DATA MINING BASED AUTOMATED SYSTEM FOR TRAFFIC CONGESTION MANAGEMENT USING TRAFFIC PATTERNS

Dissertation submitted in fulfilment of the requirements for the Degree of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

SHILPA THAKUR

11506615

Supervisor

DR. RAMANDEEP SINGH

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

MAY 2017

PAC FORM

TOPIC APPROVAL PERFORMA

School of Computer Science and Engineering

Program : P172::M.Tech. (Computer Science and Engineering) [Full Time]

COURSE CODE : CSE546 **REGULAR/BACKLOG :** Regular **GROUP NUMBER :** CSERGD0236

Supervisor Name : Dr. Ramandeep Singh **UID :** 14105 **Designation :** Associate Professor

Qualification : _____ **Research Experience :** _____

SR.NO.	NAME OF STUDENT	REGISTRATION NO	BATCH	SECTION	CONTACT NUMBER
1	Shilpa Thakur	11506615	2015	K1518	7307918640

SPECIALIZATION AREA : Database Systems **Supervisor Signature:** _____

PROPOSED TOPIC : Data Mining based Automated System for Traffic Congestion Management using Traffic Patterns

Qualitative Assessment of Proposed Topic by PAC		
Sr.No.	**Parameter**	**Rating (out of 10)**
1	Project Novelty: Potential of the project to create new knowledge	7.00
2	Project Feasibility: Project can be timely carried out in-house with low-cost and available resources in the University by the students.	7.00
3	Project Academic Inputs: Project topic is relevant and makes extensive use of academic inputs in UG program and serves as a culminating effort for core study area of the degree program.	7.25
4	Project Supervision: Project supervisor's is technically competent to guide students, resolve any issues, and impart necessary skills.	8.00
5	Social Applicability: Project work intends to solve a practical problem.	7.25
6	Future Scope: Project has potential to become basis of future research work, publication or patent.	7.50

PAC Committee Members		
PAC Member 1 Name: Janpreet Singh	UID: 11266	Recommended (Y/N): Yes
PAC Member 2 Name: Harjeet Kaur	UID: 12427	Recommended (Y/N): Yes
PAC Member 3 Name: Sawal Tandon	UID: 14770	Recommended (Y/N): Yes
PAC Member 4 Name: Raj Karan Singh	UID: 14307	Recommended (Y/N): NA
DAA Nominee Name: Kanwar Preet Singh	UID: 15367	Recommended (Y/N): Yes

Final Topic Approved by PAC: **Data Mining based Automated System for Traffic Congestion Management using Traffic Patterns**

Overall Remarks: Approved

PAC CHAIRPERSON Name: 11011::Dr. Rajeev Sobti **Approval Date:** 26 Oct 2016

ABSTRACT

Traffic congestion is a worldwide issue which results in wastage of time, energy and causes environmental pollution. Large population, poor road infrastructure, and rapidly growing economies lead to severe traffic congestion in many parts of the world. The problem is worsened by increased diversity in vehicle types and poor observance to lane discipline. Identification of congestion is the initial step for selecting an appropriate method to avoid it. To understand congestion in simple way it is classified into different categories in terms of recurring and non-recurring congestion. There are numerous potential congestion management procedures which are based on regularity measures and economic measures. Presently there exist many strategies for congestion management using computerised sensor frameworks which examine traffic density and tackle congestion depending on the congestion problem. This research work proposes an outline of congestion problem and possible ways for the identification of traffic congestion problem. A new automated data mining based solutions to manage traffic congestion, by shifting a number of lanes in the required direction according to the changing traffic pattern, is introduced in this research work. Through simulation based experiment we have found that this system can be used for recurring congestion management by taking proactive action.

DECLARATION STATEMENT

I hereby declare that the research work reported in the dissertation entitled " DATA MINING BASED AUTOMATED SYSTEM FOR TRAFFIC CONGESTION MANAGEMENT USING TRAFFIC PATTERNS" in partial fulfilment of the requirement for the award of Degree for Master of Technology in Computer Science and Engineering at Lovely Professional University, Phagwara, Punjab, is an authentic work carried out under supervision of my research supervisor Dr. Ramandeep Singh. I have not submitted this work elsewhere for any degree or diploma.

I understand that the work presented herewith is in direct compliance with Lovely Professional University's Policy on plagiarism, intellectual property rights, and highest standards of moral and ethical conduct. Therefore, to the best of my knowledge, the content of this dissertation represents authentic and honest research effort conducted, in its entirety, by me. I am fully responsible for the contents of my dissertation work.

Signature of Candidate

SHILPA THAKUR

Reg. No. 11506615

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the M.Tech Dissertation entitled "**DATA MINING BASED AUTOMATED SYSTEM FOR TRAFFIC CONGESTION MANAGEMENT USING TRAFFIC PATTERNS"**, submitted by **Shilpa Thakur** at **Lovely Professional University, Phagwara, India** is a bonafide record of her original work carried out under my supervision. This work has not been submitted elsewhere for any other degree.

Signature of Supervisor

Dr. Ramandeep Singh

Date:

Counter Signed by:

1) **Concerned HOD:**
   HoD's Signature: _____

   HoD Name: _____

   Date: _____
2) **Neutral Examiners:**

   External Examiner

   Signature: _____

   Name: _____

   Affiliation: _____

   Date: _____

   Internal Examiner

   Signature: _____

   Name: _____

   Date: _____

ACKNOWLEDGEMENT

First and foremost praises and thanks to the God, the almighty, for his showers of blessings throughout my research work to complete the research successfully. I would like to express my deep and sincere gratitude to my research supervisor, Dr. Ramandeep Singh, for giving me this opportunity to do research and providing invaluable guidance throughout this research. His dynamism, vision, sincerity and motivation have deeply inspired me. He has taught me the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honour to work and study under his guidance. I am extremely grateful for what he has offered me .I am extremely grateful to my parents for their love, prayers, caring and sacrifices for educating and preparing me for my future. Last but not least I would like to thank all my friends for being with me at each step when I need their support. This thesis would never be successful without your support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

FHWA	**Federal Highway Administration**
ECMT	**European Conference of Ministers of Transport**
ANPR	**Automatic Number Plate Reader**
SDCS	**Speed Detection Camera System**
GIS	**Geographical Information System**
GPS	**Global Positioning System**
GUI	**Graphical User Interface**

CHAPTER 1
INTRODUCTION

Traffic congestion is a major challenge for our day to day life. Traffic congestion is a state where the speed of the vehicles is slower than the actual free stream speed on a road and highway. This is after effect of expanded traffic volume. When traffic is stopped for a specific time interval then congestion situation arises.

This can be solved by implementing a well-planned process that handles the congestion in a smart manner. The first step to solving congestion problem is to understand the basic reasons for congestion in various congestion-tackling phases. It can also be done by developing the relevant congestion indicators for monitoring and utilising the existing infrastructure and creating additional capacity using new technology.

1.1 SURVEY ON TRAFFIC CONGESTION PROBLEM

1.1.1 Definition

The definition of congestion depends on the various measures in which it is measured. Various researchers presented various definitions to describe traffic congestion. There is not any universally accepted definition for traffic congestion. An overview on the various definition of congestion from the research literature is given in Table1.1. These definitions are categorized into following three categories:

- Demand capacity related,
- Delay-travel time related, and
- Cost related.

Table1.1: Definition of Congestion

	Definition
Demand Capacity Related	When traffic rate is greater than the road capacity. Congestion situation arises when there is more number of vehicles at a particular time for special reasons.

	When traffic size on a transportation road goes beyond the capacity of that road.
	Congestion is the impedance vehicles impose on each other, due to the speed-flow relationship, in conditions where the use of a transport system approaches its capacity.
	Congestion may be defined as state of traffic flow on transportation facility characterized by high densities and low speeds, relative to some chosen reference state (with low densities and high speeds).
Delay-Travel Time Related	Congestion is an imbalance between traffic flow and capacity that causes increased travel time, cost and modification of behavior.
	Traffic congestion is travel time or delay in excess of that normally incurred under light or free-flow travel conditions.
	Traffic congestion is a condition of traffic delay (when the flow of traffic is slowed below reasonable speeds) because the number of vehicles trying to use the road exceeds the traffic network capacity to handle them.
	Congestion is the presence of delays along a physical pathway due to presence of other users
	Congestion can defined as the situation when traffic is moving at speeds below the designed capacity of a roadway.
	In the transportation realm, congestion usually relates to an excess of vehicles on a portion of roadway at a particular time resulting in speeds that are slower sometimes much slower than normal or "free flow" speeds.
Cost Related	Traffic congestion refers to the incremental costs resulting from interference among road users.

1.1.2 Various Reasons of Traffic Congestion and Its Impacts

A. The increase in Population:

More cars on the roads are one of the reasons which lead to increase in congestion. As the financial growth of citizen's increase with an increase in population, the demand for personal vehicles for daily commute also rises among citizens. Hence the increase in a number of vehicles and increase in chances of congestion are directly related.

B. Flexible Working Hours:

Congestion quite often happens when individuals are going to or returning from their work place.

C. Bottleneck:

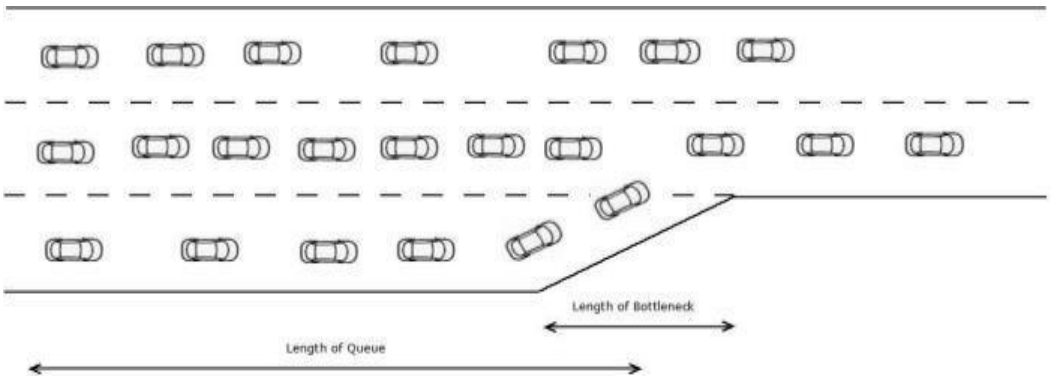


Figure 1.1: Traffic congestion due to bottleneck situation

Traffic bottleneck problems are categorised on the basis of different situations. The main categories of the bottleneck are:

- **Static bottleneck:**

  Problems like poor road alignment, narrowing of road width, on-ramps, off-ramps, and the entrance of tunnel are some of the fixed static bottlenecks which are most familiar.

- **Dynamic bottleneck:**

  The distribution in continuous flow in general traffic because of slow moving vehicles defined as dynamic bottlenecks problem in highway traffic. For example, heavy equipment pulled by a truck or a long vehicle or oversized vehicles. According to the different reports of research traffic bottleneck problem is the main reason for

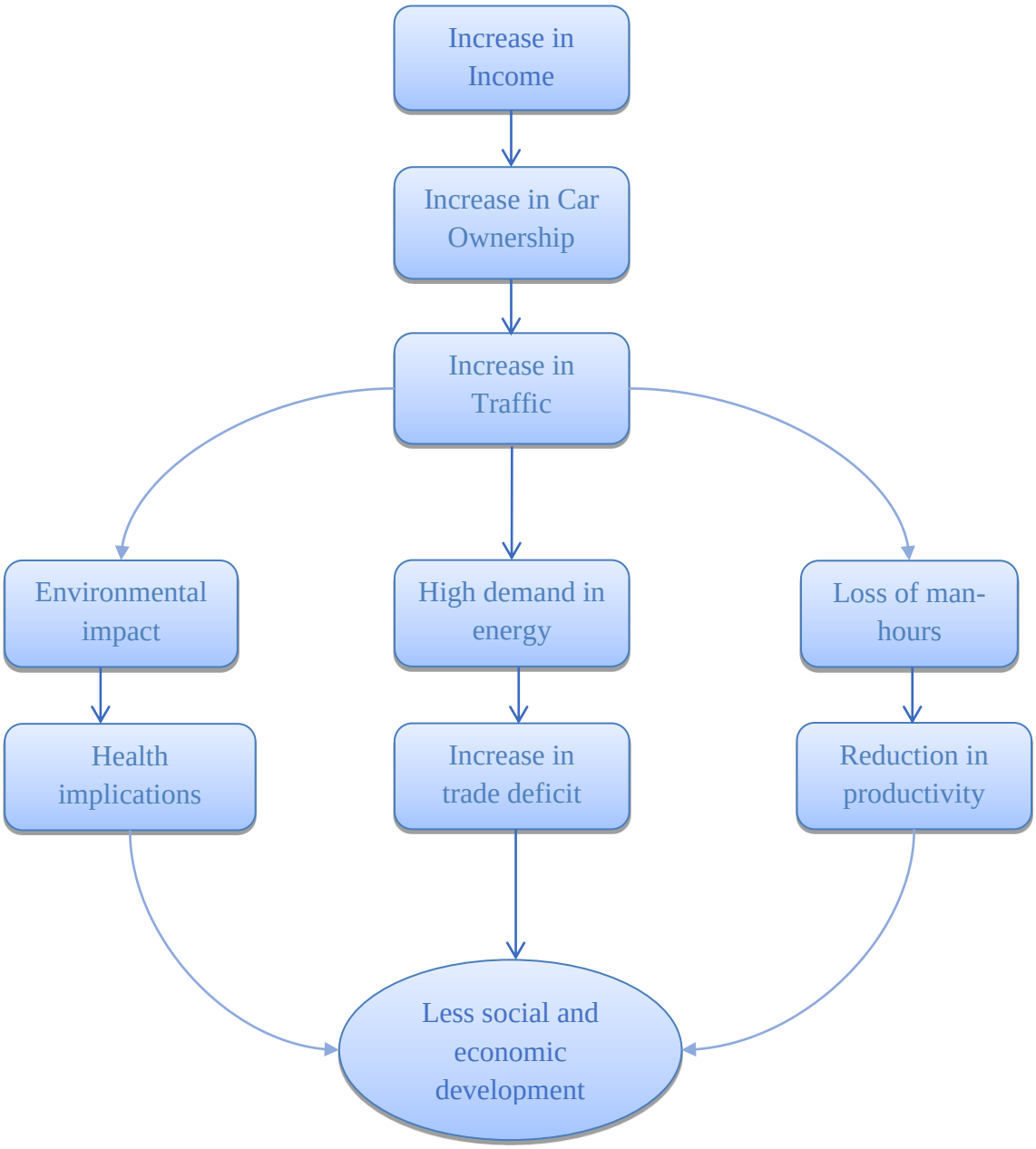


Figure 1.2: Reasons of Congestion and Impacts

the situation of traffic congestion [2]. So, to avoid this situation ones have to start with understanding bottleneck situation and its different causes.

1.1.3 Congestion Classification:

To understand traffic congestion in a complete manner, it can be further divided into the following categories:

A. **Facilities:** The distinctive sorts of roads where the congestion happens, for example, highways, streets, bridges etc.

B. **Recurring vs. Non-recurring:** In the event that congestion happens for ordinarily at the same area for the specific extent, it is known as recurring congestion, else it is categorised as non-recurring congestion. For example, congestion intensity can vary according to holidays and festivals and also due to peak working hours. According to the given report Joe Bared (FHWA) et al. [1]:

- The congested condition repeats once per day for a month because of climate situations and the delay range was between 10 to 20 minutes for every vehicle.
- Congestion repeats minimum for 20 days in a month for road maintenance (work in advancement zone) where speed decrease separates 15 to 20 percent.

1.2 SURVEY IN TRAFFIC CONGESTION MANAGEMENT

There are several potential congestion administration procedures, however, most can be categorised as one of two classes – the one is which offer new limit or free up existing limit and another one offers top, limit or generally manage traffic levels on the new or authorised perimeter[4]. Suggested two related measures are; Regularity measures and Economic measures. Access management comes under regularity measures and parking management and pricing policies come under economic measures [5]. Likewise, gather to guarantee that area use arranging, and the group targets it incorporates is composed with congestion management strategies. In this way, ECMT really depicts three periods of interest side clog administration; regulatory, money related and arrive use. These three phases are also depicted by as successful interest side clogs relief measures [6].

