815	83,673	79,038	129,547	3,405
MOTOR CYCLES	155,952	135,421	204,811	253,331	192,284	169,280	272,885	370,889	22,225
BUSES	1,180	739	2,491	4,248	3,095	1,805	3,851	4,140	236
DUAL PURPOSE VEHICLES	2,856	1,280	11,712	33,518	37,397	24,603	20,799	39,456	1,830
MOTOR LORRIES	13,588	7,823	10,803	12,446	10,445	4,525	3,851	5,356	345
PRIME MOVERS	106	183	293	573	283	139	170	316	19
LORRY TRAILERS	106	110	292	495	299	149	163	210	9
LORRY OTHERS	76	23	389	1,224	1,143	756	808	1,214	71
AMBULANCES	141	71	53	58	66	292	111	32	4
HEARSES	21	15	15	22	30	11	18	14	-
L.V.TRACTORS	24,063	13,765	17,284	20,073	18,450	10,772	7,070	9,977	882
N.A. TRACTORS	294	186	79						
L.V.TRAILERS	1,775	1,333	2,301	3,121	3,442	2,266	2,012	2,128	295
TOTAL	265,199	204,075	359,243	525,421	397,295	326,651	429,556	668,907	32,801

Total Vehicle Population

CLASS OF VEHICLE	2008	2009	2010	2011	2012	2013	2014	2015	2016
MOTOR CARS	381,448	387,210	410,282	468,168	499,714	528,094	566,874	672,502	675,982
MOTOR TRICYCLES	406,531	443,895	529,543	667,969	766,784	850,457	929,495	1,059,042	1,062,447
MOTOR CYCLES	1,760,600	1,896,021	2,100,832	2,354,163	2,546,447	2,715,727	2,988,612	3,359,501	3,381,726
BUSES	81,050	81,789	84,280	88,528	91,623	93,428	97,279	101,419	101,655
DUAL PURPOSE VEHICLES	196,236	197,516	209,228	242,746	280,143	304,746	325,545	365,001	366,831
MOTOR LORRIES	263,407	271,230	282,033	294,479	304,924	309,449	313,300	318,656	319,001
PRIME MOVERS	2,842	3,025	3,318	3,891	4,174	4,313	4,483	4,799	4,818
LORRY TRAILERS	6,150	6,260	6,552	7,047	7,346	7,495	7,658	7,868	7,877
LORRY OTHERS	1,887	1,910	2,299	3,523	4,666	5,422	6,230	7,444	7,515
AMBULANCES	2,019	2,090	2,143	2,201	2,267	2,559	2,670	2,702	2,706
HEARSES	317	332	347	369	399	410	428	442	442
L.V.TRACTORS	244,990	258,755	276,039	296,112	314,562	325,334	332,404	342,381	343,263
N.A. TRACTORS	693	879	958	958	958	958	958	958	958
L.V.TRAILERS	42,823	44,156	46,457	49,578	53,020	55,286	57,298	59,426	59,721
TOTAL	3,390,993	3,595,068	3,954,311	4,479,732	4,877,027	5,203,678	5,633,234	6,302,141	6,334,942

Table 3: Change of vehicle population in 2008-2016

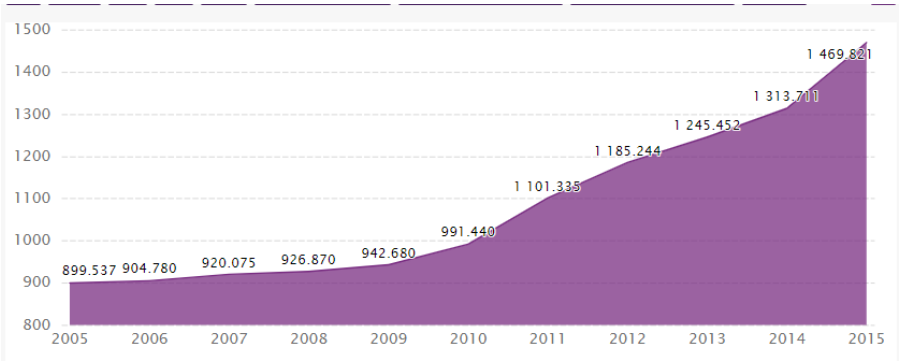


Figure 6: Growth of vehicle population

In Sri Lanka there is an extensive road network spanning the island. New roads are being built and the existing roads are upgrading every day. Vehicle consumes less fuel on a well-maintained road surface. Smooth movement of traffic on such a surface also cut the traveling time[13]. As a tropical country with two heavy monsoon seasons weather is also affects for the road traffic congestion in Sri Lanka. Lack of road monitoring and reporting system in such a situation is available.

Implementing a solution using devices is impossible in Sri Lanka. Defining a web-based solution to aware drivers about the current situation can reduce the road congestion free of charge. By considering the accuracy and availability of data, google map can be select to extract traffic data and open weather map can be used to fetch weather data. Then data can be classify using naïve bayes, k-NN algorithm and decision tree algorithm to predict the hypothesis in rapid miner.

2.7 Summary

This chapter presented a comprehensive critical review of available traffic monitoring and analyzing system with a specific reference to web services. We defined the research problem as analysis of road congestion based on weather condition. Addition to that we identify the factors that cause for the traffic congestion. We also identified the possible technology that can be used to address the research problem. Next chapter will discuss the technology adapted for solving our research problem.

[9]http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.706.8928&rep=rep1&type=pdf

Chapter 3

Technology Adapted in Analysis of Road Congestion based on Weather Conditions

3.1 Introduction

Chapter 2 described similar research work in details with related to usage of road traffic analysis on decision making. Further chapter 2 described approaches followed by other researches including tools and techniques used to collect data, monitor data analyze data to support traffic analysis. Therefore, at the end of the chapter 2 we can conclude that weather details are very important to analyze the traffic congestion.

3.2 Technologies to collect data in Road Traffic and weather

Due to availability in Sri Lanka, google map and open weather will be considered here to collect data. But limitations are there to use those tools to extract data such as google map doesn't provide historical data. Open weather provides historical data but not available for free of charge. Those tools don't provide offline data.

Application programming interfaces are available for both google map distance matrix and open weather and all the tools extract data through these APIs. Current data can be extracted using those tools. Using those API, JavaScript and P5.js library file, Customized application is developed to download data. Because google map doesn't provide any plat foam to download data.

Other than above mentioned tools, allow-control allow-origin developer tool is used to request data from the google map.

3.2.1 Google map distance matrixAPI

The Google Map Distance Matrix API is used to extract traffic data from google map. It works with a key. It features four basic map types such as roadmap, satellite, hybrid and terrain which can be modify using layers and styles, controls and events, and various services and libraries.

3.2.2 Open Weather Map API

OpenWeatherMap API provide current and historical weather data for any location including over 200,000 cities. Data is available in JSON, XML or HTML format and can be downloaded in csv format.

3.3 Data mining

Data mining is the practice of searching large stores of data to discover patterns and trends that go beyond simple analysis.

3.3.1 Data Classification

Data classification is a data mining (machine learning) technique used to predict group membership for data instances and helps to find the accuracy of the data set.

Machine learning can be further subdivided into Supervised and Unsupervised learning.

- Supervised learning - such as Regression Trees, Discriminant Function Analysis, Support Vector Machines.
- Unsupervised learning - such as Self-Organizing Maps (SOM), K-Means.

Classification has a two-step process [21]

1. Model construction- describe a set of predetermined classes and the model is represented as classification rules, decision trees, or mathematical formulae
2. Model usage- test data set is created to accuracy of the model

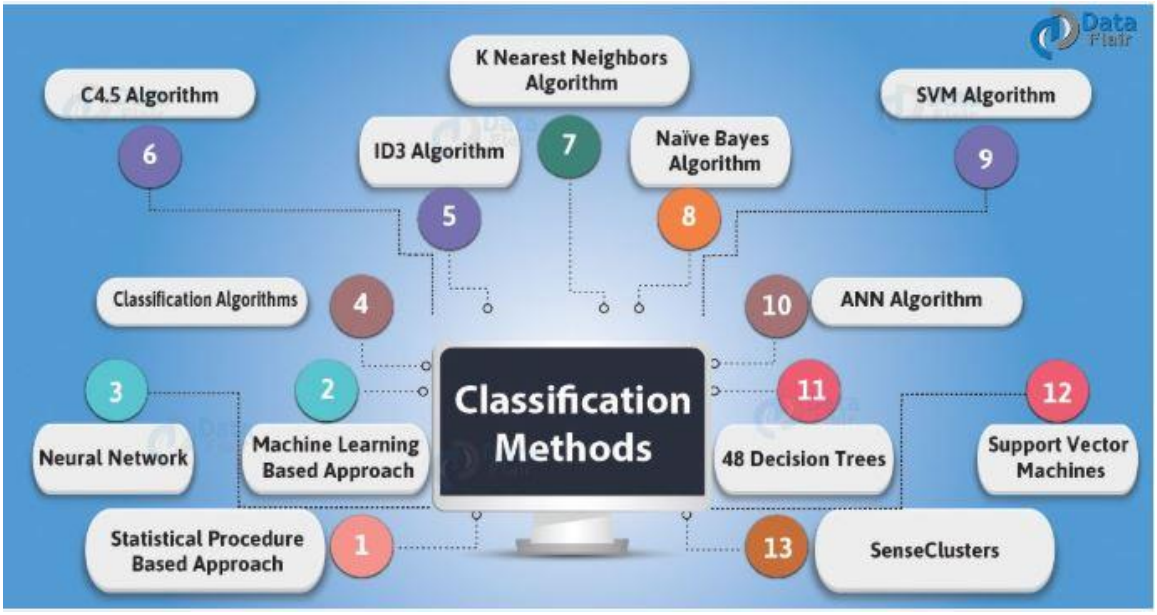


Figure 7: Data mining types[10]

[10]https://data-flair.training/blogs/data-mining-algorithms/

3.2.2 Correlation

Correlation is the mutual inter connection between two or more variables. Correlation matrix is a symmetrical matrix where ij elements in the matrix is equal to the correlation coefficient between variable i and j.