1.2.1 Regulatory measures

Administrative measures fundamentally allude to regulatory measures, strategies or controls that specifically change the explorer’s practices. Administrative measures incorporate; access administration, stopping control, confinements on vehicle use, activity calming and versatile working hours. It has numerous imperatives. Firstly, to the general population, these measures incorporate restriction on vehicle use and these are excessively strict, making it impossible to human self-sufficiency. In this way, these can't

be functional to each country. Furthermore, these measures always seriously influence financial well-creatures by changing normal traffic flows.

A. Access Management

The access arrangement restrains the vehicle access to a few districts for instance recorded focuses or to the specific street joins like ramp metering. In the circumstance of district based limitations, the traffic is block by utilizing the physical barriers as a part of the urban ranges road system (for e.g. the roads systems and one-way avenues are organized in such a way, to the point that it prevents the traffic) or through permit based systems or by traffic bans . To ensure that a delicate single measure should not responsible for handling all main of traffic reduction efforts, the restriction regions should be linked with the set of measures which are supplementary to each other. Some supplementary restrictions can result in quality public transport, pricing and parking controls. The reduction in capacity leads to prevent access to those regions.

B. Parking management

Parking management will be steady in little perspectives in the conditions where there is demand for drop off or get travellers, similar to guardians taking their youngsters from school. Due to the availability of operational flexibility and significant policy, parking control can also be implemented on the time and location basis. To ensure the desired result other measures like pricing or access control can be used for parking management. However, if parking prices are increased within the policies of parking control then profit will also get increased for further congestion reduction.

1.2.2 Economic measures

Economic measures are considered as basic interest side administration measures which can change human practices to maintain a strategic distance from congestion. Different economic measures are found through literature survey, for example, tax assessment, sponsorships, mixed use instrument streets, cordon charges, street tolls, blockage assessing/Tax, range approving the arrangement, electronic street evaluating, cordon charges, associated based pricing framework and so forth.

A. Pricing Policies

Street evaluating or blockage assessing is a fundamental monetary measure for diminishing clog. The expense on street framework charged on all drivers and it can specifically diminish driver's energy to utilise roads. It encases road tolls and congestion evaluating. Favourable position to blockage valuing is that the charges and earnings can help offering advantages for undertaking need transport ventures (e.g. in broad daylight transport, ITS establishment or street development). The public is not concurring with tax assessment and congestion evaluating in light of the fact that it is seen as an extra duty and nobody needs to pay for something that was free sometime recently. Rather than confronting such issues, numerous urban communities respect clog valuing as the last option since it is suitable for the diminishment in blockage issues.

B. Land use policies

Experience from various nations and districts have demonstrated that area use arrangements are essential for altering the activity stream lopsidedness to address clog issue over the long haul. Land utilises rises trip era and expansions local trip designs. In this way, arrive utilizations ought to be re-situated in such a way, to the point that the need and the sum to travel can be minimised. With ideal area use and advancement approaches, the interest for travel can be decreased to the slightest level. The network of urban transportation is urban traffic's backbone. With developing populace and differentiated land-utilize exercises, transportation framework should be overhauled or straightened out. Any slack between developing transportation request and system limit brings about movement blockage, accordingly creating monetary misfortune and natural debasement.

1.3 SURVEY IN VEHICLE DETECTION

A. Gatso

There are more than 4,000 settled Gatso speed cameras as of now being used by police powers and nearby specialists over the UK. The cost of introducing a Gatso speed camera is roughly £20,000, yet cost can go as high as £40,000 if situated in a country area, as the framework requires a 240v power supply. The settled Gatso camera can take

up to 400 pictures. Gatso speed cameras can also differentiate among cars/vans and HGVs distinctly.

B. Truvelo

The forward confronting Truvelo camera framework is intended to take photos of the front of a passing vehicle. This permits the photo taken to demonstrate the driver of the vehicle too. To keep up a key separation from the 'flash', which is given out by a back standing up to Gatso camera, the Truvelo structure uses an infrared glimmer. This creates no obvious "blaze" to the moving toward the driver. These depend on 3 white-lines in the street, painted just before the camera, and are activated by strips in the street, used to pick up the vehicles speed. Truvelo cameras are ending up noticeably progressively more ordinary, areas, for example, Northampton shire utilise predominately just Truvelo cameras.

C. Specs

Specs are typical speed camera systems utilise best in class video structure with Automatic Number Plate Reading (ANPR) electronic advancement. More than two cameras are utilised, every camera is fitted with infra-red illuminators which are settled on gantries over the road, so they can work day or night. Specs speed cameras deal with the vehicles typical speed; since time is running short it takes to drive between the two camera positions. Specs typical speed cameras are fitted either at the roadside or in the central reservation a set partition apart to make a speed controlled zone, or where appropriate, social affairs of cameras can be associated with make a speed controlled system. As vehicles go between the section and leave camera centres their number plates are painstakingly recorded, paying little mind to whether speeding or not. At that point, by ANPR affirmation, the photos on the video of organising number plates are consolidated up, and in light of the fact that each photo passes on a date and time stamp, the PC can then work out your ordinary speed between the cameras. There is no film used for SPECS. SPECS are conventionally used to maintain speed controls on twofold carriageways and motorways. This is in light of the fact that one SPECS gantry can screen up to four ways of action at any one time. SPECS speed camera is right now being used in lasting areas in the accompanying districts: Northampton shire, Nottinghamshire, Greater London, Northern Ireland, Strathclyde, Cornwall, Gloucestershire and South

Yorkshire. Moreover, SPECS normal speed cameras are being utilised as a part of the accompanying impermanent areas: Hertfordshire, Kent, Wiltshire/Berkshire, Devon, Perthshire, Cheshire, Greater Manchester, Staffordshire and West Midlands.

D. Peek

Peek depends upon radar innovation, like Gatso. They are likewise raised confronting due to the 'streak'. Peek traffic cameras inside the United Kingdom are in minority in created ranges.Countries starting at now using Peek Traffic cameras fuse; Leicestershire, Greater London and Berkshire.

1.3.1 Issues involved In-Vehicle Detection

This area talks about the issue definition, considering of conventional radar frameworks. These days radars are to a great degree costly. By and by, their exactness comes up short with respect to a couple of potential applications. In this manner, they should be exchanged via robotized framework keeping in mind the end goal to have enhanced exact yields, more affordable frameworks, and bar human element from the framework. There are two sorts of radars generally used:

A. Highway radars:

These radars are to a great degree costly (around 200,000 –300,000 LE). They process the speed of moving vehicles by mean of sensors getting still picture for vehicles damaging constrained speed.

B. Inner town radars:

These radars are more reasonable (around 70,000 LE) (around 70,000 LE). They compute the speed of moving vehicles by methods for sensors just and it needs an administrator to catch the pictures for the vehicles abusing speed constraint.

Laser jammers and radar jammers, these "stick" the laser by transmitting back a blended banner which can't be handled by speeding camera (these may be unlawful in a few locales). In this manner, the essential for another system to supplant the standard radars on account of the weaknesses has turned into a need.

1.4 DATA MINING

It is the way of extraction of information and patterns from the very large number (or quantity) of data. In the present day, the use of data mining is growing day by day. It is used for decision-making activities. Data mining is also known as data or Knowledge

discovery because it is also used for analysing the data from different views and brief it into useful information. Microsoft Academic Research provides data for data mining keywords which have been grown recently (used in publications, organisations etc.).

A. Association rules:

These are essential if/then statements which are utilised to display relationships between clearly free information in a relational database or some other data store.

B. Machine Learning:

Both data mining and machine learning used same methods. But there is the difference, machine learning focused on prediction, based on known properties, whereas data mining focuses on identification of unknown properties.

C. Support Vector Machine:

It is supervised learning algorithm which analyses data used for classification.

D. Cluster Algorithms:

Clustering is one of the emerging research fields in data mining due to its numerous applications.

Example:K-means,InformationRetrieval,SearchEngine,WebSearch,Indexation,Social Network. Data Mining is the process to find the hidden information as well as a pattern from a bulk amount of data i.e. the data should be coming from different sources such as data warehouse, Data mart etc.

1.4.1 Data Mining Techniques

A. Classification:

The classification is done because of exactly guess the aimed class for all case in the data. One of the examples of this model is it helps to predict the student performance.

- In inclusion, there are two stages in classification. The initial part is the learning process. In this part, the training data or facts are examined by classification algorithm and rules and design are created which are based on learned model or classifier.
- In the second part, the model is used for classification and testing data are used for gaining the accuracy of classification design.

B. Naive Bayes:

It is a conditional probability approach. In which a mention problem case is to be classified, and it can be elected by a vector x=(x1, x2……xn) representing a few *n* features. Using Bayes theorem it is written as

Posterior=prior X likelihood/evidence. Eq.1.1

C. Fuzzy logic:

It is a method to determine the "degree of facts" instead of the general "true or false" (1 or 0).Data mining uses different methods (approaches) and assumption from a broad areas or fields for the knowledge extraction from the huge amount of data But uncertainty is a general phenomenon in data mining problems. Therefore, it is applied to manage the uncertainty in the actual world.

D. Decision Tree:

Decision trees are broadly used in the classification procedure. With the help of this, the model can be predicted and classified. Decision trees show rules, which may follow by individual and used in knowledge structure like a database. Example: if the attendance is not matching the giving criteria then the chances of giving the exam is less.

Decision Tree can be represented as in Figure 1.3: It's like a flowchart. In this rectangular shapes of boxes are called node internal nodes are those nodes that have a child and the leaf node are those who don't have children. The top nodes are the root node. In the given figure outlook is a root node. Humidity and Windy is an internal node.

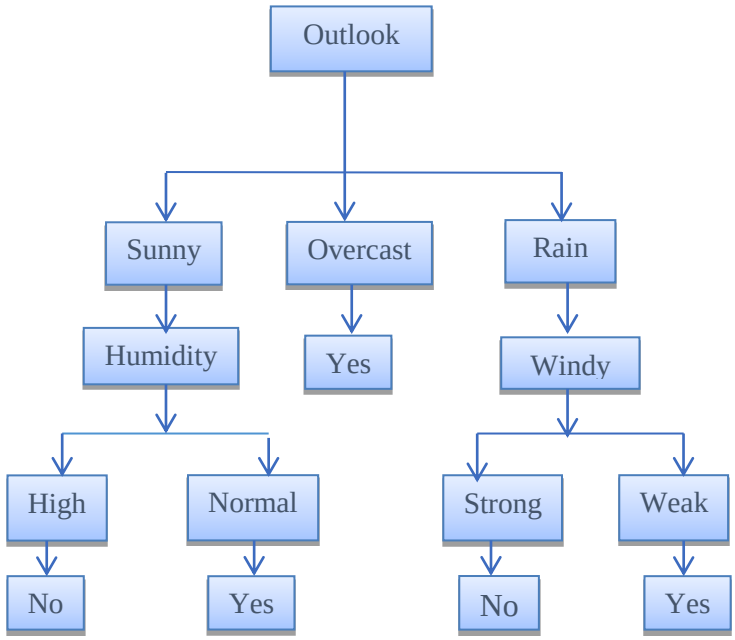


Figure 1.3: Flow Chart of Decision Tree

1.4.2 Clustering

Clustering is a procedure of dividing a group of data (or objects) into a set of significant family, called clusters. Clustering can be used as stand–alone tool to get inside into data distribution or it can be used as preprocessing step for other algorithms.

1.4.3 Association rules

Association rules are essential if/then statements which are utilised to display relationships between clearly free information in a relational database or some other data store.

1.5 RELATED WORK

There are a number of techniques available for traffic monitoring. These Technologies are categorized into intrusive or non-intrusive categories. Here in the following table, there is an overview of these technologies with their strengths and weakness.

Table 1.2: Comparison of various traffic measuring technologies

Technology	Installation	Accuracy	Flexibility	Traffic Parameters
Inductive loop	• Installation involves pavement cut. • Inappropriate establishment diminishes pavement life. • Implementation and repairs require path conclusion.	• Provides the best exactness for considering information contrasted and other normally utilised strategies.	• Insensitive to the harsh climate, for example, rain, haze and snow. • Flexible configuration fulfill expansive assortment of Applications.	• Provides fundamental traffic constraints (e.g., volume, presence, occupancy, rate, gap, and progress).
Magnetometer(two-Axis fluxgate magnetometer)	• Installation involves Pavement cut. • Inappropriate establishment diminishes pavement life. • Implementation and Support involve path Closure.	• Outcomes are less correct than loops to loads of A stream of traffic.	• Unresponsive to nasty climate, for example, snow, rain, and mist.	• Models with little identification zones require numerous units for full path recognition.
Magnetic (induction or search coil magnetometer)	• Installation requires Pavement cut. • Some prototypes are Introduced under highway without requirement for Pavement cuts. However, boring underneath highway Is compulsory.	• Less powerless than loops to anxieties of Traffic.	• Can be used on • Bridge decks where loops are not practical. • Insensitive to severe climate, for example, snow, downpour, and fog.	• Cannot recognise halted objects until unique • Detector formats and • Signal preparing programming is utilised.

Microwave radar	• Microwave radar can be fixed on directly above augmentations, posts or shaft arms transmit indicators that are replicated off automobiles Back to the detector.	• Radar fixed in the forward-looking course • Radar utilised as a part of multilane • Traffic recognition is less accurate.	• Typically, obtuse to severe climate at the moderately short ranges experienced Inactivity Administration applications. .	• Direct estimation of Speed. Continuous wave (CW) • Doppler sensors can't identify halted vehicles
Active infrared (laser radar)	• Installation and support, including intermittent focal point cleaning, require path closure.	• Provide exact data.	• Operation might be influenced By mist at the point when visibility is less than 20 feet (ft.) (6 m) or blowing snow is accessible Information. • Multiple way operations accessible.	• Transmits various bars for exact estimation of Vehicle location, speed, and class.
Passive infrared	• Installation and support, including occasional focal point cleaning, require path closure.	• Provide precise data.	• Passive radar may have lessened automobile affectability in the substantial thick rain, snow and thick haze.	• Multilane passive Sensors measure speed.

Video image processor	• Installation and support, including intermittent focal Point cleaning, require path conclusion when the camera is mounted above the highway.	• A Rich range of information accessible. • Be responsible for wide-territory identification.	• Monitors various Lanes and different discovery lane. • Easy to include and adjust location zones;	• Necessitates 30-to 50-ft • Camera mounting stature for ideal nearness identification and rate estimation.

1.6 OVERVIEW

There are many numbers of traffic management measures presented in this review. Researchers presented various automated techniques to measure traffic congestion and overcome from the situation. So in this chapter we have discussed different mechanised procedures and measures to handle congestion circumstance. So this report contains the review of various automated techniques and measures to tackle congestion situation. Depending on these measures and techniques, we are going to introduce a new data mining based automated traffic management system, which will be helpful in future to reduce congestion problem at bottleneck situations. A review of different techniques and measures is presented in the next chapter. Chapter 2 comprises of different sorts of research papers which are examined for this work. Chapter3 contain objective and sub-objective of our calculation and the hypothesis it is based on. The execution points of interest identified with the calculations additionally depicted in Chapter 3. Chapter 4 contains the Software used to running our framework. Chapter 5 contains the results and output in GUI. At last, Chapter 6 examines recommendations for further work and our decisions. Chapter 7 contains the list of publication.

CHAPTER 2

REVIEW OF LITERATURE

In this research topic, a Data Mining based Automated System with varying traffic patterns for Traffic Congestion Management is introduced. Traffic jam has been one of the challenging issues that most cities are confronting. It is trusted that distinguishing proof of clog is the initial step for selecting suitable change measures. Congestion both in observation and in all actuality impacts the development of individuals. Traffic clog consumes time, strength and causes pollution. To deal with congestion in real situation firstly we have to understand the congestion problem from different views. Various reviews are presented on congestion problem by many researchers.

2.1 Review on Traffic congestion

Aftabuzzaman et al. [7] suggested a framework for building up a measure of public transport congestion relief. This paper gives a meaning of traffic congestion which is classified into three groups; request limit related, travel/delay time related, and cost related. It additionally sets criteria for evaluating a congestion measure. Measures of traffic congestion are classified into four general groups: fundamental measures, proportion measures, the level of administration and records. Congestion measures qualities and shortcomings additionally have been analyzed.