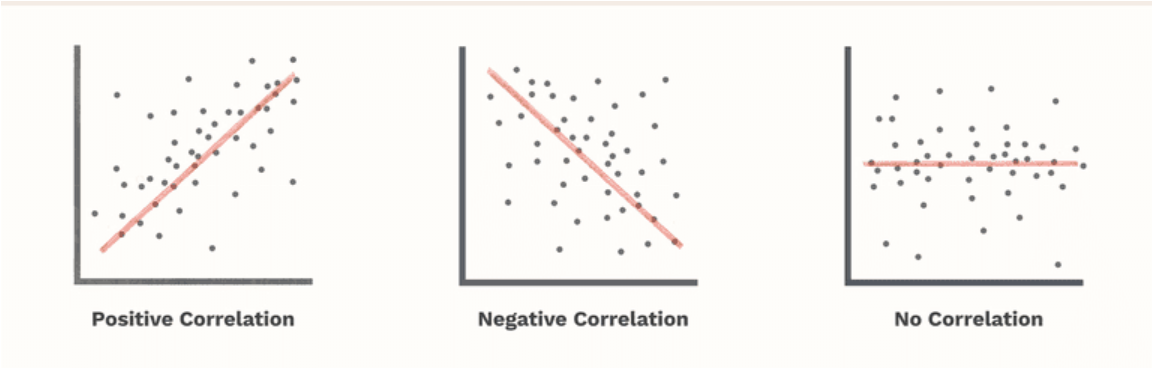


Figure 8: Correlation coefficient

Correlation is a number between +1 and -1 where +1 indicates positive correlation and -1 indicates negative correlation.

Process of correlation includes:

1. Convert numerical values to nominal; because some of the attributes have numerical values.
2. Evaluate correlation matrix: can find the relationship between variables in the data set.

3.3 Scientific Programming tools

Python is a general-purpose programming language. Python can be used to developing both desktop and web applications. Therefore, Google map data can be downloaded using Python.

3.4 Business tool kits

Rapid miner – provides data mining and machine learning procedures including: data loading and transformation (Extract, Transform, Load, a.k.a. ETL), data preprocessing and visualization, modeling, evaluation, and deployment. RapidMiner is written in Java.

3.5 Summary

This chapter presented the technology adapted for solving our research problem. We identified the possible technology that can be used to address the research problem. Next chapter will discuss about the approach.

Chapter 4

Approach for Analysis of Traffic Congestion based on Weather Conditions

4.1 Introduction

Chapter 3 describes how the technology can be adopted to collect traffic and weather data and what are the tools to analyze the collected data. This chapter describes our approach to address the problem of analysis of traffic congestion with the weather conditions. For this purpose, we describe our hypothesis, input, process, output, introduction to the design and top level of the design.

4.2 Hypothesis

We hypothesis that weather condition is effect to the demand of traffic congestion. This hypothesis is inspired by the fact that increase of road accidents is depend on the change of weather.

4.3 Input

Input for the purposed solution includes data extracted using google map and openweather. To download the data, latitude and longitude between origin and the destination need to be provide. Data can be downloaded as a csv file.

4.4 Process

Process of converting inputs to outputs is presented using a diagram in Figure below.

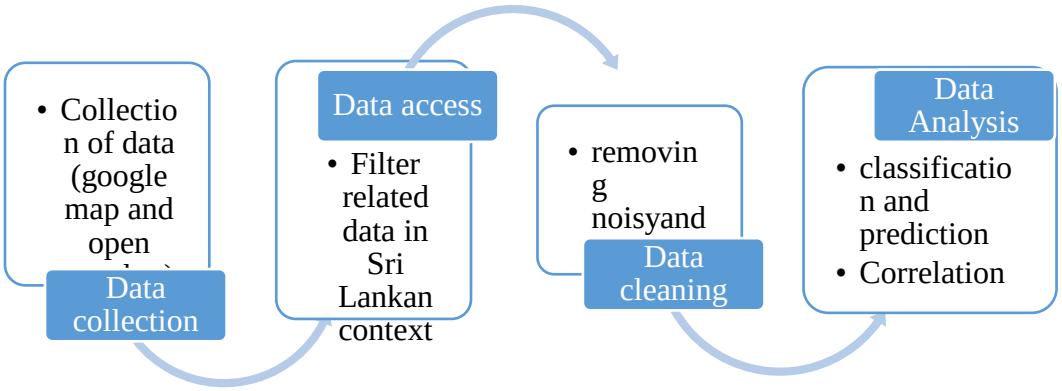


Figure 9: Approach to analyze data

4.5 Output

Visualization of present accurate information about the current situation of the traffic and provide awareness for drivers.

4.6 Introduction to the design

In chapter 4 the approach for the research is discussed with the technologies can be used to collect, clean, analyze and visualize data for analyze traffic data based on weather conditions. Design chapter will discuss architectural design and analysis of the research in a more specific manner. Analysis includes data collection and technology identification and the design will present the components of the proposed system.

4.7 Top level design of the system

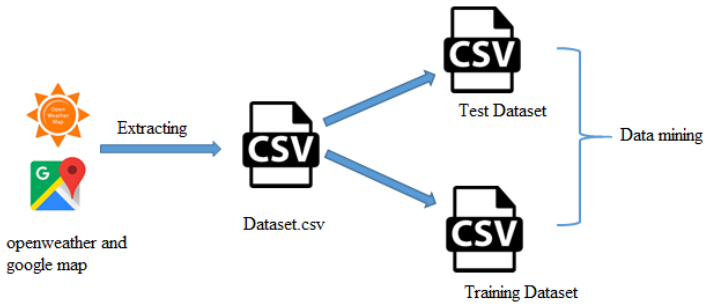


Figure 10: Top level of the system

4.8 Summary

This chapter elaborates hypothesis along with input, process, output as well as top level design of the system. System is proposed to design to deliver an output of meaningful information extracted out of google map and open weather in Sri Lankan context.

Chapter 5

Analysis and Design of Proposed Solution

5.1 Introduction

In the previous chapter, we briefly discussed the approach we have taken to solve the identified problem. This chapter describes the system design which includes three sub systems of data collection, data processing and classification.

5.2 System Design

Major components of research has been designed as data collection, data preprocessing and classification. Data collection is done using an application designed with python programming language (with APIs) and rest of the components are done using Rapid Minor.

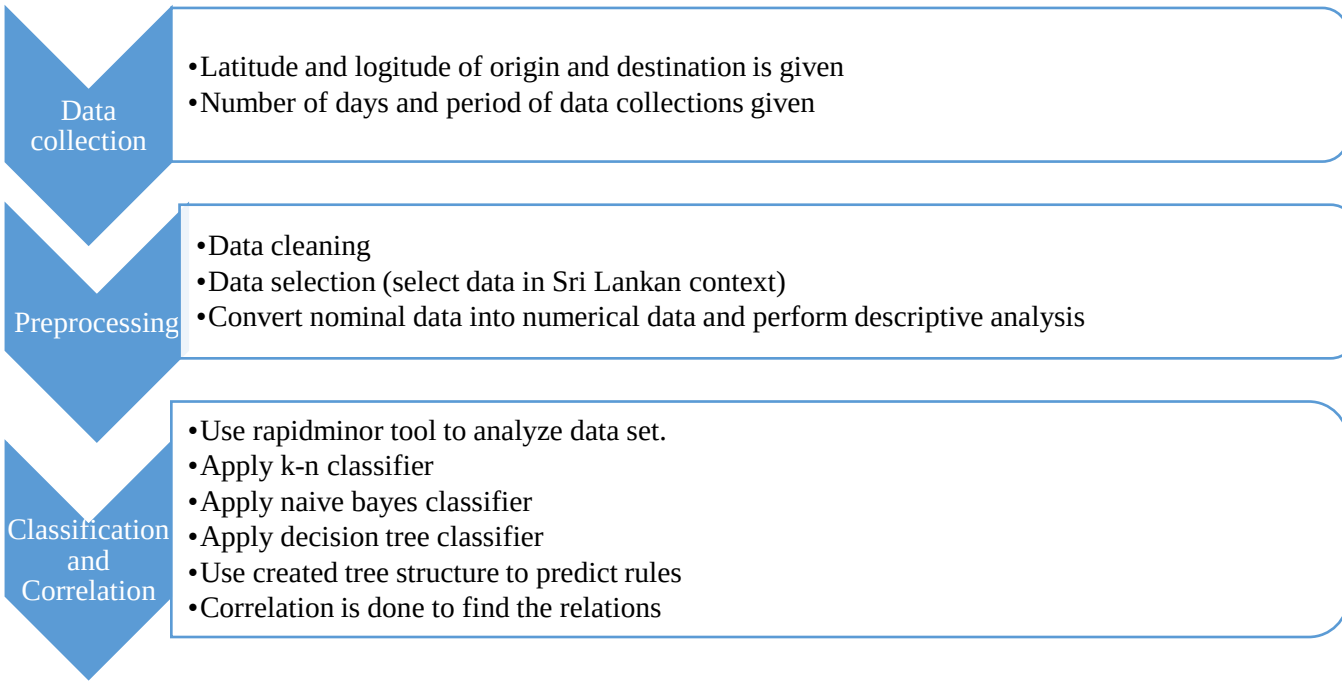


Figure 11: Design of proposed solution

5.3 Data collection and preprocessing

Data collection is done using Google Map and open weather API. These two platforms provide data separately.

5.3.1 Python

In order to collect data a tool is developed using python language. Both google map and open weather data can be access using API keys. Both keys are provided in Jason format. Python code will be implemented as following flow chart:

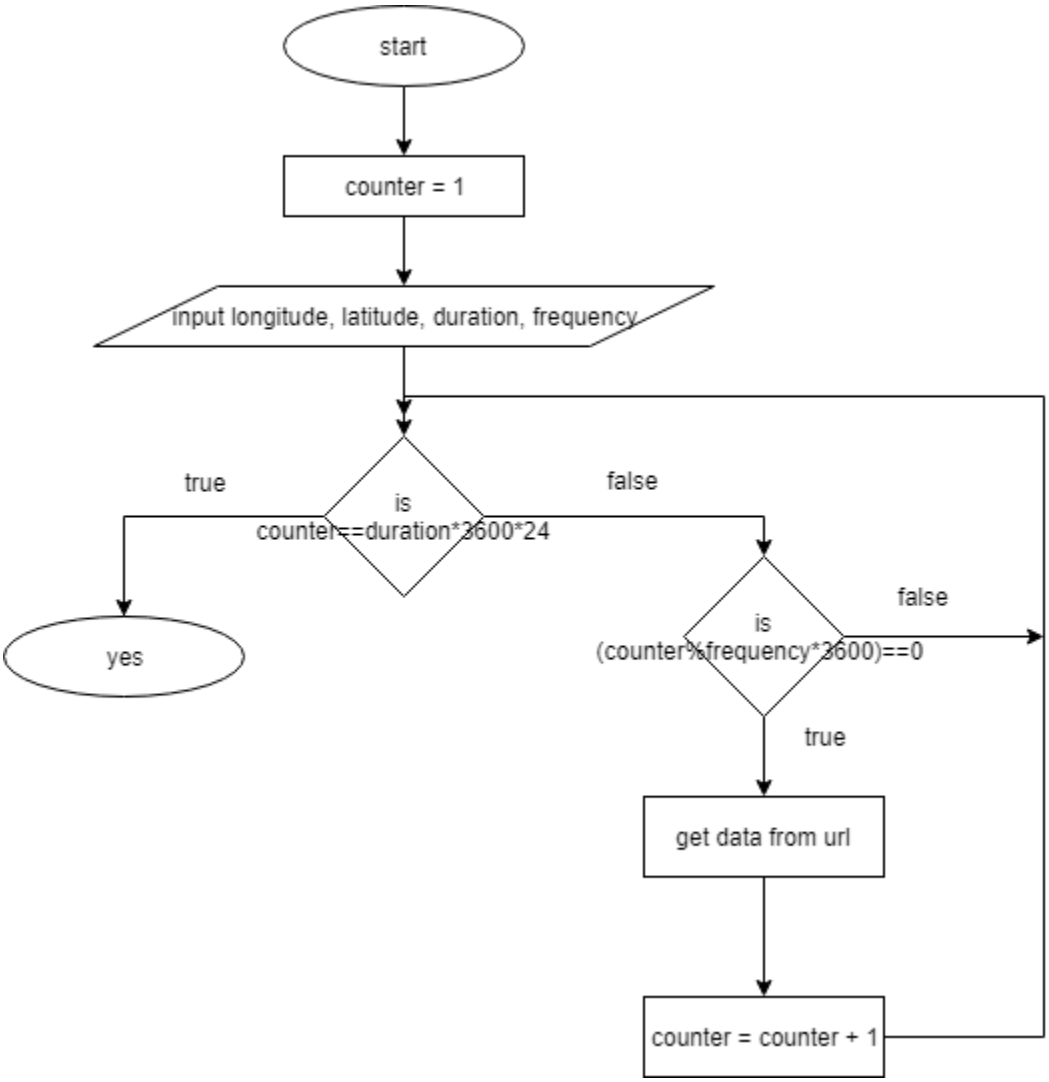


Figure 12: Flowchart of the code which are used to collect data

Collected data has 9 classes including time, duration (average duration), duration with traffic, weather, temp, pressure, humidity, wind-speed and rain. By comparing duration with the traffic duration another column is created using following logic.

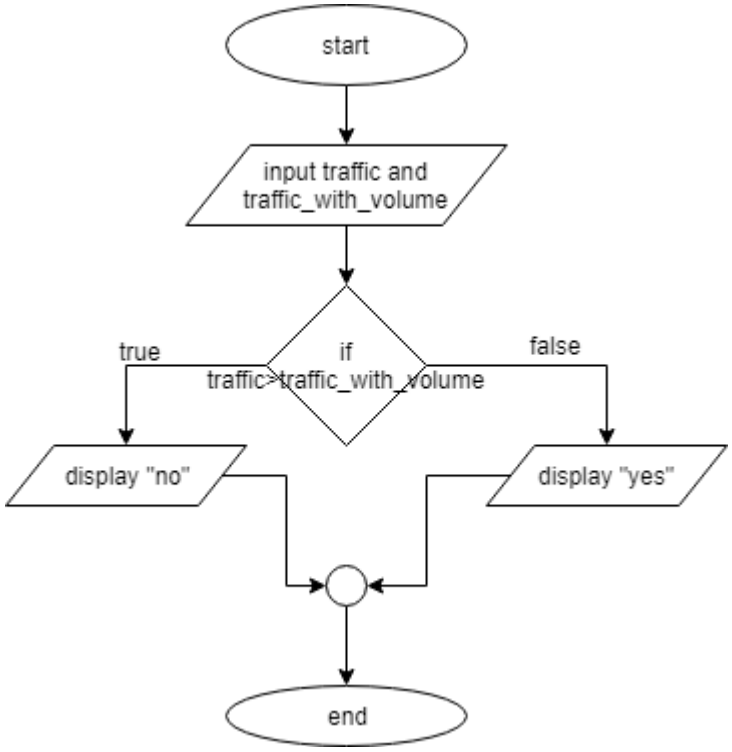


Figure 13: Generate traffic column and traffic with volume into one column

Modified data will be used to do classification and correlation process using RapidMiner

5.4 Data classification and Correlation

RapidMiner tools used to estimate the accuracy of the model using classification. Decision Tree algorithm, Naïve Bayes k-NN algorithms are used here. RapidMiner can be used to do the process of correlation and it provides correlation matrix.

Steps of correlation

1. Convert numerical values to nominal values because Rain attribute and traffic attribute has nominal values.
2. Evaluate the correlation using correlation matrix in RapidMiner

5.5 Summary

In detail analysis on system design is presented here. After evaluating the correlation values, relationship between attributes can present.

Chapter 6

Implementation

6.1 Introduction

In this chapter implementation of the proposed solution will be elaborated. In order to derive situational awareness during extreme weather conditions it is very important to have a real data set.

Flow of traffic is count minutely based (Frequency is 6 minutes) because weather culture of Sri Lanka is changing rapidly. Research have not focused on road traffic congestion during disasters. Research is based on extreme weather conditions such as mist, temperature, rain, humidity and pressure. Open weather provide data in several parameters. We will be using only few selected parameters which is collected during December, 2018-January,2019 by observing weather culture in Sri Lanka.

6.2 Challenges in proposed system implementation

The very first challenge faced was find a suitable data set. Because there are few numbers of real time traffic monitoring system in Sri Lanka. Such as mapmyride, google maps. We have selected google maps by considering popularity in Sri Lanka. Google map provide current traffic data according to the given origin and destination. But historical data is not provided. To fetch the current data small application is created using python programming language.

To capture traffic data, Maharagama to Nugegoda route (distance is about 3.6km) is selected. Because the flow of traffic is consistently changing in that area. We have selected only one route to fetch the data because effect of other parameters (such as road condition) will be same during that period.

Figure 14: Selected area for fetching data

Open weather provide data including date, city id, temperature, wind, humidity, snow, rain and cloud. Only few parameters were selected to fetch because some parameters are inapplicable in Sri Lankan context.

6.3 Downloading data using customized application

To download the data, application is implemented as below figure.

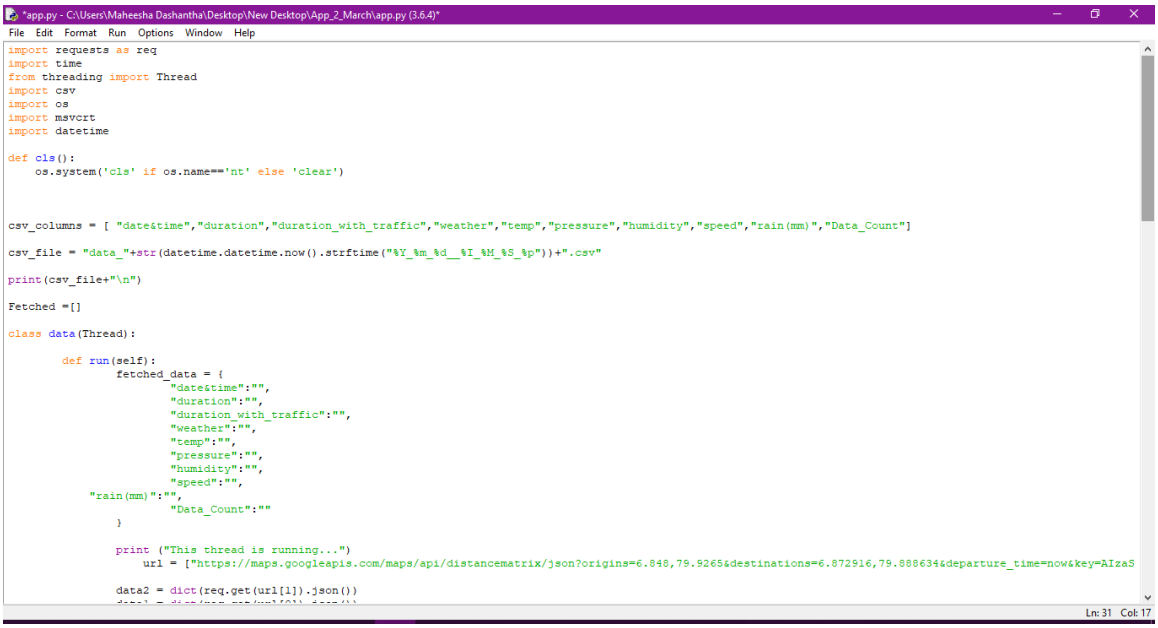


Figure 15: Implemented .py file to download data- Part l

Latitude and Longitude is given to specify origin and destination. Parameters which we want to fetch from the google map and open weather are selected.

Here duration can be defined in number of days. Frequency can be defined in hours.