Similarly, Rao et al. [8] also present a review on measuring the urban traffic congestion. There are two major principle categories of congestion problem are introduced in this review, they are (a) micro-level components, which are identified with traffic on the highway and (b) macro-level variables that identify with general demand for the highway. Various author's declaration statements on congestion with respect to different countries have also been included by the author. The author also provided his critic on the different congestion methodology from the past studies, like speed, travel time and suspension, capacity, the level of service, cost related, demand capacity relative and

others. Some data collection methods to measure traffic congestion are also presented. Author classified these techniques into two categories; one probe vehicles furthermore, others that are making utilisation of fixed sensors. Depending on various indices the strength and the weakness of various traffic measures also presented in Table3.

Table 2.1: Advantages and Disadvantages of various traffic measures

Macro Level Indices	**Strength**	**Weakness**
Travel time and delay	• Time-based blockage events help in recognising significant issues by empowering strategy producers to comprehend issues inside the state in the better path and with arrangements having a most prominent effect. • Travel time list has the upside of communicating movement blockage as far as both space and time. • The main concepts of this index are easily understandable.	• Travel time list requires partition of repeating and nonrepeating delay. It can be hard to quantify nonrepeating information. • Travel time measures are not receptive to irregular conditions like climatic occasions, mischances or development movement impedances. • Total deferral could likewise permit transportation experts to gauge how enhancements inside a transportation framework influence a specific passageway or the whole framework.
Volume and LOS	• Many non-technical people are able to understand LOS measure easily. • This is a standard variable in traffic stream examination. • It is generally utilised in light of the fact that this information gathered without much of stretch.	• LOS can't give a consistent scope of estimations of congestion and these strategies can't separate between various levels of clog if once congestion circumstances are emerging. • Lane-Mile Duration Index can't demonstrate the impact of having Diverse expressway capacities on traffic congestion.

Speed	• Congestion length can likewise be figured by evaluating the lessened travel speeds at a specific timeframe. • Therefore, the setting of an edge that is straightforwardly identified with travel velocities is generally proper. ✓ A speed-based limit represents a greater a number of the effects of congestion than would an edge.	• The utilising scope of the rate for the study zone shows the absence of the consensus in the urban regions with respect to the suitable threshold reflecting local conditions**.** • The outcome of this is regarding free stream speed, which is the understanding problem for drivers.

LUO Qingyu[9] explains the main reasons behind the congestion pricing. The main reasons for congestion pricing are the analysis of external cost in a qualitative and quantitative way. The external costs of activity stop up are sorted into four segments: extra travel time costs, natural contamination costs, car accident costs, and fuel usage costs. A careful overview on the quantitative estimations for these four outside expenses is conducted. Likewise, a model case is introduced to demonstrate the practicality of this strategy. Congestion affects country socially as well as economically.

ChaoWang et al.[10] explain main goals of transport arrangement producers to diminish the impacts traffic clog on road. The relationship between traffic clog and highway catastrophes, conversely, is not clear and less considered. The principle point of this examination is to find the traffic blockage impact on the recurrence of highway mishaps. This review is conducted on M25 London orbital motorway which is confined to 70 portions. A specific congestion measurement has been employed by this study. As per aftereffects of this review, traffic blockage has slight or, then again no effect on the occasion of parkway setbacks on the M25 motorway. All other critical components gave outcomes reliable with current reviews. To collect relevant traffic data many techniques also introduced.

Antoniou et al.[11] provide a framework for developing facts gathering innovations and their effect on movement administration applications. To collect traffic data several already existing surveillance technologies are being used. These advances have distinctive specialised attributes and working standards, which characterises the

sorts of information gathered, the exactness of the estimations, levels of improvement, possibility and cost, and framework scope. About present best in the class of traffic modeling is discussed, with regards to utilising rising information hotspots for better arrangements, operations and lively administration of highway systems. There are many traffic data collection techniques already present, which are used by many researchers to collect traffic data. A different data collection technology also helps to learn about nature of congestion. By understanding the nature of congestion one can avoid congestion situation. There are many traffic data collection techniques already present, which are used by many researchers to collect traffic data.

Cesar A. Quiroga[12] presents a review of geographic information system (GIS) technology standards like issues, measures, and examples of the application of to the advancement of congestion administration frameworks. Global positioning system (GPS) innovation is utilised for the accumulation of travel time and speed information. To process the collected data from the field the use of GIS stages and tester user interfaces is also described. Global Positioning System further revised by another researcher.

Michael A.P. et al.[13] describe the nature of traffic congestion by investigation of congestion stages on two equivalent ways in urban Adelaide, South Australia. It gives a general significance of traffic congestion and clarification of parametric measures of the clog. The computation of these measures conducted on the premise of data assembled from the joined Global Positioning System GPS/Geographical Information System GIS structure is likewise characterised. Global Positioning System (GPS) is utilised to gather on-road traffic information from inspecting the vehicle. GPS has been consolidated with the motor administration arrangement of the vehicle to convey time marked information on GPS area and speed, separate secured, increasing speed, fuel ingestion, engine execution, and air contamination discharges on the premise of every second time. In GIS a database administration platform is utilised for the joining, presentation, and examination of the information gathered from GPS. This database conveys a huge source of data for the investigations of different clog measures like travel times/delays, congestion levels, and energy outflows. This information can be used in future work to calculate congestion level. Speed perform index also used to calculate traffic congestion.

Lu Ma et.al[14] analysed traffic congestion in Beijing arterial highway network. To determine the existing traffic systems circumstances of congestion speed performance index was assumed. This review isolated the movement state into four classes; broad blockage, minor clogs, smooth and curiously smooth. This review did Beijing framework qualities examination, highway blockage evaluation, and system congestion estimation, which gives essential information to upcoming traffic organisation. Various researchers provide many ideas to deal with specific congestion issue in a specific way. To mitigate the congestion effect the initial step is to identify or collect congestion data. There are many researchers which provide various innovative techniques for measuring traffic data or dealing with congestion.

Benjamin Coifman [15] introduces an algorithm for coordinating the individual vehicle's relating estimations which are taken earlier by upstream located detector with the measurements taken by highway detector. This paper measures the vehicle length with the help of already existing dual-loop detector system. Detectors are not created to quantify vehicle length so there is an error in these estimations. To beat this issue, the calculation exploits driver's inclinations to hold their positions inside dense detachments. In this algorithm, there is no need to deploy new detector technologies to find out travel time data. Simulation systems are broadly introduced to present congestion issue and to provide a relevant solution to deal with congestion.

Karel A. et al.[16] give an overview of new driver support systems. In congestion circumstances, drivers might be guided by a "Congestion Assistant", a structure that affiliations the elements of a Congestion Warning System and a Stop and Go framework. The review was allocated in a modern driving test system. Driving with the Congestion Assistant while in stop up probably result in a decrement of driver mental workload, however just before blockage starts, i.e. growing essentially observable, the framework could help the workload of the main impetus. Acknowledgement is normally high while encountering the framework, in spite of the fact that not in all regards.

W. Wen [17] proposed a structure for a changing as well as programmed traffic flow light controller framework consolidated by the simulation model, that is made of 6 sub-models hinted in Field to break down traffic flow issue. This model receives and

sends timing of entry and timing of simulating the feat ranged vehicles on highways. Inside the trial, every sub-model tells about ways that have 3 crossing points. Information is gathered from RFID collectors. In this way, the greatest and next best traffic signal intervals displays which were created with the help of simulation. Robotized and changing road traffic signal accomplished framework will administrate extended traffic flow signals for the movement change.

Wei-Hsun Lee[18] proposes knowledge based travel time estimation show, which contains real-time and historical travel time forecasters. These forecasters decide traffic designs from the rough information of area based organizations by information mining techniques and after that by transforming them into travel time expectation rules. In like manner, dynamic weight mix of the two forecasters by meta-standards is proposed to give continuous activity occasions response framework to enhance the precision of the travel time expectation. By evaluating the congestion value, one can deal with congestion circumstance.

W. Wen[19] presents a traffic administration expert system with RFID innovation. This design system comprises with data frame of a private computer, RIFD, separated tag, consolidated electromagnetic sensors, and lastly a speedy server. The framework computes the average of stream and speed information of every road in the town. Then it communicates messages from the full streets in an exceptional locale space. Through a flooding algorithmic program, each server of extremely zone concentrates on trades as well as overhauls facts with all nearby servers in option locality, consequently, servers in various focused area will catch the most current jam communications in an exceptional town. In this way, a dynamic route framework will see the briefest way that maintains a strategic distance from full streets. In future for selecting a more robust response for e-plates, 3 sorts of labels can be compared.

L.C. Davis[20] introduces in his research a new technique to lessen the congestion effect at bottleneck situation. Alleviation of blockage on a two-path expressway with an exit ramp and an entrance ramp is recreated with three-stage movement hypothesis. Propelled travel data the normal speed of vehicles close to the on-ramp jam is further used to involve transports on off-ramp upstream. The presentation of path limitations

notwithstanding preoccupation considerably decreases and basically wipes out the clog, re-establishing stream to about free-stream conditions.

Chris Bachmann[21] analysed estimation methods of seven number of multi-sensor information combination. All methods were executed and paralleled as far as their capacity collect information starting of loop identifiers to test transportations for estimating exactly road traffic volume. The micro-simulation model contains loop detectors and a recently applied traffic observing framework that distinguishes Bluetooth empowered gadgets travelling previous roadside and Bluetooth recipients, taking into consideration an automated technique for test vehicle information collection. To estimate the right traffic speed, all vehicles in the model of micro-simulation are outfitted with gadgets of GPS. Coming towards outcomes, maximum data fusion systems enhance speed on solo sensor methods. The improvement by information combination is affected by the technical system, number of test vehicle, and lastly the road traffic situations.

Xiao Zhang et al.[22] made an HFRBS (Hierarchic Fuzzy Rule-Based System) advanced using Genetic Algorithms for building up a right and durable clog forecast framework utilising a larger than usual assortment of information factors. The arranged framework lessens the size of the concerned info factors and rule base though keeping greater accuracy level. For understanding this, various leveled information structure made from FRBSs is upgraded using Steady-State of GA, that mergers different decision and positioning, parallel direction of enrolment capacities, as well as advancement of the rule base. Video/Image processing is likewise considered as the best method to guide traffic on a specific road.

V. Kastrinaki[23] Video sensors turn out to be especially essential in traffic applications mostly because of their quick reaction, simple establishment, operation and upkeep, and their capacity to screen wide regions. Systems created for these zones must incorporate, among their different assignments, the examination of their static surroundings and the location of static or move impediments inside their space of intrigue. This paper exhibits an audit of picture get ready and examination of instruments used as a piece of these applications and these apparatuses are related to finishing systems, created for particular traffic applications. Preparing techniques are arranged in

light of the inborn association of their information and the space of handling.

2.2 Review on Vehicles Video image processing

There are various automated technoques are already exist for traffic survillience.But installation of these sytems and maintainance are costly.Video monitoring can be used for traffic survillence.There are various kind of cameras are exist which are very effective for this purpose and are less costly than sensor techniques. Video image processing technique is used for counting vehicles and finding speed.There are various researchers who gave various algorithms in video processing to detct carious measures of vehicles.

Jen-Chao Tai[24] demonstrates an image tracking system which is detected accidents at road intersections and is used to for traffic monitoring. In this model, real-time data is collected with by taking dynamic contour framework approach. Estimation of the image looks more consolidated through Kalman filtering methods to track singular transport movement. For introducing picture location of vehicles at an intersection, the researcher proposes a contour format method. Considering a simply planned circuit card, a complete picture tracker is composed then made on behalf of automatic traffic observation. This framework effectively accomplished a timeframe following picture of multi-lane transportation. The intriguing trial comes near region unit talked about showing the effectiveness of the anticipated system.

Diminish Reinartz[25] shows the capacity of utilising picture time arrangement from cameras to drive movement parameters on the guideline of single question estimations. The techniques clarified in the exploration chooses a couple activity parameters for a solitary question and protest bunches incorporate recording and considering different advanced or straightforward ethereal pictures from a high height and with a broad total field of view. The exactness and conceivable outcomes of the systems are researched and presented and also the utilization of an advanced street database for improving the following calculation and for consolidating the results for further activity applications.

Osman Ibrahim[26], demonstrates another Speed Detection Camera System (SDCS)

that is significant as a radar elective. SDCS uses a couple picture dealing with methodology on video stream in online mode - gotten from single camera-or separated mode, which makes SDCS prepared for learning the speed of moving things avoiding the traditional radars' issues. SDCS offers an en-exorbitant other choice to standard radars with a comparative precision or by a long shot superior.SDCS is realized and attempted in many investigations; it exhibited to have achieved a pleasing execution.

H.A.Rahimt et. al[27] presents vehicle speed identification calculation and its application for brilliant reconnaissance framework utilizing PC workstation at the chose street path. Commonplace worldly determination of 30ms to 40ms of ordinary camera has been used. The determination is constrained to the casing rate of camera where the most extreme speed that can be registered relies on upon two components: the planning determination and the dislodging determination in pixels on picture.Exploratory outcomes are exhibited to show the proficiency of the proposed calculation where it has been tried on a genuine video arrangement and the outcomes are broke down.

Chomtip in this research[28] plans to build up the vehicle speed discovery framework utilizing picture preparing procedure. General works are the item headway of a system that requires a video scene, which includes the going with portions: moving vehicle, starting reference point and finishing reference point. The system is expected to recognize the position of the moving vehicle in the scene and the position of the reference centers and figure the speed of each static picture layout from the perceived positions. The vehicle speed recognizable proof from a video diagram system includes six critical sections: 1) Image Acquisition, for social affair a movement of single pictures from the video scene and securing them in the short stockpiling. 2) Image Enhancement, to improve a couple of characteristics of the single picture with a particular ultimate objective to give more exactness and better future execution. 3) Image Segmentation, to play out the vehicle position identification utilizing picture separation. 4) Image Analysis, to break down the position of the reference beginning stage and the reference finishing point, utilizing an edge procedure. 5) Speed Detection, to find out the speed of each vehicle in the single picture edge using the area vehicle position and the reference point positions, and 6) Report, to pass on the data to the end client as intelligible data. The

experimentation has been made so as to evaluate three qualities: 1) Usability, to demonstrate that the framework can decide vehicle speed under the particular conditions laid out. 2) Performance, and 3) Effectiveness. The outcomes demonstrate that the framework works with most astounding execution at determination 320x240.It takes around 70 seconds to recognize a moving vehicle in a video scene.

Jianping Wu[29] composed and actualized with Visual C++ a product video stream handling framework to understand the continuously programmed vehicle speed recognition and vehicle number checking. In perspective of geometric optics, at first acquaint an adjusted system with absolute plot orchestrates in picture territory into certifiable space. The second part is fixated on the vehicle acknowledgement in electronic picture traces in the video stream. This framework tried on a tablet controlled by an Intel Centrino-2 (1.2GHZ) CPU and 1GB RAM. The preparing rate is 18.0 edges/second. Investigate moreover shows the system can in the meantime perform vehicle speed distinguishing proof and vehicle number looking at changed ways.

WangJing-Zhong[30] research based on PC vision guideline and advanced picture differential recognising innovation, this shows an ongoing strategy for vehicle's speed discovery. In the first place, align parameters of the camera, catch advanced flag of pictures and take pre-handling on pictures. At that point, select foundation picture and set straight projection of speed. At last, recognise the snapshot of passing the direct projection of speed, and process the speed of the vehicle.

Fumio Yamazaki et. al[31]A new question-based strategy is produced to separate the moving vehicles and in this manner recognise their velocities from two back to back computerised elevated pictures naturally. A few parameters of Grey values and sizes are inspected to characterise the items in the picture. The vehicles and their related shadows can be separated by expelling huge questions, for example, streets. To recognise the speed, initially, the vehicles and shadows are separated from the two pictures. The relating vehicles from these pictures are connected in light of the request, estimate, and their separation inside an edge. At last, utilising the separation between the comparing vehicles and the time slack between the two pictures, the moving velocity can be distinguished. Their test demonstrates a promising aftereffect of distinguishing the

moving vehicles' rates. Promote improvement will utilise the proposed technique for a couple of Quick Bird panchromatic and multi unearthly pictures, which are at a coarser spatial determination.