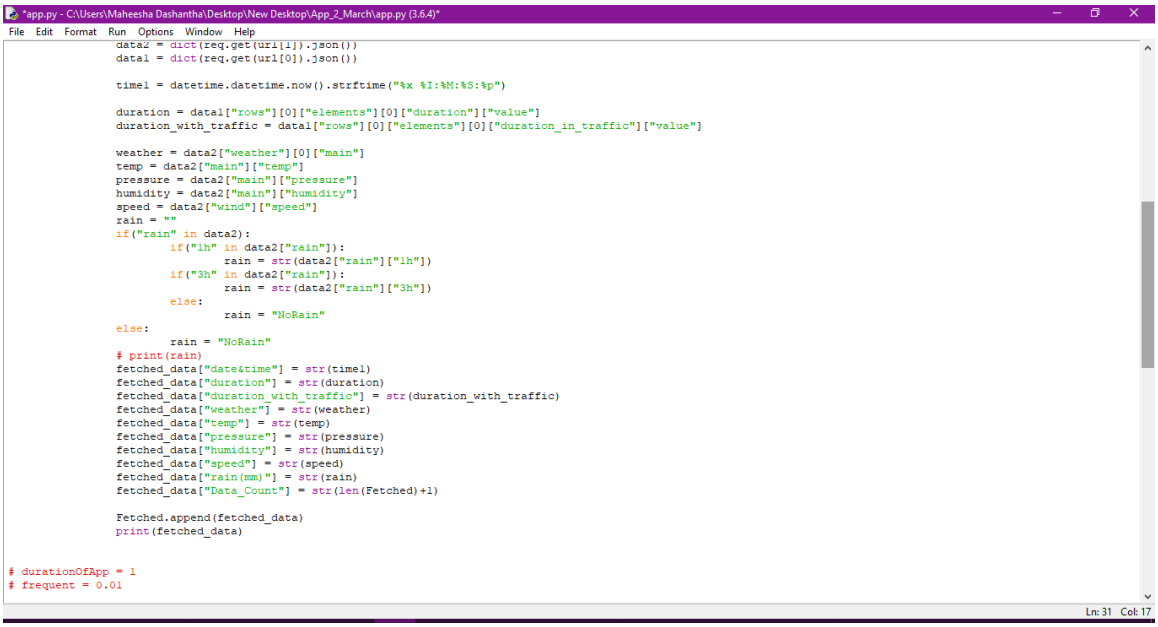


Figure 16: Implemented .py file to import data - Part ll

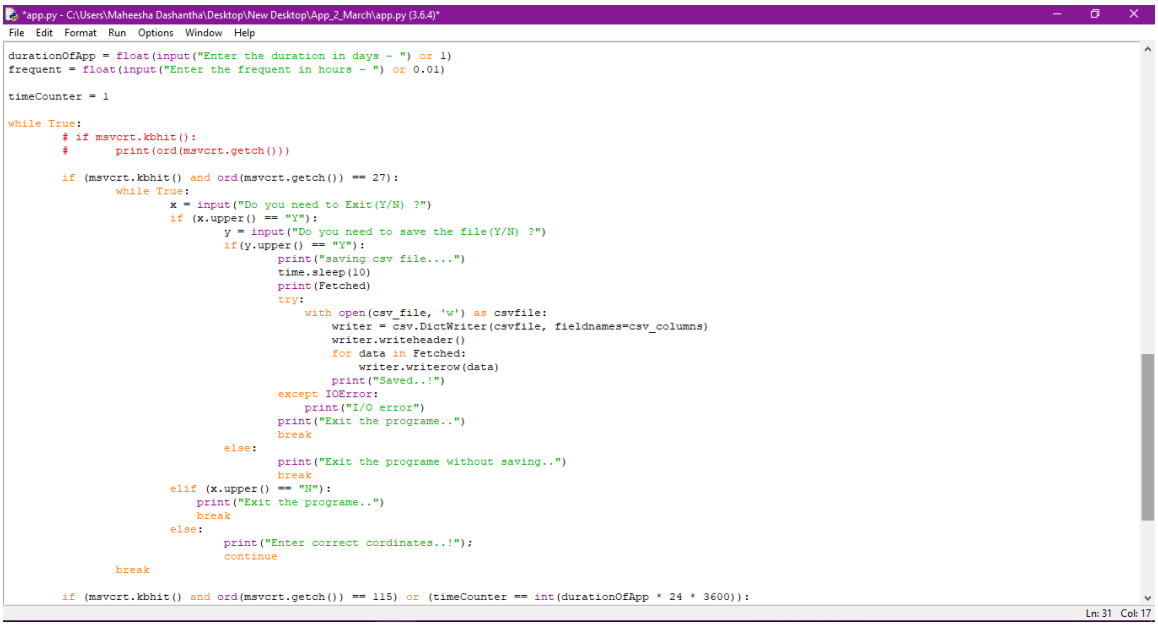


Figure 17:Implemented .py file to fetch and download data - Part lll

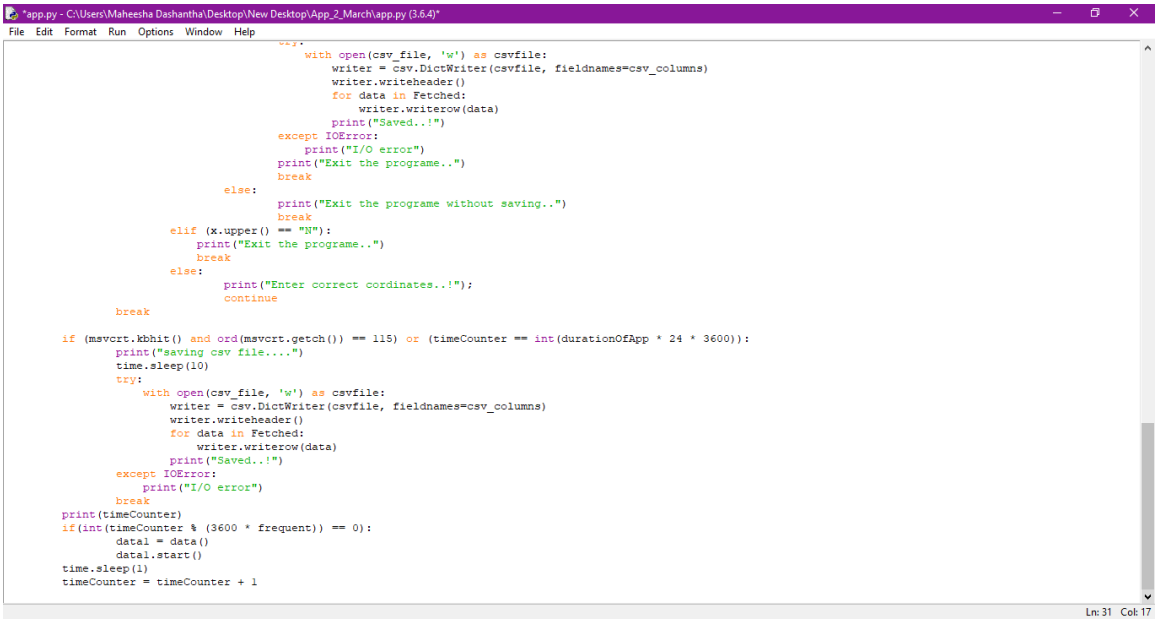


Figure 18: Implemented .py file to fetch data - Part lV

Then run the python file as below (Figure 19) diagram.

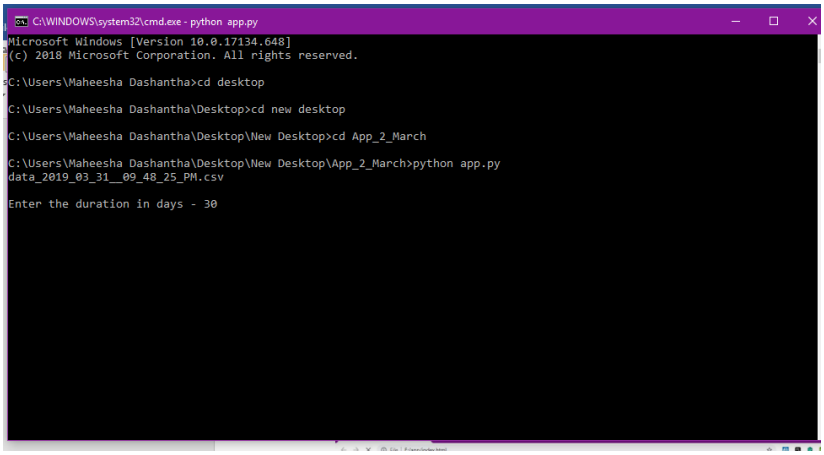


Figure 19: Shows how the application is run

Duration should be defined first. Duration is counted in number of days which the data is fetched. Long duration is not included at once. Because the internet connection must be available to collect the data during the period.

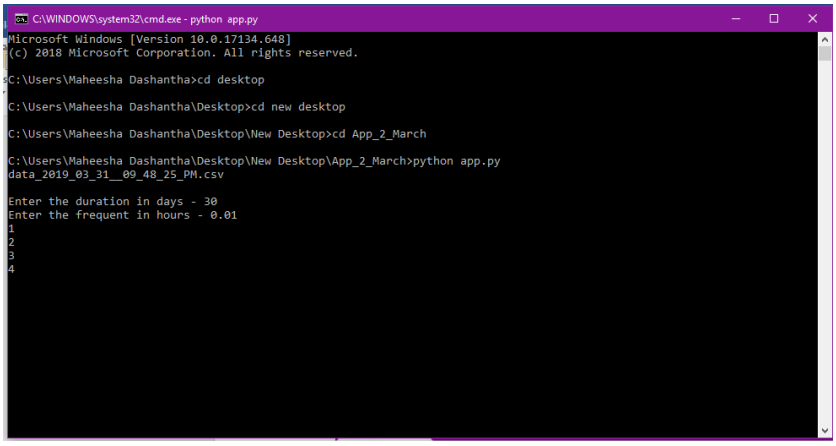


Figure 19: Frequency can define in hours

Figure 22 shows how the data is capturing according to the given frequency. Frequency is defined in hours in the application. Google maps provides current data. Therefore, to maximize the accuracy of data, frequency should be minimized. Therefore, frequency is included as 36 seconds.

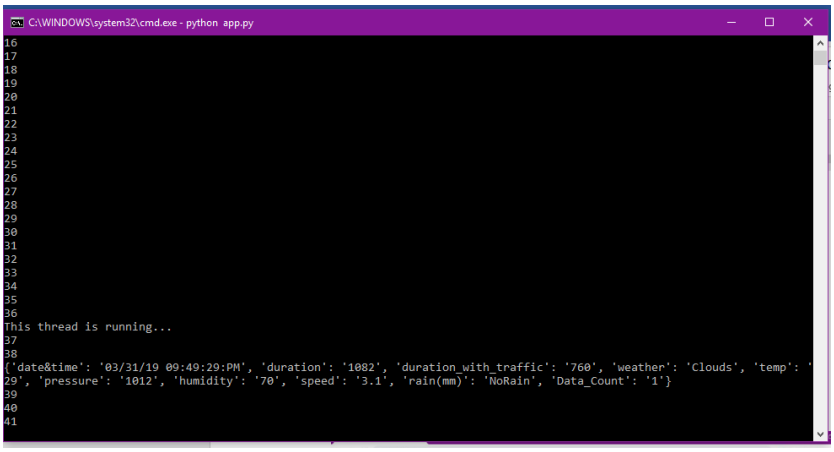


Figure 23: Sample of capturing data

After the given period, csv file is automatically downloaded. Or stop the process by clicking ctrl + s and the csv file automatically downloaded.

All the downloaded data is labeled (Figure 24)

	A	B	C	D	E	F	G	H	I
1	Time	duration	duration v	Weather	Temp	Pressure	Humidity	Wind-Spe	Rain(mm)
2	21:19:32	1084	754	Clouds	26	1013	88	1	NoRain
3	21:20:08	1084	747	Clouds	26	1013	88	1	NoRain
4	21:20:44	1084	732	Clouds	26	1013	88	1	NoRain
5	21:21:20	1084	729	Clouds	26	1013	88	1	NoRain
6	21:21:56	1084	728	Clouds	26	1013	88	1	NoRain
7	21:22:32	1084	733	Clouds	26	1013	88	1	NoRain
8	21:23:08	1084	730	Clouds	26	1013	88	1	NoRain
9	21:23:44	1084	751	Clouds	26	1013	88	1	NoRain
10	21:24:20	1084	750	Clouds	26	1013	88	1	NoRain
11	21:24:56	1084	744	Clouds	26	1013	88	1	NoRain

Figure 24: Sample of downloaded data

6.3 Classification using Rapid Miner

Using classification data set can be organized into categories for its most effective and efficient use. Classification helps to make predictions using the extracted knowledge from existing data. Here classification is done by using decision tree algorithm, naïve bayes algorithm and k-nn algorithm.

6.3.1 Classification by Decision tree model

In order to evaluate the algorithm, first data set is import to Rapid Miner tool

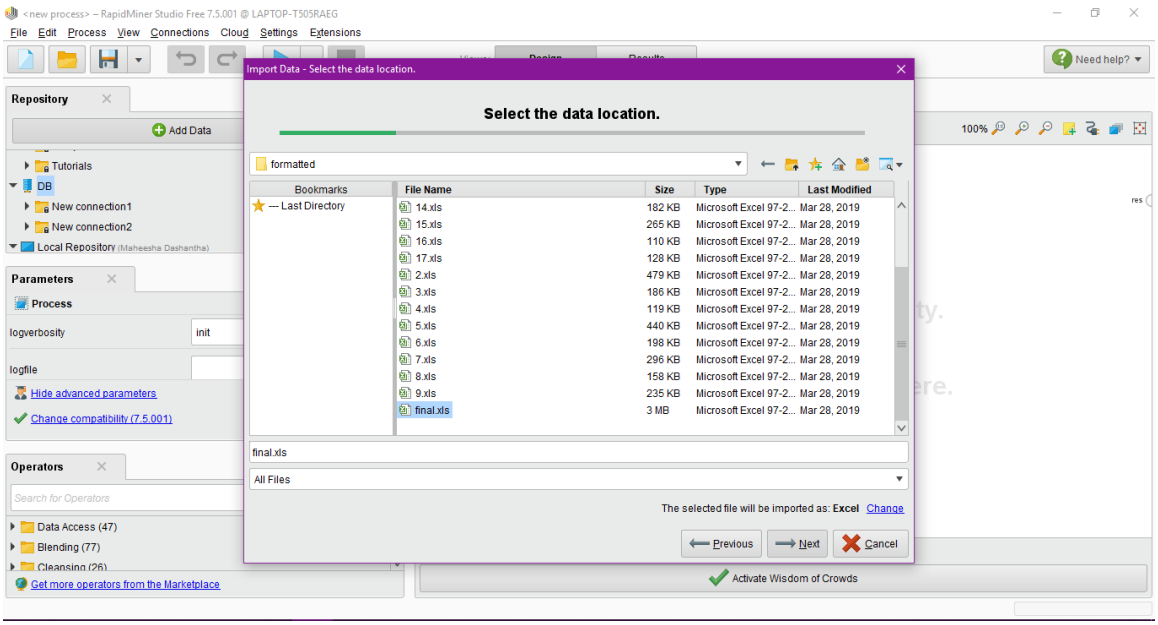


Figure 25: Importing data to rapid miner

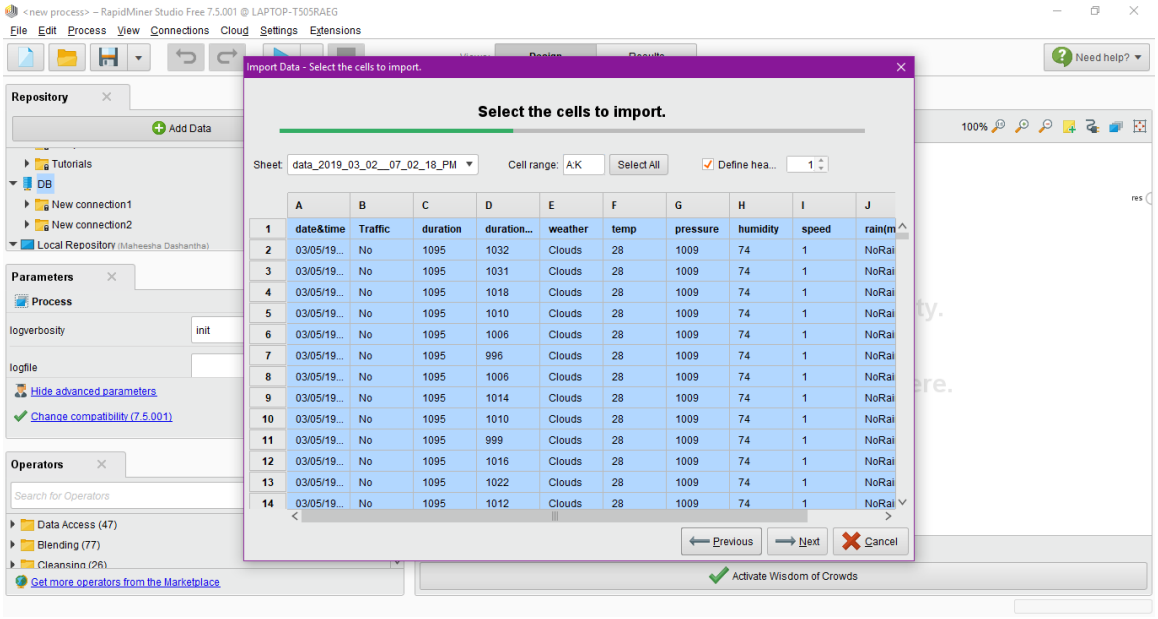


Figure 26: Importing data to rapid miner

To convert data into meaningful way, label column and Temp column (with numerical value) are removed.

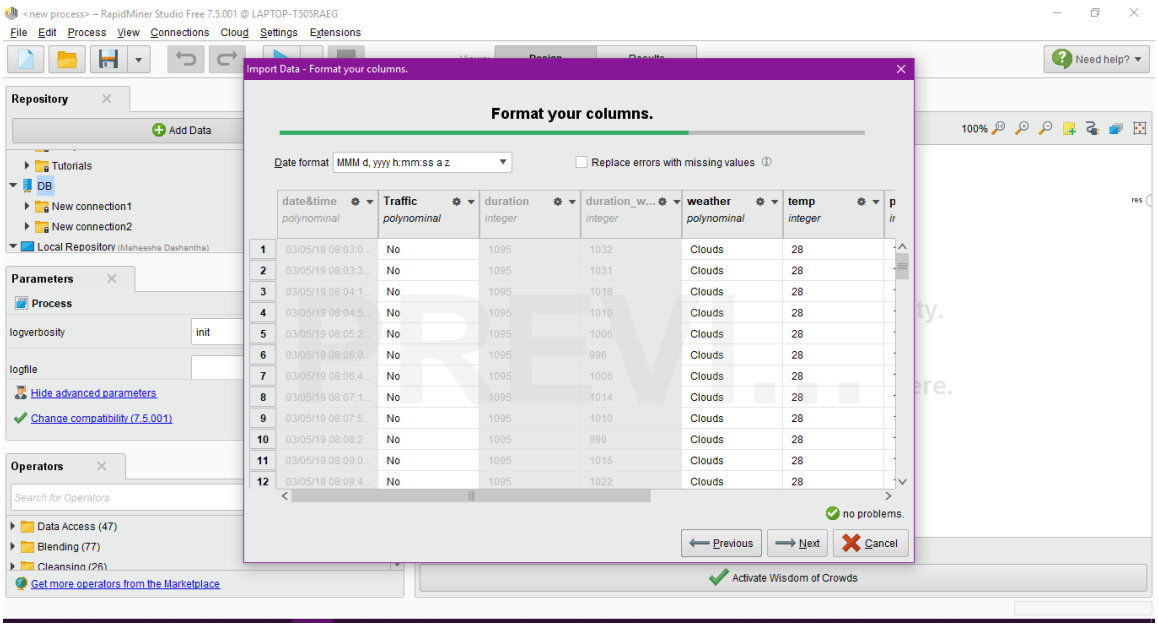


Figure 27: Remove duplicates

Then draw the process to conduct the decision tree classification.