Chomtip Pornpanomchai et. al[32] plans to build up the vehicle discovery and checking framework utilising picture handling. General works are programming improvement of a framework that requires a video stream and catches to a video outline. They comprise of the accompanying segments: foundation street with no moving vehicle and the casing with moving vehicles. The framework is intended to discover the separation which is the moving vehicles and locate the quantity of moving vehicles from the video outline. The vehicle identification and tallying framework comprises of four noteworthy segments: 1) Image Acquisition, 2) Image Analysis, 3) Object location and Counting, and 4) Display Result The test has been directed keeping in mind the end goal to get to the accompanying qualities: 1) Usability, to demonstrate that the framework can decide vehicle recognition and check under the particular condition layout. 2) Efficiency, to demonstrate that the framework can work with high exactness.

Muzaffar Djalalov1 et. al[33], propose a vehicle recognition and following calculation. Center filtering is used for establishment extraction which is later subtracted from the development plots for question acknowledgement. Morphological directors are used for blob extraction. In this way, disagree disclosure is master utilising focus separating and morphological shutting operation. Kalman isolating is used for challenge taking after which uses the territory of blobs. One of the upsides of this structure is that each vehicle in the packaging is described into different shading boxes. We present preliminary research occurs that will finally provoke the unmistakable evidence of the took after vehicle.

A. Jović et. al[34] depicts the attributes of six most utilised free programming devices for general information mining that are accessible today: RapidMiner, R, Weka, KNIME, Orange, and scikit-learn. The goal is to outfit the captivated examiner with all the fundamental points of interest and drawbacks as to the usage of a particular device. An examination of the finished tallies covering all locales of information mining (organise, descend into sin, gathering, backup guidelines, consolidate affirmation,

assessment criteria, perception, and so on.) is given. Besides, the gadgets' support for the more advanced and focused research subjects (huge data, data streams, content mining, et cetera.) is laid out, where material. The instruments are in like way emerged from respect with the social occasion bolster, in the context of the accessible sources. This multidimensional audit as the ace paper on data mining devices underlines the way of RapidMiner, Weka, and KNIME stages moreover perceive the enormous movements made in substitute gadgets.

Tina R. Patil et. al[35] demonstrates that order is a vital information mining procedure with expansive applications to group the different sorts of information utilised as a part of about each field of our life. The order is used to mastermind the thing according to the components of the thing with respect to the predefined set of classes. This paper put a light on execution evaluation in light of the privilege and off-kilter events of data game plan using Naïve Bayes and J48 portrayal figuring. Guiltless Bayes computation relies on upon probability and j48 count relies on upon decision tree. The paper sets out to make a relative evaluation of classifiers NAIVE BAYES and J48 with respect to bank dataset to extend bona fide positive rate and farthest point false positive rate of defaulters as opposed to finishing the simply higher course of action accuracy using WEKA instrument. The examinations occur showed up in this paper are about request precision, affectability and specificity. The results in the paper on this dataset also show that the viability and precision of j48 are better than that of Naïve Bayes.

R.Arora et .al[36] This paper represents the classification algorithms analysis on various datasets by using WEKA tool. Data mining is one of the most important coming research areas to resolve various problems. In data mining field classification is one of the main problem areas. In this paper, two classification algorithms J48 and multilayer perceptron alias MLP of the Weka interface are used. These are used for testing various datasets. The testing results of both algorithms are analysed to choose the best algorithm depending on the requirements of the datasets. Datasets for this testing are selected from UCI Machine Learning Repository. On comparing the performance results of both algorithms the performance of Multilayer Perceptron is found better in many cases. So Multilayer Perceptron can be used for the classification in data mining for better results.

Rohit et.al[37] This paper provides review on new data mining techniques and predicting the student's academic performance which is the main part of education management. There are various factors which can affect a student performance. The main purpose of this paper is to discover the accuracy of data mining techniques. In this research paper 100 under graduate student's record is collected. By using WEKA tool Naive Bayse algorithm and Decision tree algorithm are evaluated to predict the student performance. On this dataset Decision tree algorithms accuracy is more than the Naive bayes algorithm. This work will be helpful for Educational Institution for predicting the student's performance. For our research work prediction of congestion can also be predicted from these algorithms.

Tribhuvan et.al[38] In this paper a review is presented on accident and congestion of traffic in USA. According to this papers survey almost 90 people on average lose their lives every day and more than 250 are injured every hour. Accidents also results in traffic congestion, which affects the life of people in big cities. In this research the big data of accidents and Big Data analytics tools are used to obtain important insights to improve road safety and decrease traffic accidents.H2O and WEKA mining tools are used in this research. Feature selection method is used to find the most important interpreters. Class imbalance issue is resolved by using various quality measures. This analysis will be helpful for decision makers and experts for developing new traffic rules and policies by which road accidents can be prevented, and roadway safety also increased.

2.3 Summary

In this review various researchers' views are presented on traffic congestions, reasons for congestion and their effects. These reviews are very helpful for this research work to understand congestion problem. Various traffic measuring techniques are also discussed which are very useful to detect different types of vehicles. In this research work video image processing system is used. Various video image processing techniques are discussed with their advantages and disadvantages. Data mining is the main feature of this system. Review on various data mining techniques also is presented. Data mining techniques are used for mining various kinds of data. So by using these features traffic congestion data can also be analyzed.

CHAPTER 3

PRESENT WORK

3.1 PROBLEM FORMULATION

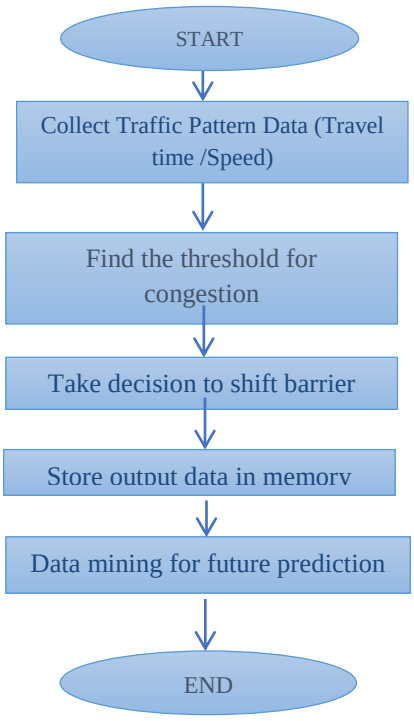
In this research works a new data mining based automated traffic congestion management system with the concept of changing the number of lanes is introduced. This chapter discusses the framework used to design this system. This research work is designed into two stages

Stage 1

In first stage, the numbers of vehicles are detected with the video image processing system. By counting a number of vehicles on both sides at fixed time interval will detect the congested side of the road and take the decision to shift barrier towards the less congested side of the road. This data is stored for further processing.

Stage 2

In second stage data analysis is done to take future decision on the collected data from first stage.



.

Figure 3.1: Problem Formulation

3.2 OBJECTIVES OF THE STUDY

The proposed system is designed for managing the traffic congestion by changing the road area /size for traffic moving in a particular direction. The primary objective of this work is to deal with congestion problem in a smarter way by using a smart automated system. Following objectives define the scope of the proposed work.

1. This system specifically designed for managing recurring congestion at bottleneck road infrastructures.
2. To form easy and convenient system to make the current traffic condition adaptable by including least of physical infrastructural changes.
3. Maximizing the transportation system's efficiency by reducing travel time and congestion.

3.3 RESEARCH METHODOLOGY

3.3.1 Video Image Processing

This proposed system identifies the speed and direction of different vehicles in a video with stationary background in the consecutive frame scanning (as shown in figure 3.2).This part introduces our approach of creating a system of vehicle detection and counting from the consecutive video frame. This work will start with overall framework of the system.

A. Block diagram

1. Capture the video, loading the video from pc to Matlab.
2. After then normalization in video try to correct for lighting effect over the images intensities which is store in buffer, which will approximate the average level across the image. Showing video frame wise
3. The two subsequent frames can be subtract which are clicked by the camera, the part of image which does not change (background) get subtracted to give zero intensity.
4. After subtracting the current image from the current background, the resultant difference image has to be threshold to get the binary.
5. Apply the morphological operation preserve the image shape and make it simple and increase image shape and make it simple and increase the quality of the image. Morphological operations are used generally for the image preprocessing.

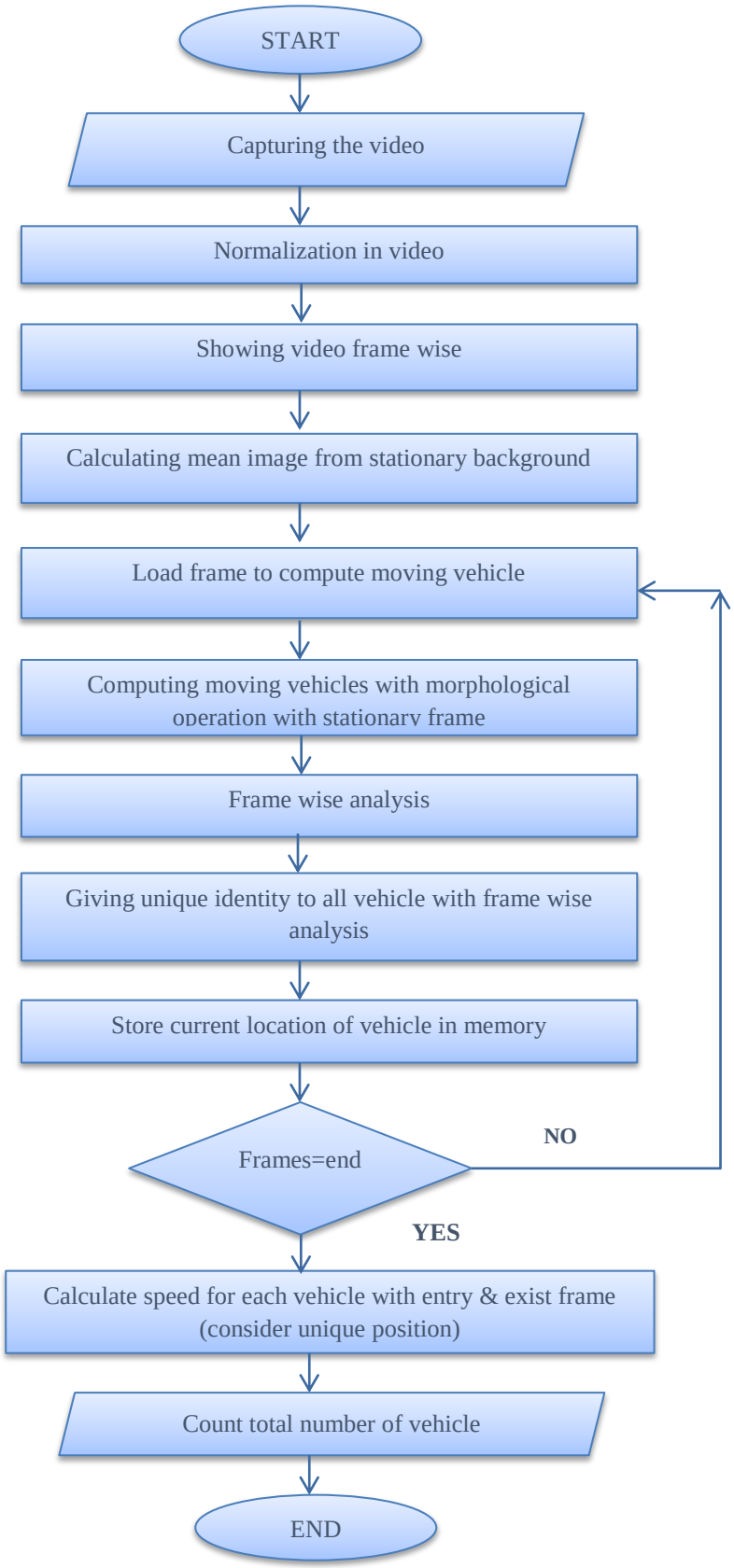


Figure 3.2: Vehicle Detection

6. "Consecutive image" which will traverse the whole pixels of background image and it has to compare the RGB value of each pixel at the same time during traversal in the consecutive image process
7. Frame wise analysis, and giving unique identity to all vehicles with frame wise analysis for detection of vehicle crossing on the road.
8. To Store current location of each vehicle in the buffer , when number of frames end time per frame is computed which is 1/20 per frame and calculate speed of multi vehicle with enter and exist frame calculate the direction and plot the graph showing the speed in km/hour and meter/seconds.

B. Model Description:

The first and foremost step in the vehicle detection system by using consecutive frame scanning loading video from pc to Matlab. Calculate the mean image from stationary background after then frame wise analysis and last calculate the speed.

C. Algorithm Outline

- ✓ DATA ACQUISITION PROCESS
    1. MEAN IMAGE
    2. THRESHOLD
    3. MORPHOLOGY OPERATION
    4. NORMALIZATION OF FRAMES
- ✓ CONSECUTIVE FRAME ANALYSIS ALGORITHM
- ✓ VEHICLE DETECTION AND UNIQUE NUMBER
- ✓ SPEED CALCULATION AND DIRECTION OF THE VEHICLE

- **Data Acquisition Process**

The first component is image acquisition. The processes in this component are "video stream" which is an input from the video camera "single-sequence image" which comes from capturing the video stream to single-sequence image in one frame, "store image in buffer" which is used to store the images in the buffer, and "display captured image" which will show the image in picture box.

1. Mean Image

When two subsequent frames got subtracted clicked by our cam, the part of image which does not change (background) gets subtracted to give zero intensity (black). Only the part of the image moved (moving object) don't get reduced to zero as intensity of pixels of two consecutive frames are different. So it provides non zero intensity of pixels corresponding moved objects. Rest is simple. Just convert the image into binary and obtain the centroid of largest area of connected pixels.

- Loop for taking image for R frames
- Loop for taking image for G frames
- Loop for taking image for B frames
- Creating mean image from R, G& B frames

2. Threshold

After subtracting the current image from the current background, the resultant difference image has to be threshold to get the binary. Since the background changes dynamically, a static threshold is not used to compute the object .The object is used for updating the current background. In this manner we require an approach to change the threshold as the current background changes. The difference image is used to update the threshold. In this manner an approach require to refresh the limit as the present foundation changes. Therefore the difference image would consist of a large number of pixels having low values, and a small number of pixels having high values. This observation is used for deciding the threshold. The brightness of surrounding environment is one of the factors that can affect the result because the brightness of environment affects the color value of the road, for this purpose to set the threshold value

3. Morphology Operation

Morphological operations preserve the image shape and make it simple, and increase the quality of the image. Morphological operation are used generally for the object structure improvement, image preprocessing. Morphology operations apply a structuring element to an input image, creating an output image of the same size. In a morphological operation, the value of each pixel in the output image is based on a comparison of the

corresponding pixel in the input image with its neighbor. The most basic morphological operation is dilation and erosion. Demising of lower frequencies, smoothing of edges and refilling remaining holes. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on the object boundaries.

4. Normalization Of Frames

Correct the lighting effect over the image intensities, which will approximate the average level across the image. By withdrawing this shading plane from the novel image, normalize lighting situations transversely the image.

For color RGB images, simply repeat the process on each channel separately, or even apply it to a different color space.

A method for normalizing the lighting changes in an input image having various illuminated intensities.

- **Consecutive Frame Analysis Algorithm**

The processes in this component are “consecutive image” which will traverse the whole pixels of background image and it has to compare the RGB value of each pixel at the same time during traversal in the consecutive image process, “store the different RGB value” which is a process that stores the values that come from consecutive image process and it will store value as binary number and, “display binary image” which is used to show the image after previous processes are done, and the image is displayed in black and white. The brightness of surrounding environment is one of the factors that can affect the result because the brightness of environment affects the color value of the road.

- **Vehicle Detection And Unique Number**

In this thesis a program is designed to detect multi vehicles on the road which are recorded in video file. Moreover the program can detect the cars. The program contains picture of the road that have no cars. As a background picture and then the program can differentiate the cars from the road and make car detection with a captured image. After the car detection process is done, the program will give a unique number to car and show the number of cars in the program. In order to distinguish moving pixels from stationary pixels, first apply a frame wise analysis on video. The intensity of the objects (in motion)

increases or decreases sharply. This gives us the background image without any objects in motion.

After the training phase, initial background model for each pixel is obtained. The resulting background image is then used in the background subtraction process. Background subtraction results in an image containing only the moving objects for every frame in that scene. After subtraction convert the resulting image to a binary image.

- **Speed Calculation and Direction of the Vehicle**

By noting down the position of a vehicle in each frame, velocity of the vehicle can be calculated. A unique id is provided to each vehicle for further calculations

Frame numbers are used to know the direction of the vehicles. These frames further can save numbers both when vehicle enters and when it leaves the screen.