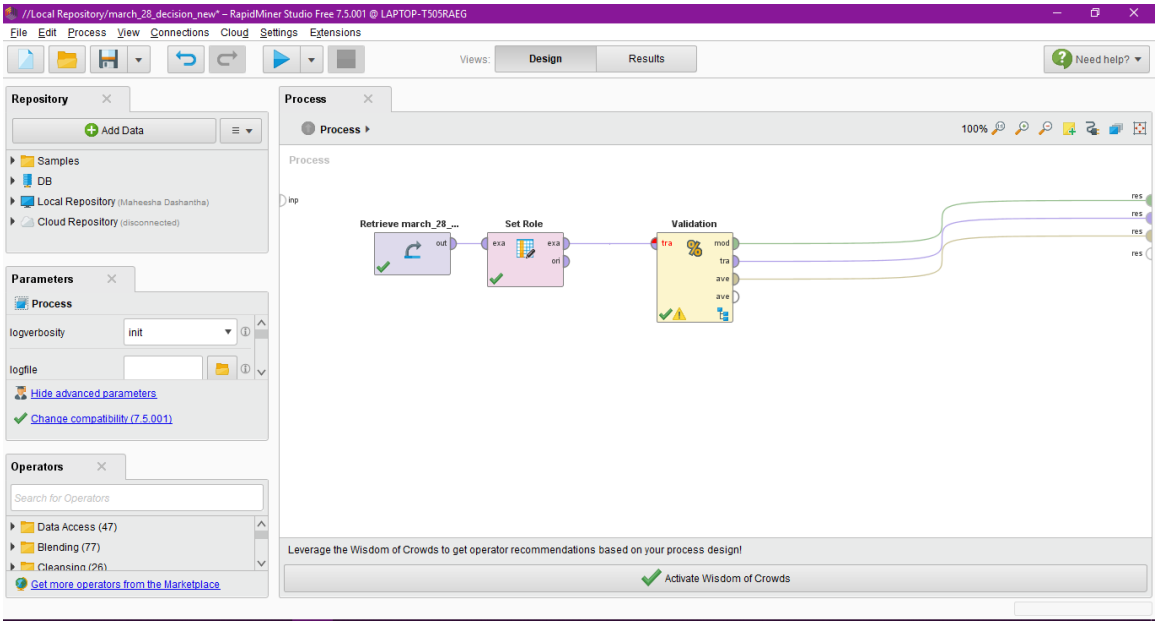


Figure 28: Draw the process – Part l

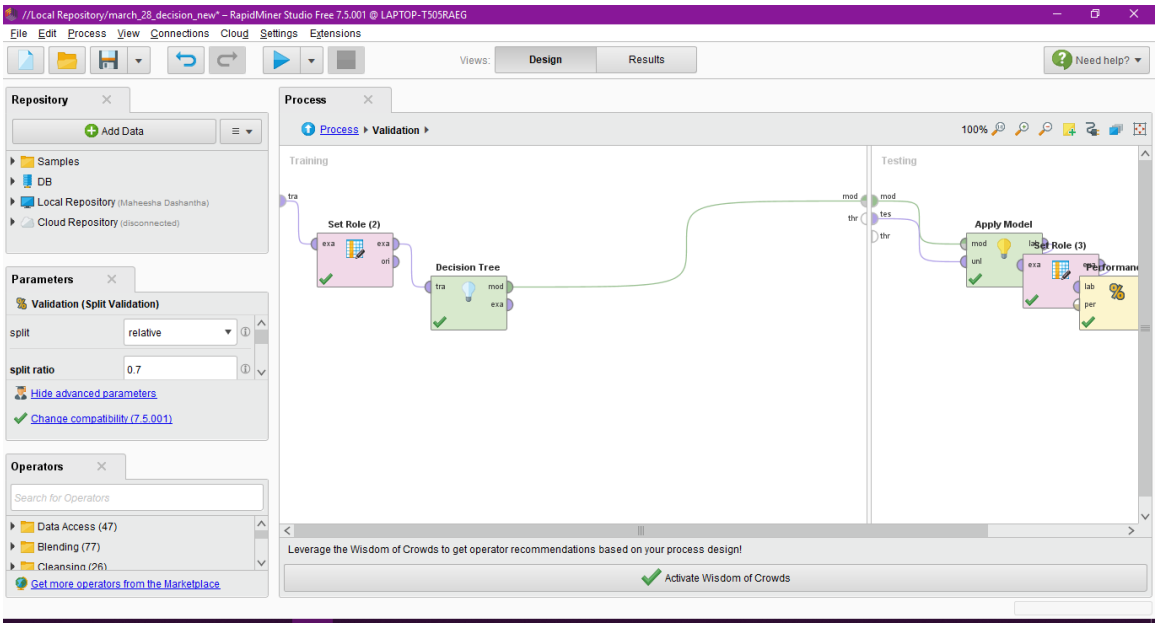


Figure 29: Draw the process -Part ll

Results will be display as a tree and the hypothesis can be predicted using performance. To find the performance of each classifier, classification has done using naïve bayes algorithm and k-nn algorithm.

6.4 Correlation using RapidMiner

Correlation Matrix helps to understand between various variables. To find the co-relation between weather parameters and the traffic, correlation is done using RapidMiner.

There are nominal and numerical values in our dataset. To convert nominal to numerical values and find the correlation, process id drawn as below figure.

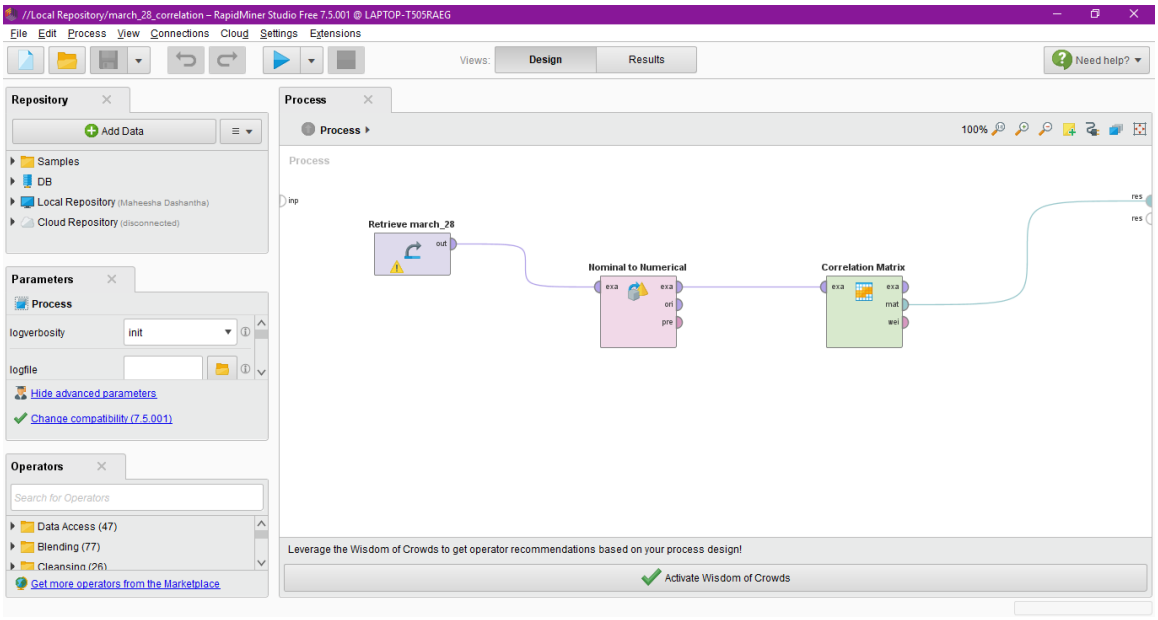


Figure 30: Process of Correlation

Results will be display as a matrix with correlation number.

6.5 Summary

This chapter described how the proposed solution has been implemented from the starting point of the downloading data to classification and correlation. Next chapter describes results of models generated by each module and provide a justification.

Chapter 7

Evaluation

7.1 Introduction

The previous chapter described the details on implementation of the proposed solution. This chapter justifies and evaluates the results.