Title : Calculation of Speed

Time for one frame=1/y

Frame taken=length (position vehicle)

*Time taken= frame taken *time for one frame*

Speed= distance/time

*Speed 1= speed * 3.6 km/hour*

Where, Distance=x

Frame per second=y

3.3.2 Detecting Congestion and Shifting Barrier Algorithm:

Title: Algorithm for Lane Shifting :

START

STEP 1:

Check number of vehicles in both lanes.

STEP 2:

If NL-NR>Threshold value

Then Display message: Shift barrier toward right.

Else If

STEP 3:

If NR-NL>Threshold value

Then Display message: Shift barrier toward left

Else

STEP 4:

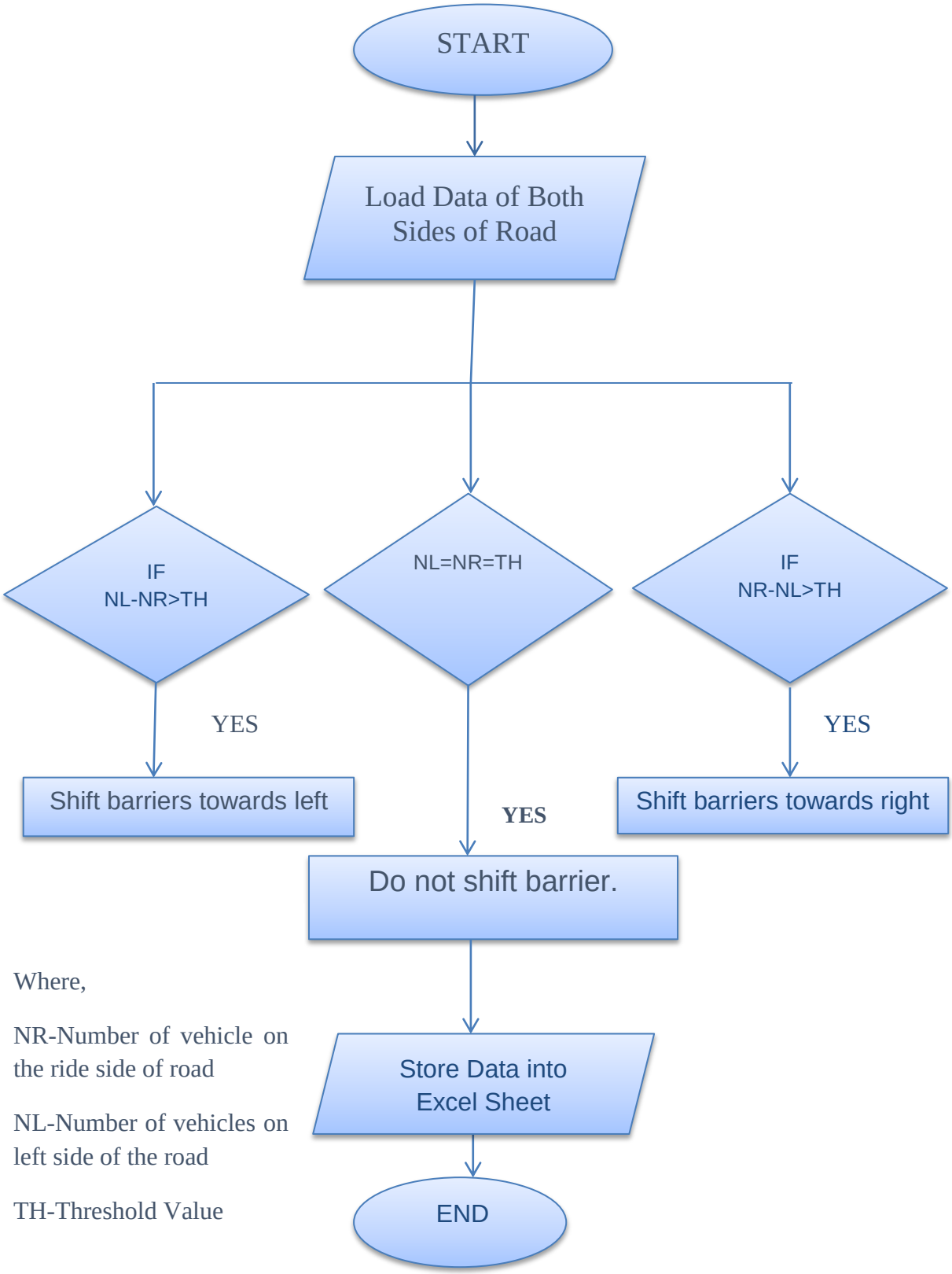
Display Message: Don't shift barrier.

END

Where, NR-Number of vehicles on right side of road

NL-Number of vehicles on left side of road

After storing the data, analysis on vehicles density on both side of the road is done to take a decision to adjust barrier on a particular side of the road. By fixing a threshold value of the vehicles difference on each side of road system is taking decisions to shift barrier in particular direction. Then decision is taken to shift the barrier in particular direction with respect to threshold value as shown in Figure3.3.



Where,

NR-Number of vehicle on the ride side of road

NL-Number of vehicles on left side of the road

TH-Threshold Value

Figure 3.3: Decision System

3.3.3 Data Mining:

In traffic data analysis process previous week's traffic decisions are classified into three categories.

1. Shift barrier towards right side at a particular time for pre-defined time duration in the morning time.

2. Shift barrier towards left side at a particular time for pre-defined time duration in the evening time.

3. Barrier will remain stationary and number of lanes on either side will not be changed.

Depending on these decisions future lane shifting operations are scheduled. On these decisions it is decided that when should the barrier shifting should take place and in which duration.

3.3.3.1 Data Mining Tool

For this research work WEKA tool is selected for data mining. A variety of formats: WEKA's ARFF format, CSV format, C4.5 for-mat, or serialized Instances format. We select ARFF format here. Practically, WEKA tool supports to build a broad range of algorithms and also supports for very large data sets, so we decided to use WEKA tool.

3.3.3.2 Training Dataset

The initial step in data mining is to collect data. Here data is collected from video image processing system. In this training set of three days vehicles data from videos is collected. In pre-processing on collected data appropriate classes are designed Information get for each attribute is calculated.

A. Naïve Bayes:

It is a conditional probability approach. In which a mention problem case is to be classified, and it can be elected by a vector x=(x1, x2……xn) representing a few *n* features. Using Bayes theorem it is written as

Posterior=prior X likelihood/evidence. Eq 3.1

- **Naive Bayesian Classifiers**

Naive Bayesian classifiers assume that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is called class conditional independence. It is made to simplify the computations involved and, in this sense, is considered "naive." Bayesian belief networks are graphical models, which unlike Naive Bayesian classifiers allow the representation of dependencies among subsets of attributes. Bayesian belief networks can also be used for classification [39].

- **WEKA classifiers bayes Naive Bayes**

Result

A total of 43records were taken for the analysis. The flat file is used in arff (Attribute Relation File Format).

- **Viewer of WEKA**

The Result is Split into Several Sections

- **Run information**

A list of information giving the learning scheme options, relation name, instances, attributes and test mode that were involved in the process.

- **Classifier model (full training set)**

A textual representation of the classification model that was produced on the full training data.

The results of the chosen test mode are broken down thus.

- **Summary**

A list of statistics summarizing how accurately the classifier was able to predict the true class of the instances under the chosen test mode.

Following results have been imported from the WEKA tool

Run information in Naïve Bayse

Scheme: weka.classifiers.bayes.NaiveBayes

Relation: congestion_prediction

Instances: 43

Attributes: 4

Left Lane

Right Lane

Time of Congestion (Morning=1(8:30 to10:00) to, Evening=2(4:45pm to 6:00pm)

Direction of barrier (right=1,left=0,stable=2)

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Naive Bayes Classifier

	Class		
Attribute	0	1	2
	(0.35)	(0.39)	(0.26)
==========			
Left Lane			
mean	3.3333	8.7059	4.4545
std. dev.	0.8692	1.1254	1.1571
weight sum	15	17	11
precision	1	1	1
Right Lane			
mean	9.0667	3.6303	3.8442
std. dev.	0.9757	0.8969	0.8816
weight sum	15	17	11
precision	1.1429	1.1429	1.1429
Time of Congestion(Morning=1(8:30 to10:00) to, Evening=2(4:45 to 6:00)			
1	1.0	18.0	6.0
2	16.0	1.0	7.0
[total]	17.0	19.0	13.0

Time taken to build model: 0.01 seconds

Evaluation on training set

=== Summary ===

Correctly Classified Instances	42	97.6744 %

Incorrectly Classified Instances	1	2.3256 %
Kappa statistic		0.9648
Mean absolute error	0.0224	
Root mean squared error	0.1178	
Relative absolute error	5.093 %	
Root relative squared error	25.0625 %	
Total Number of Instances	43	

B.J48 Decision Tree

Decision tree J48 is the implementation of algorithm ID3 (Iterative Dichotomiser 3) developed by the WEKA project team. J48 is the open source implementation of the C4.5 algorithm.

Run information in J48

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: congestion_prediction

Instances: 43

Attributes: 4

Left Lane

Right Lane

Time of Congestion(Morning=1(8:30 to10:00) to,Evening=2(4:45 to 6:00)

Direction of barrier(right=1,left=0,stable=2

Test mode: evaluate on training data

=== Classifier model (full training set) ===

J48 pruned tree

Right Lane <= 5

Left Lane <= 6: 2 (12.0/1.0)

Left Lane > 6: 1 (16.0)

Right Lane > 5: 0 (15.0)

Number of Leaves: 3

Size of the tree: 5

Time taken to build model: 0 seconds

Evaluation on training set

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances	42	97.6744 %
Incorrectly Classified Instances	1	2.3256 %
Kappa statistic	0.9648	
Mean absolute error	0.0284	
Root mean squared error	0.1192	
Relative absolute error	6.4872 %	
Root relative squared error	25.4819 %	
Total Number of Instances	43	

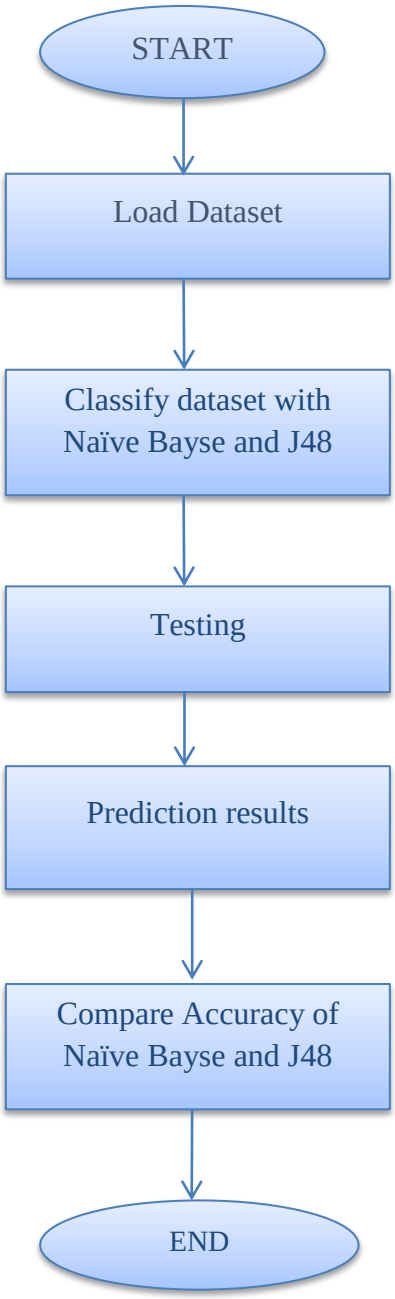


Figure 3.4: Prediction Analysis

3.4 PROPOSED SYSTEM

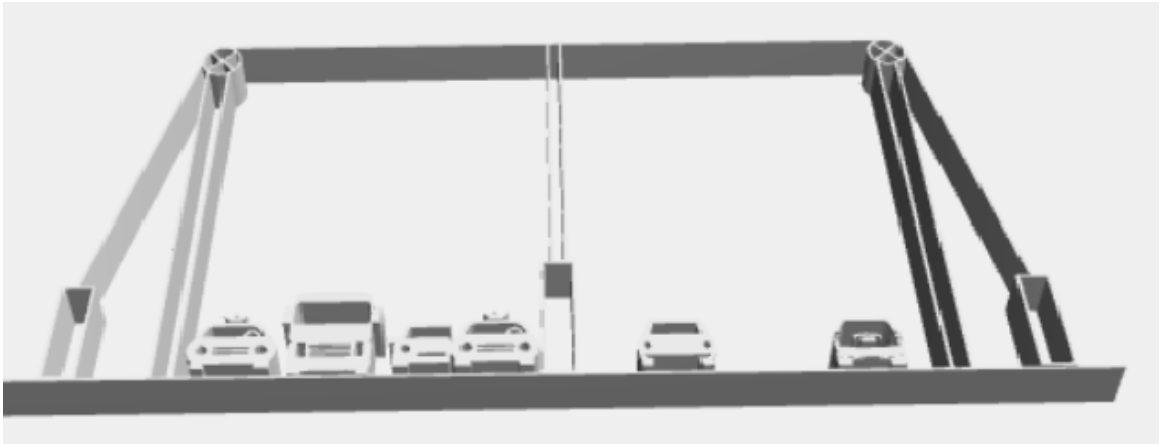


Figure 3.5: Before shifting Barrier

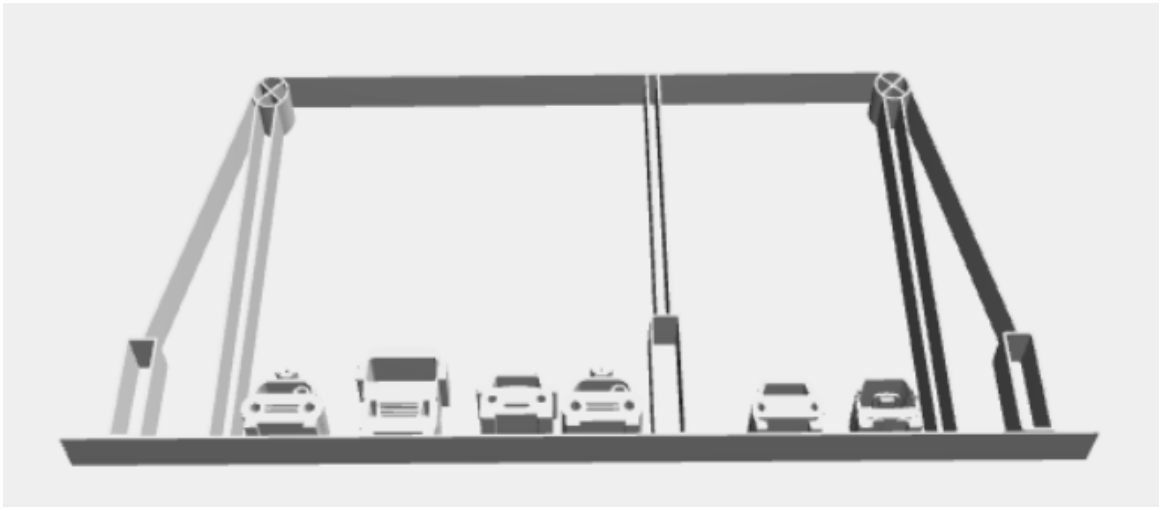


Figure3.6: After Shifting Barrier

CHAPTER 4

SOFTWARE USED

4.1 MATLAB 8.1

MATLAB is a cooperative system whose basic element is an array that does not require dimensioning. It was originally written to provide easy access to matrix software developed in the LINPACK and EISPACK project.

Features of Matlab

1. High-execution propels calculations for numerical counts, particularly in the field of lattice polynomial math.
2. A substantial gathering of pre-characterized scientific capacities and the capacity to characterize one's own particular capacities.
3. 2D and 3D design for plotting and showing information.
4. Powerful, lattice/vector-arranged, abnormal state programming dialect for individual applications.
5. Compatible with projects written in different dialects and for bringing in and sending out designed information.
6. Toolboxes accessible for taking care of cutting edge issue in a few application ranges.

Application of Matlab

1. Math and computational
2. Algorithm improvement
3. Modeling, reproduction and prototyping.
4. Data investigation, representation.
5. Scientific and building illustrations.
6. Application improvement, including GUI building.

4.2 WEKA 3.8.1

WEKA is a collection of machine learning algorithms for data mining tasks.