7.2 Evaluation of different classifiers

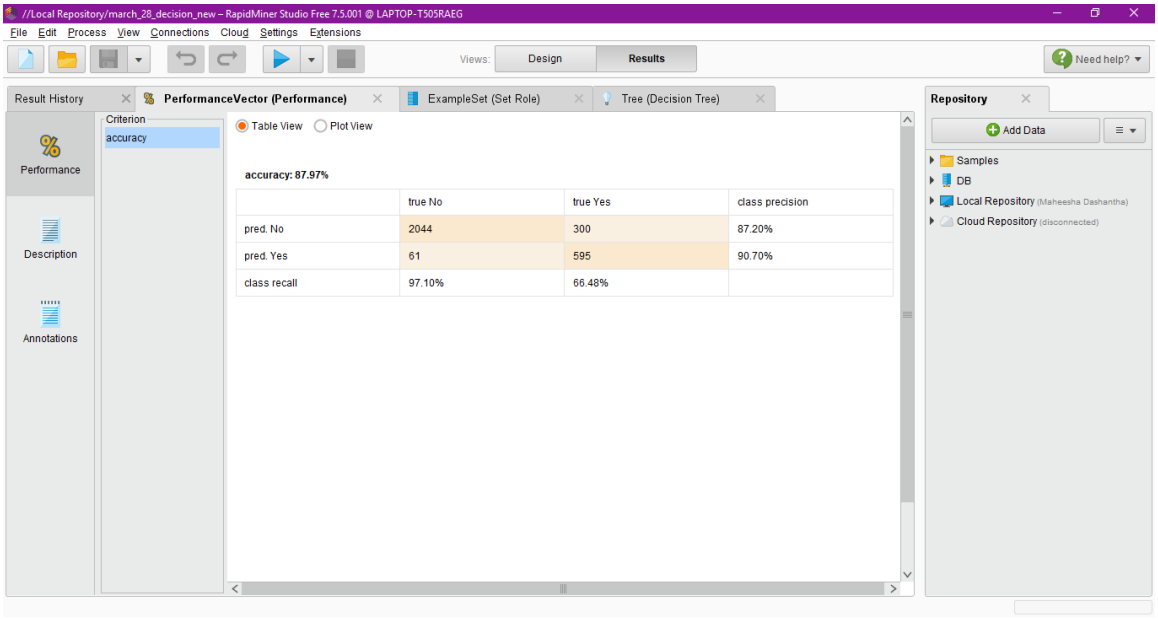


Figure 31: Performance of Decision Tree Algorithm

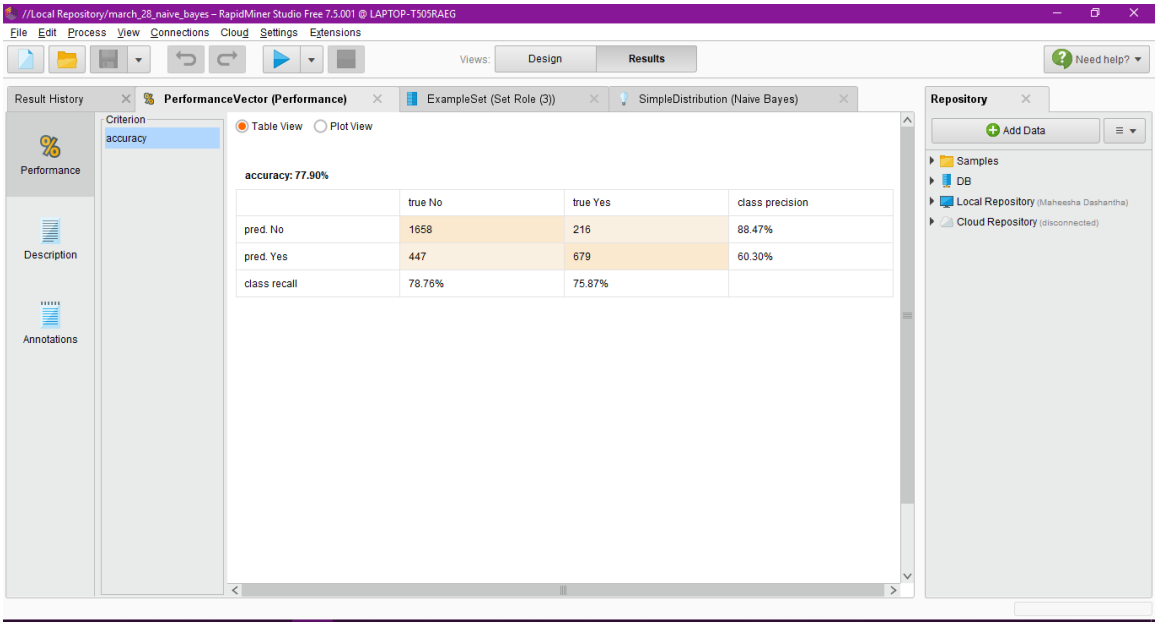


Figure 32: Performance of Naive Bayes Algorithm

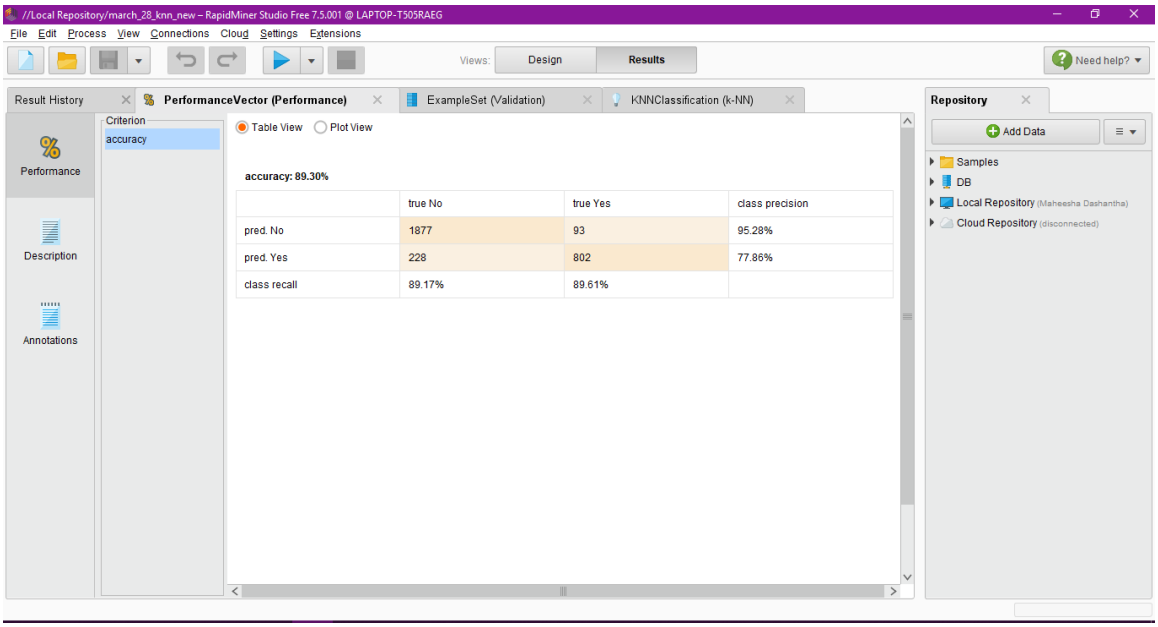
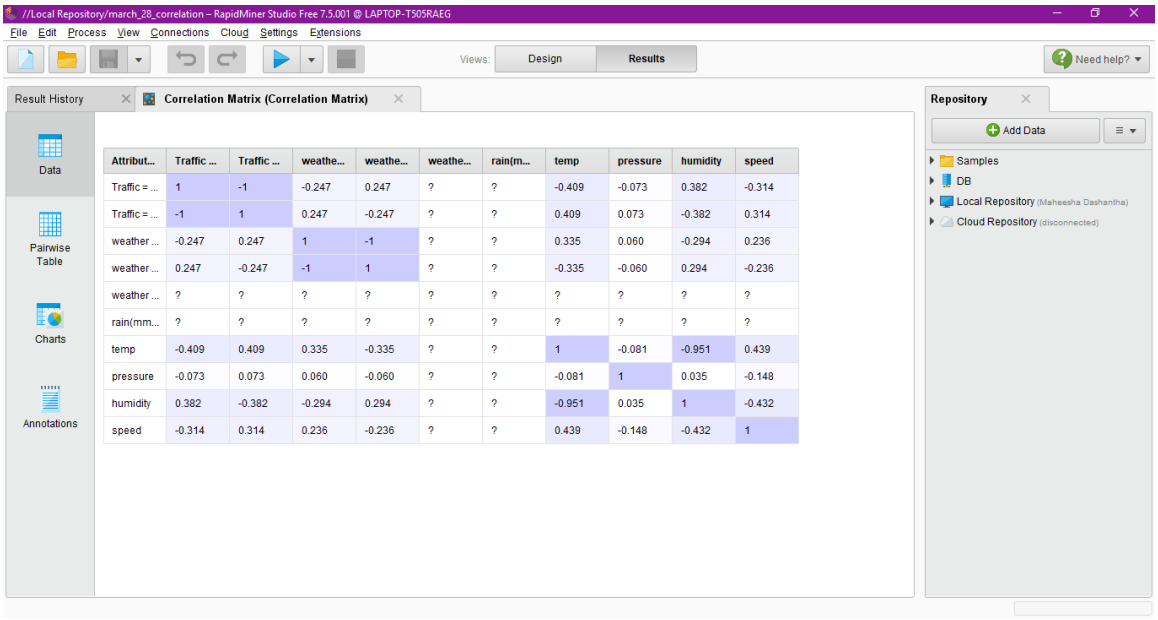


Figure 33: Performance of k-NN Algorithm

7.3 Evaluation of Correlation



Attribut...	Traffic ...	Traffic ...	weathe...	weathe...	weathe...	rain(m...	temp	pressure	humidity	speed
Traffic = ...	1	-1	-0.247	0.247	?	?	-0.409	-0.073	0.382	-0.314
Traffic = ...	-1	1	0.247	-0.247	?	?	0.409	0.073	-0.382	0.314
weather ...	-0.247	0.247	1	-1	?	?	0.335	0.060	-0.294	0.236
weather ...	0.247	-0.247	-1	1	?	?	-0.335	-0.060	0.294	-0.236
weather ...	?	?	?	?	?	?	?	?	?	?
rain(mm...	?	?	?	?	?	?	?	?	?	?
temp	-0.409	0.409	0.335	-0.335	?	?	1	-0.081	-0.951	0.439
pressure	-0.073	0.073	0.060	-0.060	?	?	-0.081	1	0.035	-0.148
humidity	0.382	-0.382	-0.294	0.294	?	?	-0.951	0.035	1	-0.432
speed	-0.314	0.314	0.236	-0.236	?	?	0.439	-0.148	-0.432	1

Figure 34: Correlation Matrix on each attribute

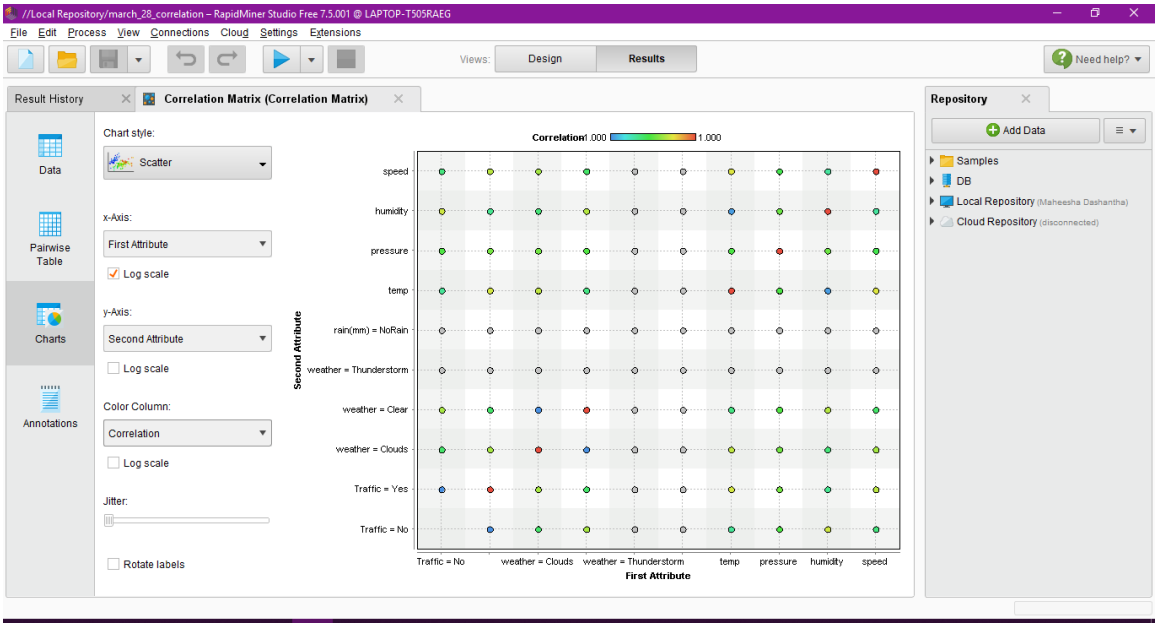


Figure 35: Correlation chart

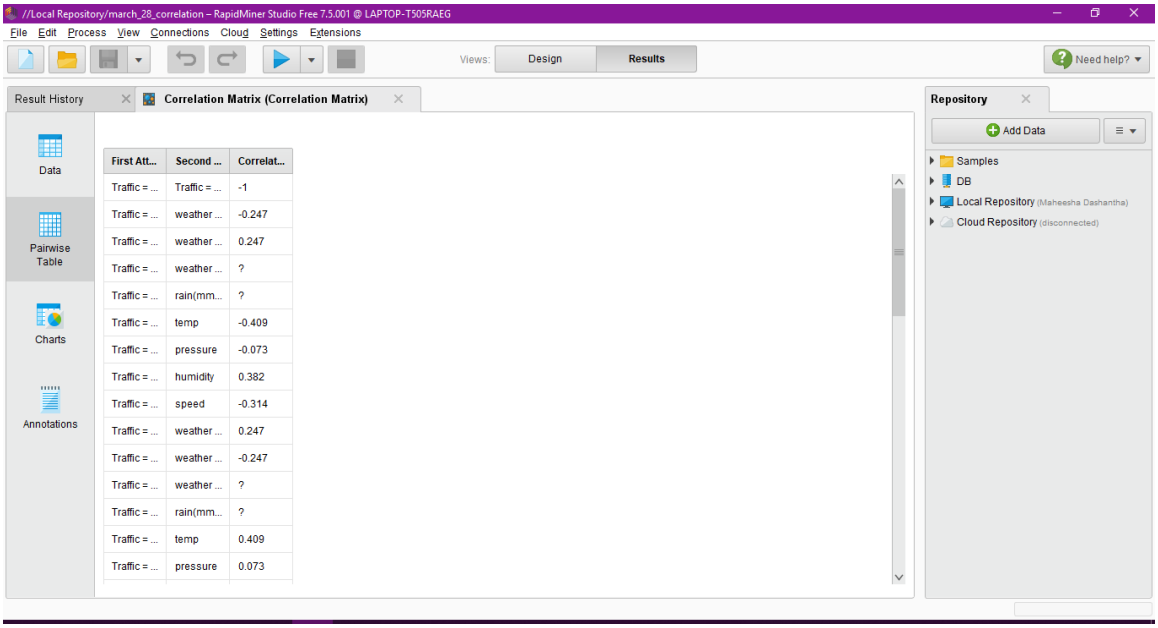


Figure 36: Correlation pairwise table

7.4 Summary

Following table shows the comparison of accuracy of the model using three classification algorithms.

	k-NN algorithm	Naïve Bayes algorithm	Decision tree
Accuracy of the model	89.30%	77.90%	87.97%

Table 4: Accuracy of different classifiers

These three algorithms presented high accuracy, therefore model can be used in future work of this study. According to the pair wise table in correlation matrix, it shows that there is a positive correlation between selected attributes. Therefore, it is predicted that weather conditions are affected to change the traffic in Sri Lanka. Here selected weather parameters are analyzed with the traffic conditions. But during the data collection period, rain data is not fetched. Therefore, rain is mot applied to the analysis. According to the decision tree graph, when the temperature is increased, it affects to the traffic if the pressure and the humidity is high. It is predicted that some weather conditions indirectly affects to the traffic flow while some conditions are directly affecting.