Features

1. WEKA is used for research, education, and applications.
2. WEKA is written java language and it is compatible with each platform.
3. WEKA is informal to usage and also useful at several different levels.
4. It is platform independent & portable.
5. It is freely available under GNU.

Applications

1. WEKA accumulates a far reaching set of information pre-handling devices, learning calculations and assessment techniques, GUIs (incl. information perception) and condition for contrasting learning calculations.
2. WEKA is extensible and has turned into a gathering of machine learning calculations for taking care of certifiable information mining issues.
3. There are three actualized plots in WEKA.
   - Implemented plans for arrangement.
   - Implemented plans for numeric forecast.
   - Implemented "meta-plans".
4. WEKA as a matter of course contains a substantial assortment of soft wares which are valuable in pre-preparing of datasets. This software helps in perusing the information from records, executing separating calculation and giving code to assess the outcomes.

4.3 Operating Environment

WINDOWS and its higher versions can be used as the operating environment. For this system windows operating system by Intel i3 and 2 GB ram is used.

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 EXPERIMENTAL RESULTS

In this research work an attempt is made to design a data mining based automated system depending on the varying traffic patterns.

In first phase a multi vehicle identification system is designed which detects the number of vehicles and their speed. A program is created, to distinguish multi vehicles which are recorded in video document. This system can also detect small vehicles. To save the picture shape morphological operations are used. A morphological operation increases the quality of the picture and also maintains the standardization of video which is captured through camera. By detecting number of vehicles in particular direction this system takes decisions to shift barrier in the less congested side of the road.

In second phase prediction analysis is performed depending on the previous week's results. In prediction analysis the direction are forecasted in which barrier should be shifted for a particular time interval. Two classification algorithms are used for this analysis. The accuracy of both algorithms is 97.6744%. For showing results of traffic data prediction analysis GUI is used. For representing the accuracy of the used algorithms bar graph is also used.

The results of this research work are obtained by using Matlab8.1 on windows operating system by Intel i3 and 2 GB ram. GUI is used for showing the output of vehicle detection and decision making.

Overall, this data mining based automated framework performed well under preparatory testing. As the analysis was directed on a little scale one must be careful in the investigation of the outcomes.

5.1.1 Results Using GUI

A. Congestion detection and Decision Making:

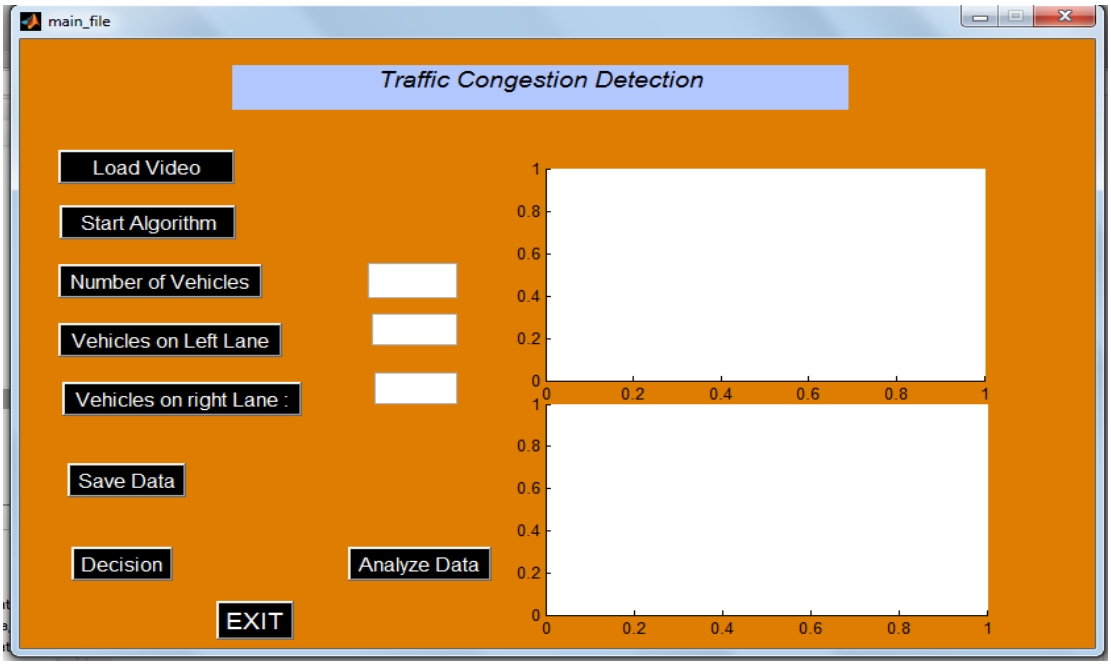


Figure5.1:Main window

In Figure5.1 Main GUI window for processing video of the vehicles is displayed. This main window contains different buttons for performing various operations of this research work.

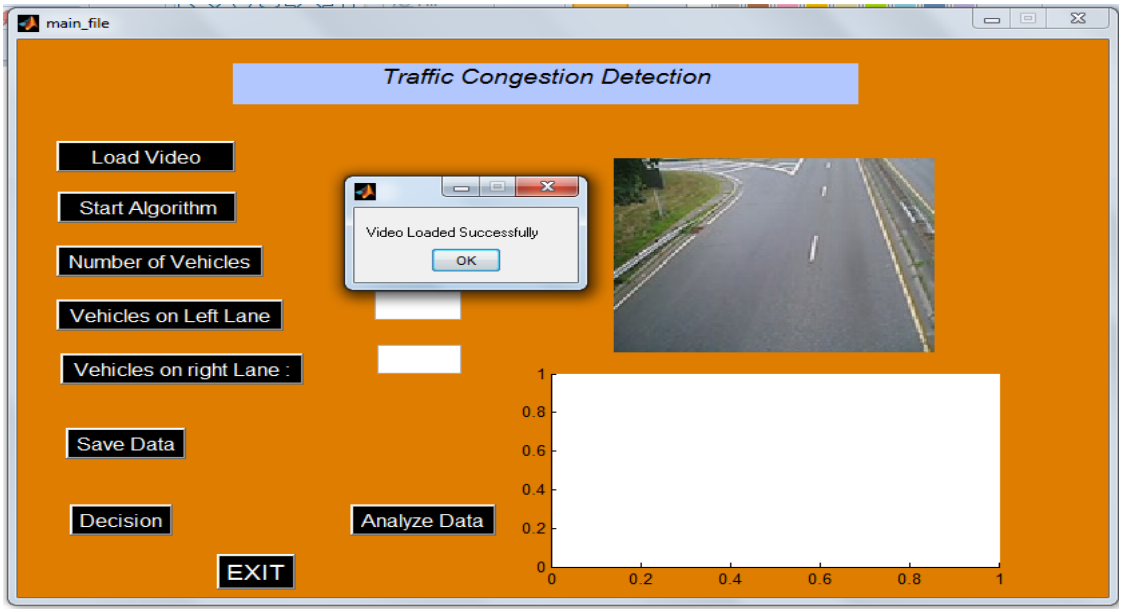


Figure 5.2: Frame for Video sequence

In Figure 9 sample frame for video sequence is displayed.

Figure 5.3: Start Algorithm

On starting algorithm user will be able to draw manually two points. The user may first draw on the upper side and second point lower side road. Pointer is displayed in Figure5.3.

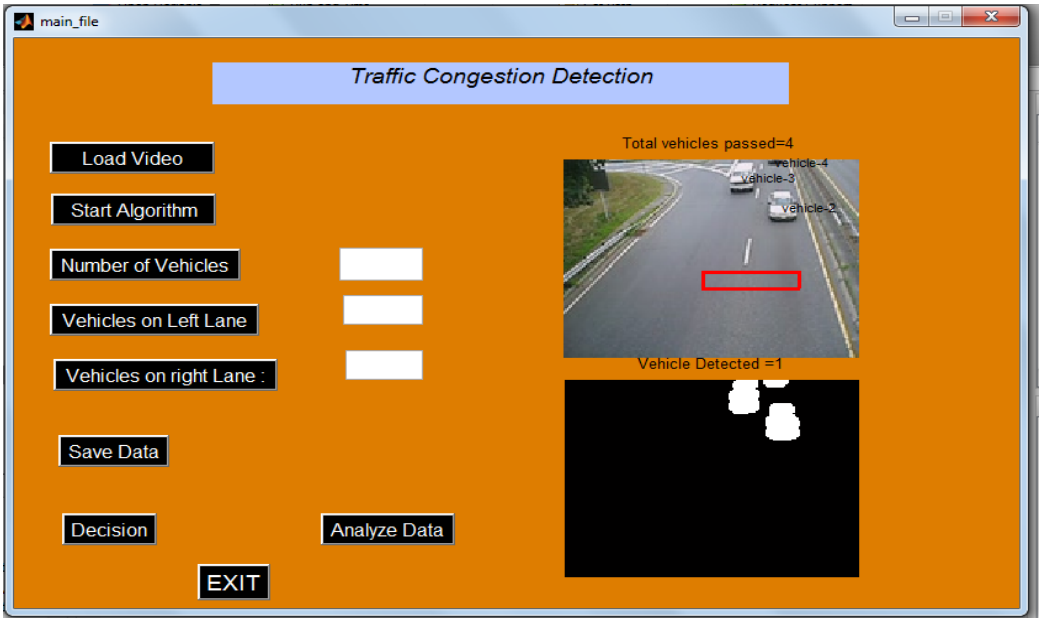


Figure5.4: Vehicle detection and mean image

Special red lines spots can be indicated on the image, a rectangular box counts the vehicles which are passing from the road and mean image as shown in Figure 5.4.

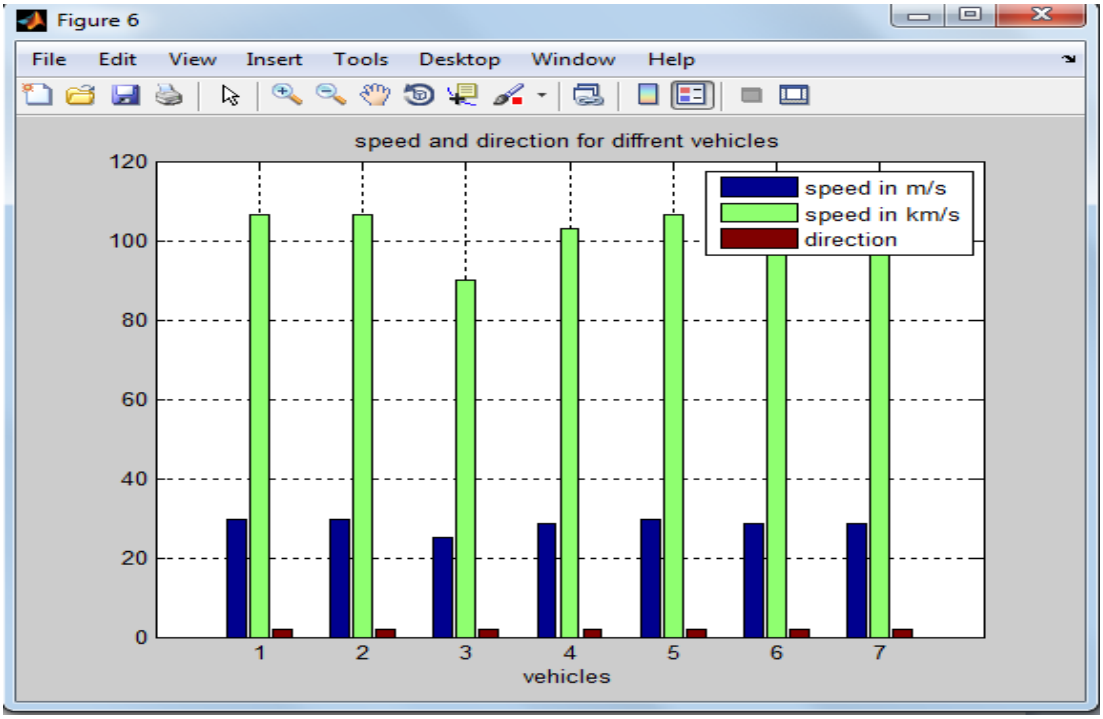


Figure 5.5: Speed and direction for different vehicles.

Table 5.1 : Speed and direction for different vehicles

Vehicles	**Speed m/s**	**Speed Km/hour**	**Direction N-S**
1	30.76923	110.7692	2
2	30.76923	110.7692	2
3	29.62963	106.6667	2
4	29.62963	106.6667	2
5	25	90	2
6	25.80645	92.90323	2
7	34.78261	125.2174	2

Table 4 shows the output of the video processing system. This system will detect the vehicles for both side of the road.

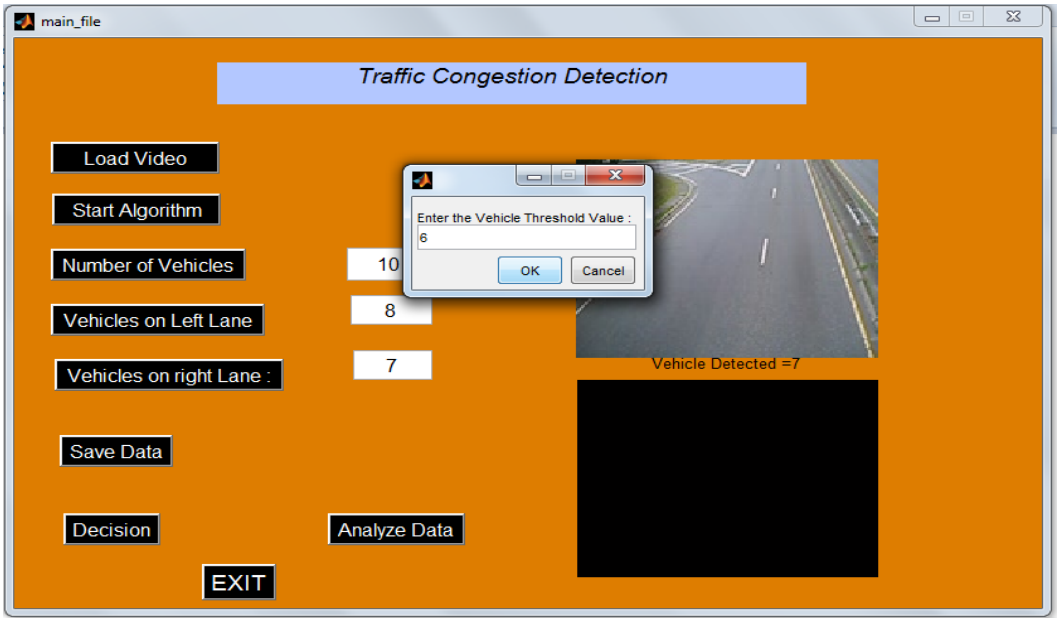


Figure 5.6: Decision to adjust barrier

User will have to enter a threshold value to take decision. As shown in Figure 5.6

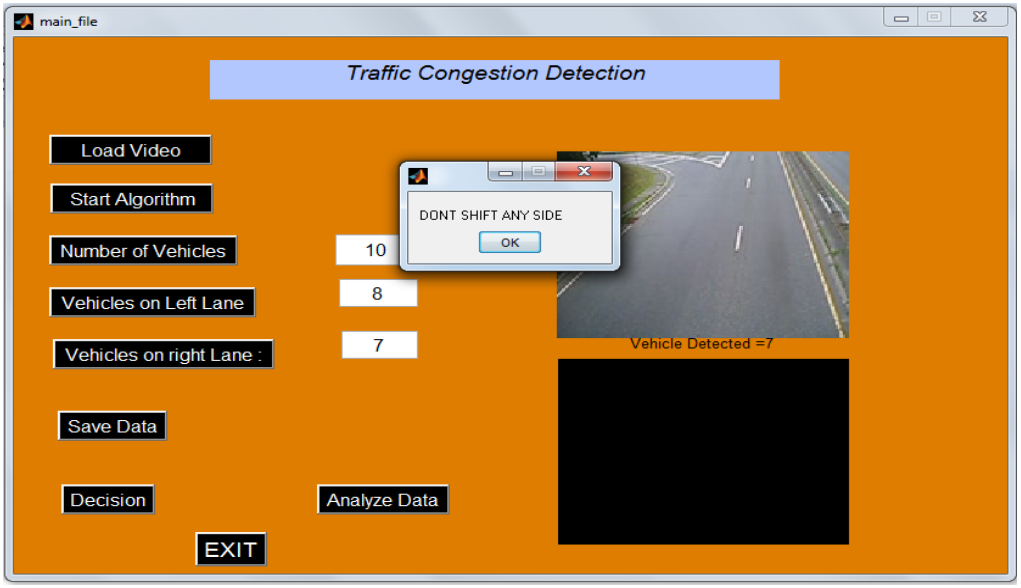


Figure 5.7: Decision taken by the system

In Figure 5.7 The system will display the message by taking the decision.

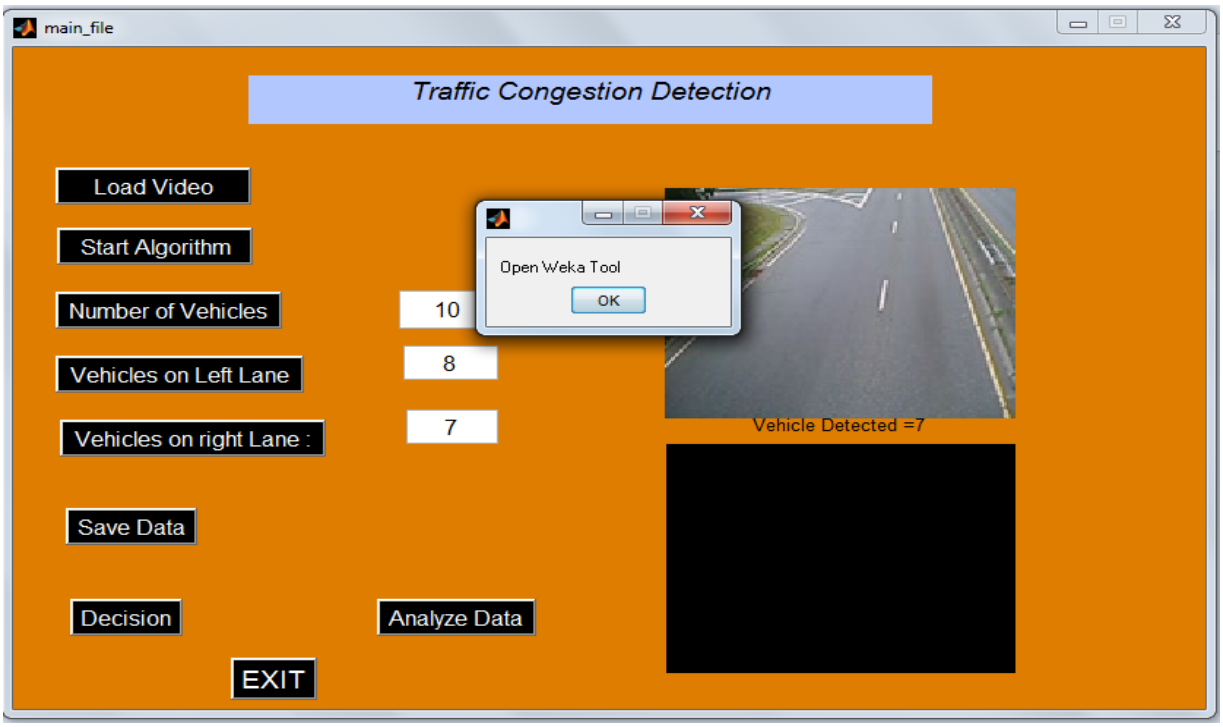


Figure 5.8: Open WEKA Tool

For further processing data will be analysed in WEKA Tool. So system will display message to open WEKA.

B. Prediction Analysis:

Table 5.2 : Data set in nominal form

SR. NO	LEFT LANE	RIGHT LANE	TimeofCongestion(Morning=1(8:30am to 10:00am)to, Evening=2(4:45pm to 6:00pm)	Directionofbarrier(right=1,left=0,stable=2
1	10	2	1	1
2	8	3	1	1
3	10	3	1	1
4	9	4	1	1
5	9	2	1	1

6	8	4	1	1
7	6	5	1	2
8	3	4	1	2
9	4	9	2	0
10	3	10	2	0

Dataset in .csv form and converted into nominal form. Convert dataset into arff form with WEKA for further processing.

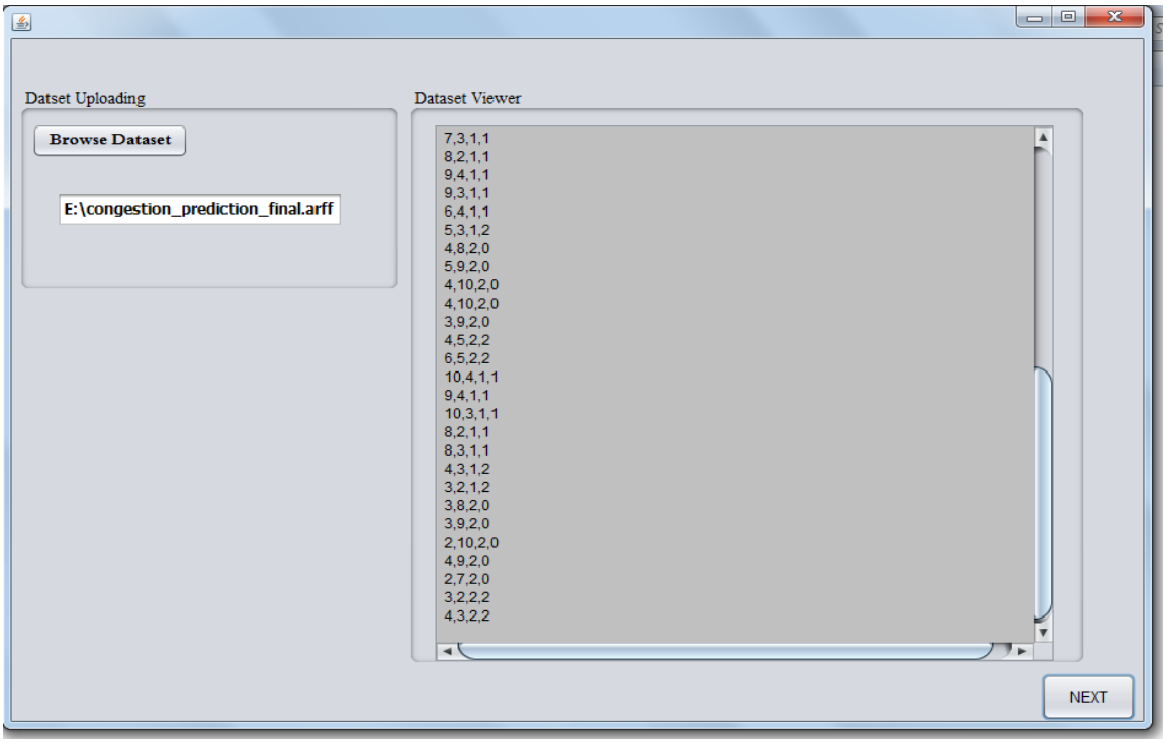


Figure 5.9: Browse Dataset

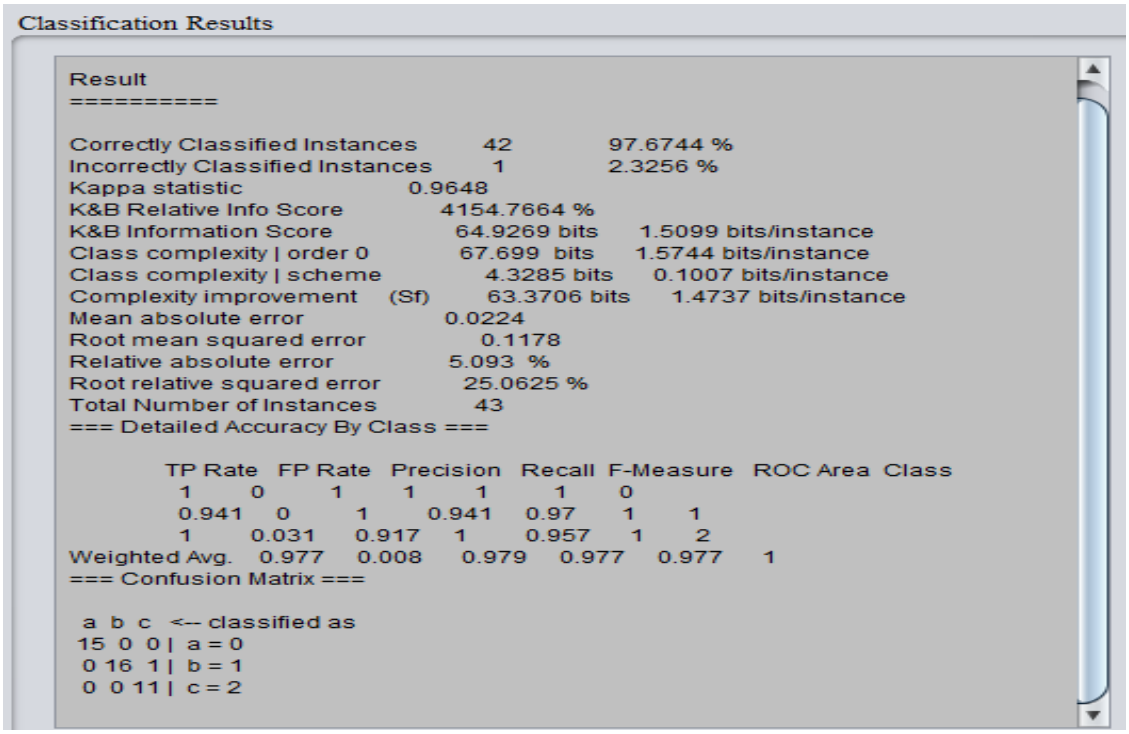


Figure 5.10: Classification Results for Naive Bayse

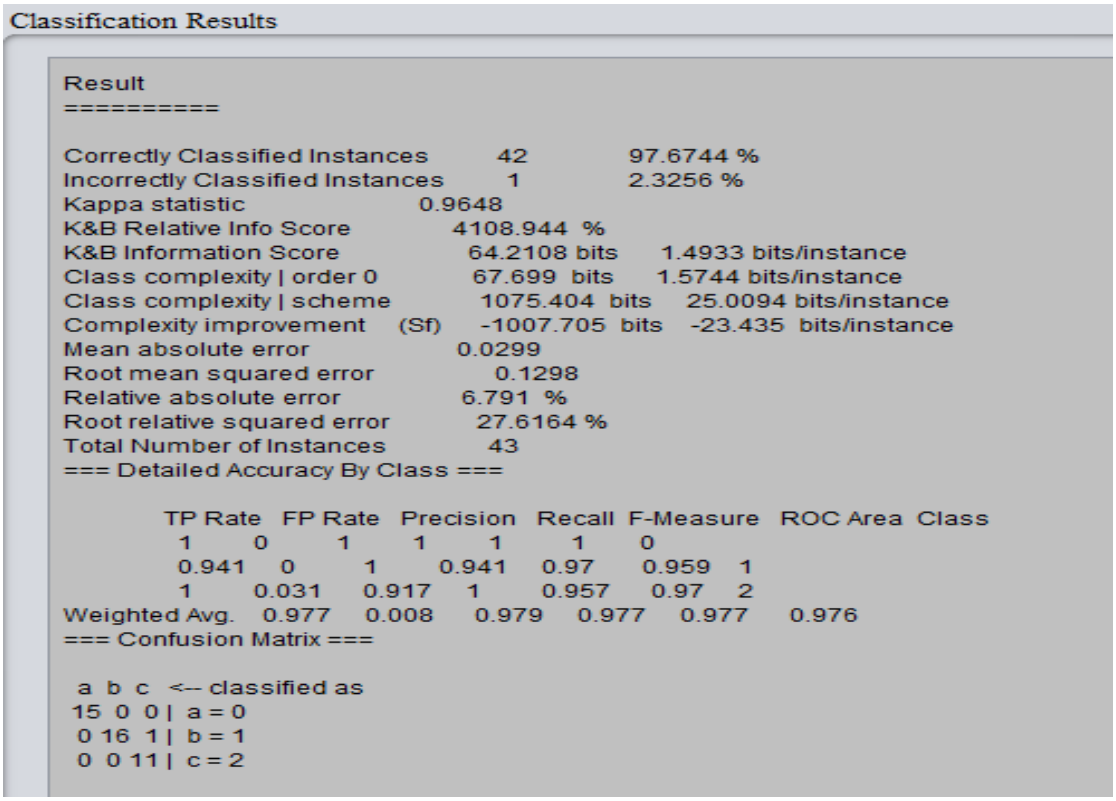


Figure 5.11: Classification Results for J48

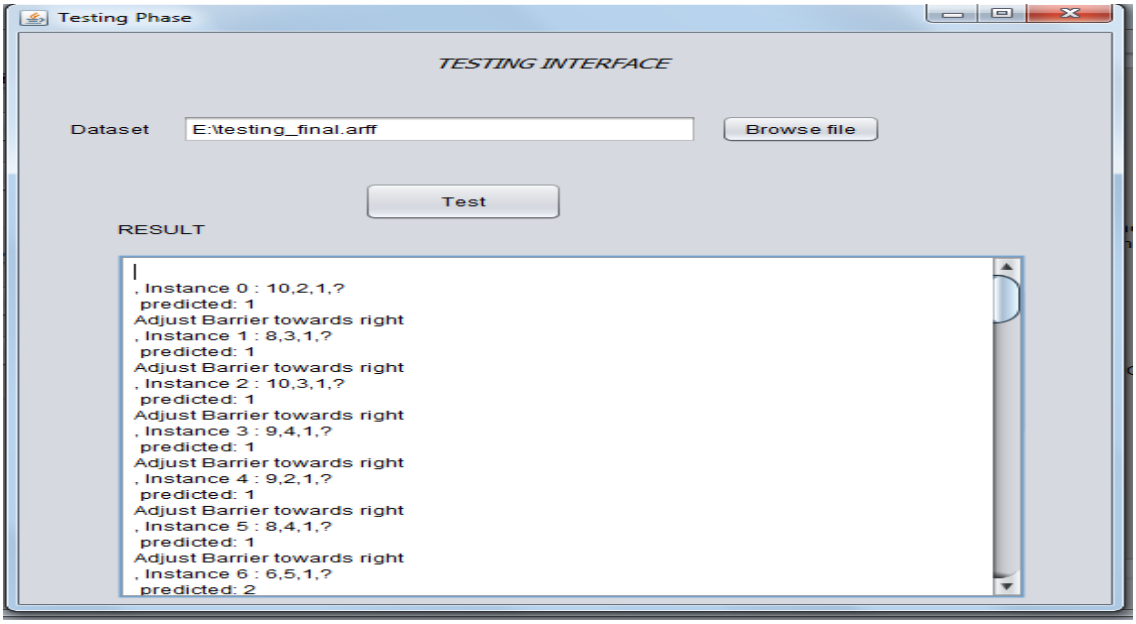


Figure 5.12: Testing

Browse the testing dataset arff file for processing and test the dataset with respect to both algorithms.

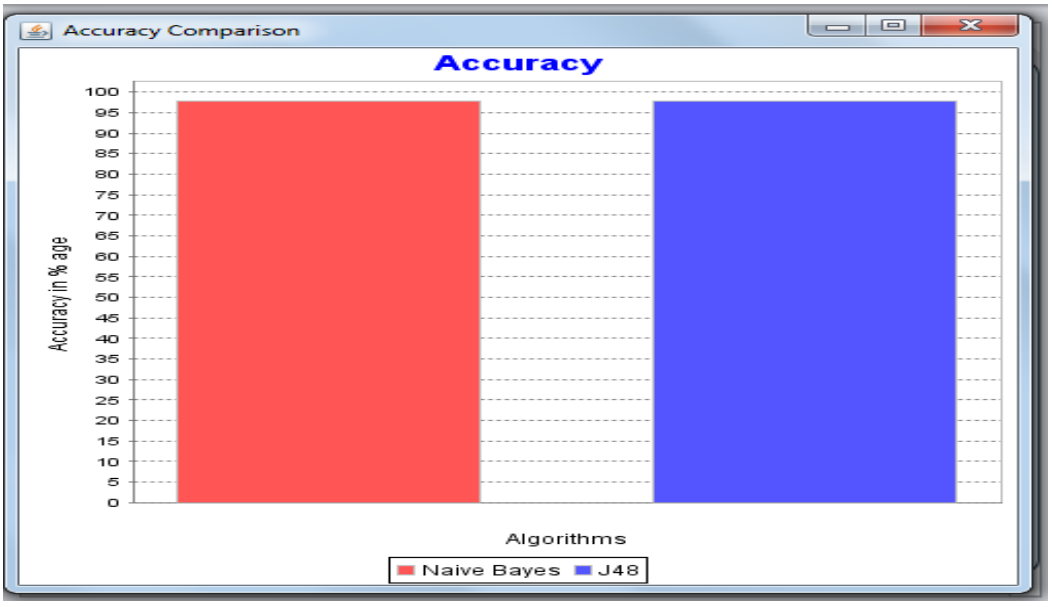


Figure 5.13:Accuracy comparison

Accuracy graph for Naive Bayse and J48 are shown in Figure5.13.