Chapter 8

Discussion

8.1 Introduction

All previous chapters discussed the problem identified and the proposed solution as well as its implementation. This chapter discusses some limitations and future work, improvements which can be proposed for other researchers.

Traffic data and weather data can be taken from various sources. But considering the popularity of Sri Lankan, well known platforms are used here. Developed module can be used and implemented to any platforms. Therefore, adopting to modified module does not take long time. Drivers can get information about the route and select the best route using this module.

It was found that there is a positive correlation of weather conditions with the traffic conditions. But some conditions are not affecting to the traffic flow directly.

8.2 Limitations

Module is used in small area to neglect other factors which are effect for traffic such as road condition.

According to the type of the road, correlation between attributes can be changed. To fetch the data from urls, internet connectivity should be established. There are many weather applications which provide weather conditions. But the key point is we are using open weather map data to collect weather information and google maps data to collect traffic information. All the research depends on the data provided by those platforms.

8.3 Future work

As future extensions of this study we can proposed to expand the data collection in various routes which includes different types of roads using different web application platforms. By analyzing different data collection platforms, we can merge those into one platform and introduce a best solution for Sri Lankans.

8.4 Summary

By making use of available traffic and weather related data and data mining techniques this study aims to find the correlation between those factors. Analyzed information can be used for many researches in this field of study.

References

[1] M. Fernando and S. Fernando, Reducing Traffic Congestion in Colombo Metropolitan Area through Adaptation of Alternative Personal Transportation Methods: Barriers and Long-term Strategies. 2016.

[2] A. Abdelrahman, N. Abu Ali, and H. Hassanein, On the Effect of Traffic and Road Conditions On Drivers' Behavior: A Statistical Analysis. 2018.

[3] A. K. Jägerbrand and J. Sjöbergh, "Effects of weather conditions, light conditions, and road lighting on vehicle speed," SpringerPlus, vol. 5, pp. 505–505, Apr. 2016.

[4] Y. Peng, Y. Jiang, J. Lu, and Y. Zou, "Examining the effect of adverse weather on road transportation using weather and traffic sensors," PLOS ONE, vol. 13, no. 10, p. e0205409, Oct. 2018.

[5] N. V. Hung, L. C. Tran, N. H. Dung, T. M. Hoang, and N. T. Dzung, "A traffic monitoring system for a mixed traffic flow via road estimation and analysis," in 2016 IEEE Sixth International Conference on Communications and Electronics (ICCE), 2016, pp. 375–378.

[6] M. Lewandowski, B. Płaczek, M. Bernas, and P. Szymała, "Road Traffic Monitoring System Based on Mobile Devices and Bluetooth Low Energy Beacons," Wirel. Commun. Mob. Comput., vol. 2018, p. 12, 2018.

[7] S. Priya Biswas, P. Roy, A. Mukherjee, and N. Dey, Intelligent Traffic Monitoring System, vol. 380. 2015.

[8] C. Bian, C. Yuan, W. Kuang, and D. Wu, "Evaluation, Classification, and Influential Factors Analysis of Traffic Congestion in Chinese Cities Using the Online Map Data," Math. Probl. Eng., vol. 2016, p. 10, 2016.

[9] O. Drissi-Kaïtouni and A. Hameda-Benchekroun, "A Dynamic Traffic Assignment Model and a Solution Algorithm," Transp. Sci., vol. 26, no. 2, pp. 119–128, May 1992.

[10] N. Buch, S. A. Velastin, and J. Orwell, "A Review of Computer Vision Techniques for the Analysis of Urban Traffic," IEEE Trans. Intell. Transp. Syst., vol. 12, no. 3, pp. 920–939, Sep. 2011.

[11] R. Cucchiara, M. Piccardi, and P. Mello, "Image analysis and rule-based reasoning for a traffic monitoring system," in Proceedings 199 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems (Cat. No.99TH8383), 1999, pp. 758–763.

[12] C. ChakradharaRao, P. Pushpalatha, and N. AdityaSundar, "GPS Based Vehicle Navigation System Using Google Maps," Almanac, vol. 4, p. 5, 2013.

[13] K. De Zoysa, C. Keppitiyagama, G. P. Seneviratne, and W. Shihan, "A public transport system based sensor network for road surface condition monitoring," in Proceedings of the 2007 workshop on Networked systems for developing regions, 2007, p. 9.

[14] X. Li et al., "WebVRGIS based traffic analysis and visualization system," Adv. Eng. Softw., vol. 93, pp. 1–8, Mar. 2016.

[15] T. H. Maze, M. Agarwal, and G. Burchett, "Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow," Transp. Res. Rec., vol. 1948, no. 1, pp. 170–176, Jan. 2006.

[16] R. Arnott, A. de Palma, and R. Lindsey, “Does providing information to drivers reduce traffic congestion?,” Transp. Res. Part Gen., vol. 25, no. 5, pp. 309–318, Sep. 1991.
[17] X. Yan, X. Li, Y. Liu, and J. Zhao, “Effects of foggy conditions on drivers’ speed control behaviors at different risk levels,” Saf. Sci., vol. 68, pp. 275–287, Oct. 2014.
[18] A. Ghasemzadeh and M. M. Ahmed, “Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric MARS and parametric logistic regression modeling approaches,” Transp. Res. Part C Emerg. Technol., vol. 90, pp. 379–392, May 2018.
[19] M. Cools, E. Moons, and G. Wets, “Assessing the Impact of Weather on Traffic Intensity,” Weather Clim. Soc., vol. 2, no. 1, pp. 60–68, Jan. 2010.
[20] A. K. SOMASUNDARASWARAN, SL ACCIDENT STATISTICS IN SRI LANKA, vol. 30. 2006.
[21] Preprint Version of: Automatic classification of atypical lymphoid B cells using digital blood image processing - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Two-steps-classification-process-First-3-different-classes-of-data-were-obtained_tbl1_281448292 [accessed 13 Feb, 2019]

Appendix A

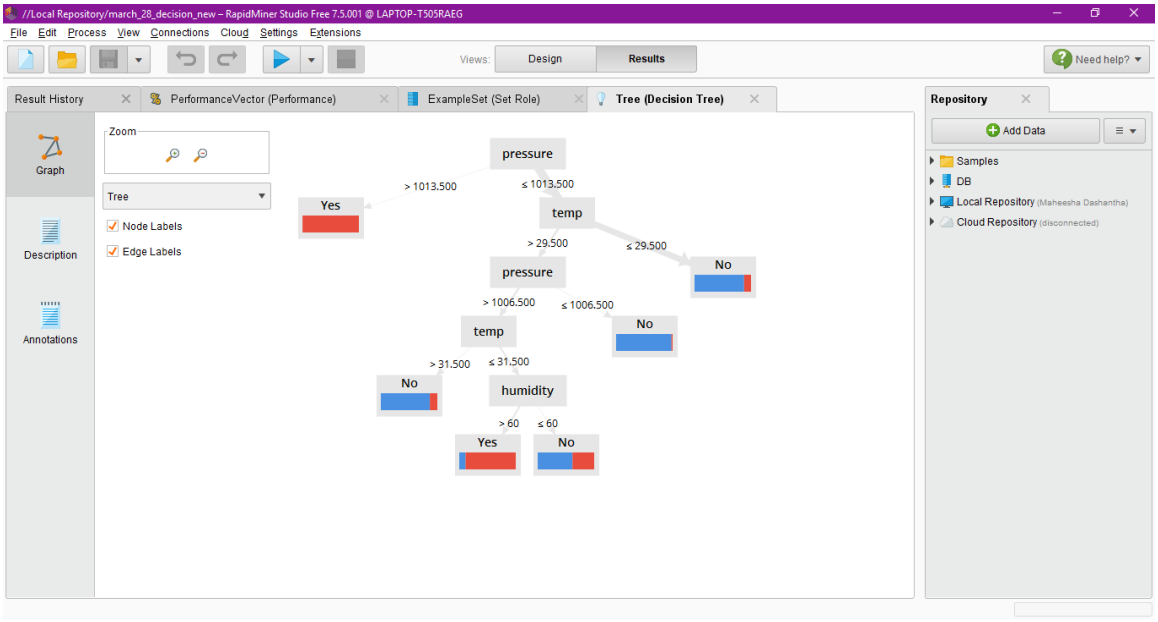


Figure 37: Results of Decision Tree

Appendix B

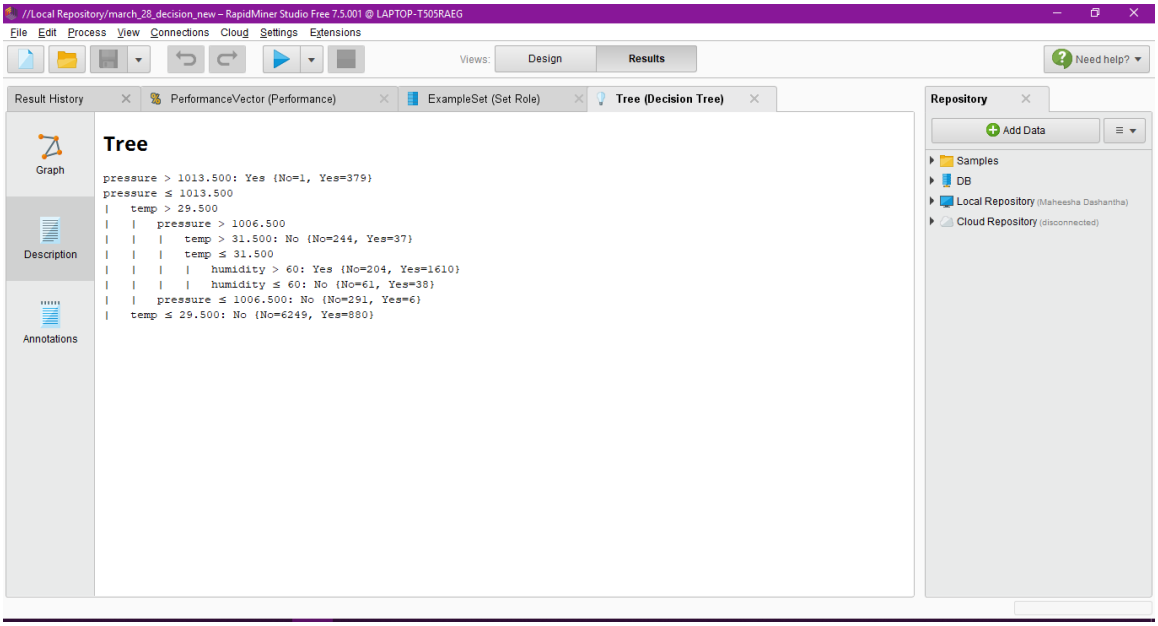


Figure 38: Description of results in decision tree