The accuracy of Naive Bayse and J48 is equal (97.644%)

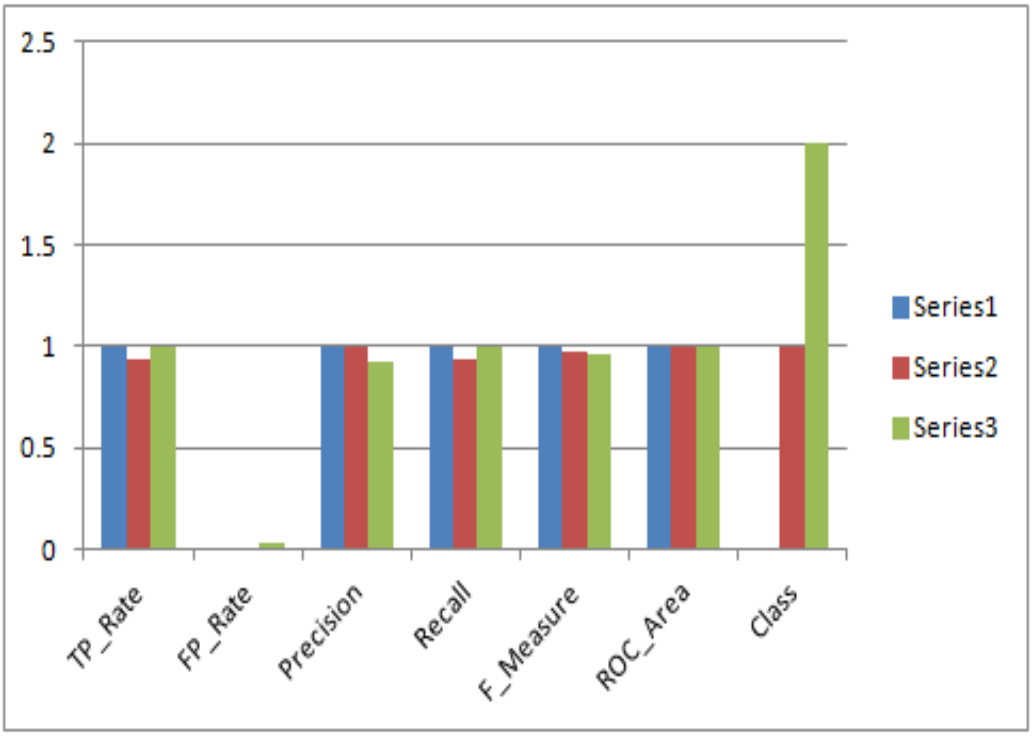


Figure 5.14: Accuracy by class in Naive Bayse

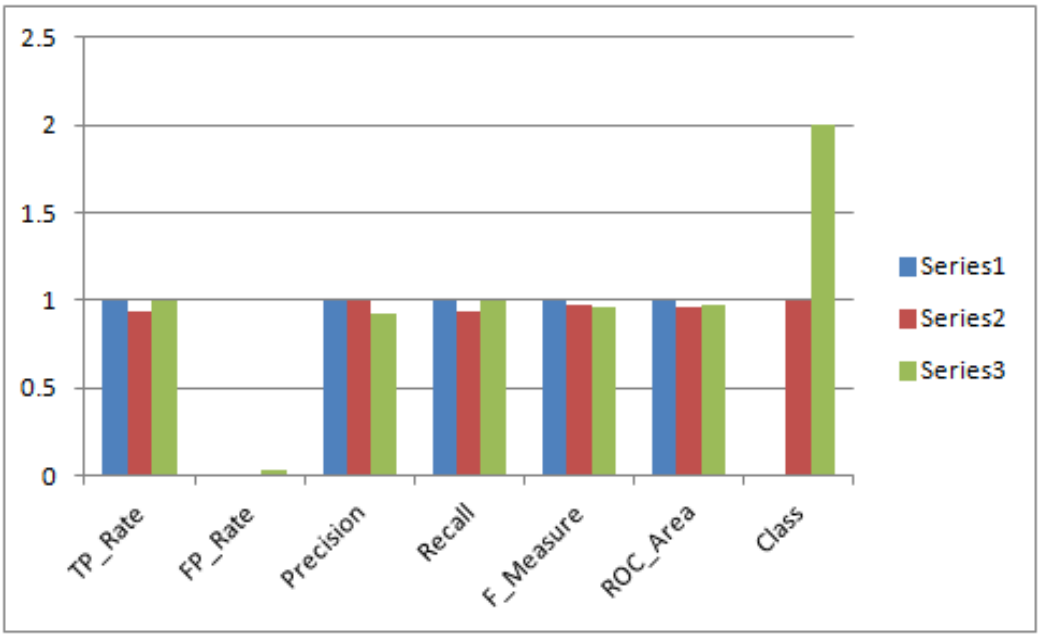


Figure 5.15: Accuracy by class in J48

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

6.1 CONCLUSION

In this research, a new automated system for shifting a number of lanes depending on the varying pattern of traffic is introduced. Before this system various congestion managing systems were available but till now no system is fully successful in managing congestion. This system is basically designed to deal with congestion which takes place because of bottleneck situation at peak traffic times. Congestion mainly occurs at particular peak interval of time like employees going to the workplace, students going to school at morning time and returning to their home at evening. It results in delay, energy loss, and economic loss as well. So this system will help to overcome from these bad influences of congestion and deal with congestion issue by using the intelligent system in a smart way.

This system replacing most commonly used active sensors for traffic measurement with video image processing vehicle detection system. It is concluded that the correctness of the video image processing results depends on frame wise analysis. This video image processing system cannot differentiate some vehicle size that is too small or too large. The system may not differentiate some vehicles of which the colour is similar to that of the road in the background image by using normalization of lighting condition and fixing threshold point. In prediction analysis part two famous approaches are used which are: Naive bayse and J48 .These are common famous algorithms which are popularly used in data mining. The Accuracy of both algorithms in the proposed work is 97.6744%.

6.2 FUTURE SCOPE

The scope of this research work is to design a data mining based automated system which will help in managing traffic congestion, by adapting to the varying traffic patterns especially at the bottleneck situations. This new automated traffic congestion

control system is based on shifting the number of lanes in a particular direction as per the changing extent of congestion. An additional number of lanes will be added to the congested side of the road. The proposed system will also use the calendar effects to decide the number of lanes in a particular direction of the bottleneck situation. The primary objective of the proposed method is to utilize the available road infrastructure in an efficient and effective way. Places like narrow bridges on a highway can benefit a lot from the proposed system. In future the vehicle detection system and congestion detection system can be replaced with new advanced traffic or congestion detection technologies.

CHAPTER 7
LIST OF PUBLICATIONS

- Shilpa Thakur, Ramandeep Singh **"A Review of Traffic Congestion Problem and Various Automated Traffic Measurement Sensors and Techniques"** Indian journal of science and technology , Volume 9, Issue 47, December 2016

LIST OF REFERENCES

[1] Ram Jagannathan (Leidos), Hesham Rakha (VTTI), Joe Bared (FHWA), Neil Spiller (FHWA) VASITE 2014 Annual Meeting – VA Beach, June 26, 2014

[2] Yamamoto, S., Hieida, Y., & Tadaki, S. (2006). Effects of bottlenecks on vehicle traffic. *Journal of the Physical Society of Japan*, *75*(11), 114601. doi:10.1143/jpsj.75.114601

[3] Article title. https://people.hofstra.edu/geotrans/eng/methods/bottlenecks.html. Date accessed: 04/07/2016.

[4] Hasan, Mohamad K. "A Framework For Intelligent Decision Support System For Traffic Congestion Management System". Engineering 02.04 (2010): 270-289. Web.

[5] European conference of ministers of transport, OCSE, OECD, & ECMT. (2007). *Managing urban traffic congestion*. Paris, France: OECD Publication.(OECD Transport Research Centre, 2007)

[6] *Evaluation Of Traffic Congestion Relieving Options With Using Cost-Benefit Analysis*. 1st ed. University of Hong Kong (Pokfulam Road, Hong Kong), 2005. Print.

[7] A. C. Review, "Measuring Traffic Congestion- A Critical Review 30 th Australasian Transport Research Forum 30 th Australasian Transport Research Forum," pp. 1–16, 2004.

[8] A. Mohan Rao and K. Ramachandra Rao, "Measuring Urban Traffic Congestion – a Review," Int. J. Traffic Transp. Eng., vol. 2, no. 4, pp. 286–305, 2012.

[9] Q. LUO, Z. JUAN, B. SUN, and H. JIA, "Method Research on Measuring the External Costs of Urban Traffic Congestion," J. Transp. Syst. Eng. Inf. Technol., vol. 7, no. 5, pp. 9–12, 2007.

[10] C. Wang, M. A. Quddus, and S. G. Ison, "Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England," Accid. Anal.Prev., vol. 41, no. 4, pp. 798–808, 2009.

[11] C. Antoniou, R. Balakrishna, and H. N. Koutsopoulos, “A Synthesis of emerging data collection technologies and their impact on traffic management applications,” Eur. Transp. Res. Rev., vol. 3, no. 3, pp. 139–148, 2011.

[12] B. Coifman and M. Cassidy, “Vehicle reidentification and travel time measurement on congested freeways,” Transp. Res. Part A Policy Pract., vol. 36, no. 10, pp. 899–917, 2002.

[13] W.-H. Lee, S.-S.Tseng, and S.-H. Tsai, “A knowledge based real-time travel time prediction system for urban network,” Expert Syst. Appl., vol. 36, no. 3, pp. 4239–4247, 2009.

[14] C. A. Quiroga, “Performance measures and data requirements for congestion management systems,” Transp. Res. Part C Emerg. Technol., vol. 8, no. 1–6, pp. 287–306, 2000.

[15] M. A. P. Taylor, J. E. Woolley, and R. Zito, “Integration of the global positioning system and geographical information systems for traffic congestion studies,” Transp. Res. Part C Emerg. Technol., vol. 8, no. 1–6, pp. 257–285, 2000.

[16] K. A. Brookhuis, C. J. G. van Driel, T. Hof, B. van Arem, and M. Hoedemaeker, “Driving with a congestion assistant; mental workload and acceptance,” Appl. Ergon., vol. 40, no. 6, pp. 1019–1025, 2009.

[17] W. Wen, “A dynamic and automatic traffic light control expert system for solving the road congestion problem,” Expert Syst. Appl., vol. 34, no. 4, pp. 2370–2381, 2008.

[18] F. He, X. Yan, Y. Liu, and L. Ma, “A Traffic Congestion Assessment Method for Urban Road Networks Based on Speed Performance Index,” Procedia Eng., vol. 137, pp. 425–433, 2016.

[19] W. Wen, “An intelligent traffic management expert system with RFID technology,” Expert Syst. Appl., vol. 37, no. 4, pp. 3024–3035, 2010.

[20] L. C. Davis, “Mitigation of congestion at a traffic bottleneck with diversion and lane restrictions,” Physica A, vol. 391, no. 4, pp. 1679–1691, 2012.

[21] C. Bachmann, B. Abdulhai, M. J. Roorda, and B. Moshiri, “A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling,” Transp. Res. Part C Emerg. Technol., vol. 26, pp. 33–48, 2013.

[22] X. Zhang, E. Onieva, A. Perallos, E. Osaba, and V. C. S. Lee, “Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction,” Transp. Res. Part C Emerg. Technol., vol. 43, pp. 127–142, 2014.

[23] V. Kastrinaki, M. Zervakis, and K. Kalaitzakis, “A survey of video processing techniques for traffic applications,” Image Vis. Comput., vol. 21, no. 4, pp. 359–381, 2003.

[24] J. C. Tai, S. T. Tseng, C. P. Lin, and K. T. Song, “Real-time image tracking for automatic traffic monitoring and enforcement applications,” Image Vis. Comput., vol. 22, no. 6, pp. 485–501, 2004.

[25] P. Reinartz, M. Lachaise, E. Schmeer, T. Krauss, and H. Runge, “Traffic monitoring with serial images from airborne cameras,” ISPRS J. Photogramm.Remote Sens., vol. 61, no. 3–4, pp. 149–158, 2006.

[26] Ibrahim, Osman, Hazem ElGendy, and Ahmed M. ElShafee. "Speed Detection Camera System using Image ProcessingTechniques on Video Streams." *International journal of computer and electrical engineering* 3, no. 6 (2011): 771.

[27] Rahim, Hasliza A., UsmanUllah Sheikh, R. Badlishah Ahmad, A. S. M. Zain, and W. N. F. W. Ariffin. "Vehicle speed detection using frame differencing for smart surveillance system." In *Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on*, pp. 630-633. IEEE, 2010.

[28] C. Pornpanomchai and K. Kongkittisan, "Vehicle speed detection system," *2009 IEEE International Conference on Signal and Image Processing Applications*, Kuala Lumpur,2009,pp.135-139.

[29] Jianping Wu, Zhaobin Liu, Jinxiang Li, Caidong Gu, Maoxin Si and Fangyong Tan, "An algorithm for automatic vehicle speed detection using video camera," *2009 4th International Conference on Computer Science & Education*, Nanning, 2009, pp. 193-196.

[30] Wang Jing-zhong and XuXiaoqing, "A real-time detection of vehicle's speed based on vision principle and differential detection," *2009 IEEE/INFORMS International Conference on Service Operations, Logistics and Informatics*, Chicago, IL, 2009, pp. 493-496.

[31] Yamazaki, Fumio, Wen Liu, and T. Thuy Vu. "Vehicle extraction and speed detection from digital aerial images." In *Geoscience and Remote Sensing Symposium, 2008.IGARSS 2008.IEEE International*, vol. 3, pp. III-1334.IEEE, 2008.

[32] Pornpanomchai, C., Liamsanguan, T. and Vannakosit, V., 2008, August. Vehicle detection and counting from a video frame.In *Wavelet Analysis and Pattern Recognition, 2008.ICWAPR'08.International Conference on* (Vol. 1, pp. 356-361).IEEE.

[33] Djalalov, M., Nisar, H., Salih, Y., & Malik, A. S. (2010, June). An algorithm for vehicle detection and tracking.In *Intelligent and Advanced Systems (ICIAS), 2010 International Conference on* (pp. 1-5).IEEE.

[34] Jovic, A., Brkic, K., &Bogunovic, N. (2014, May). An overview of free software tools for general data mining. In *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on* (pp. 1112-1117). IEEE.

[35] Patil, Tina R., and S. S. Sherekar. "Performance analysis of Naive Bayes and J48 classification algorithm for data classification."*International Journal of Computer Science and Applications* 6, no. 2 (2013): 256-261.

[36] Arora, Rohit. "Comparative analysis of classification algorithms on different datasets using WEKA."*International Journal of Computer Applications* 54, no. 13 (2012).

[37] Tribhuvan, A. P., P. P. Tribhuvan, and J. G. Gade. "Applying Naive Bayesian Classifier for Predicting Performance of a Student Using WEKA."*Advances in Computational Research* 7, no. 1 (2015): 239.

[38] Al Najada, Hamzah, and ImadMahgoub."Big vehicular traffic data mining: Towards accident and congestion prevention." In *Wireless Communications and Mobile Computing Conference (IWCMC), 2016 International*, pp. 256-261. IEEE, 2016.

[39] Jiawei, Han, and MichelineKamber. "Data mining: concepts and techniques." San Francisco, CA, itd: Morgan Kaufmann 5 (2001).

[40] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten,I. H. The weka data mining software: An update. SIGKDD Explor.Newsl. 11, 1 (Nov.2009), 10-18.

[41] Barria, J., and Thajchayapong, S. Detection and classi_cation of tra_c anomalies

using microscopic tra_c variables. Intelligent Transportation Systems, IEEE Transactionson 12, 3 (Sept 2011), 695-704.

[42] Rosenbloom, S. (1978) Peak-period traffic congestion: a state-of-art analysis and evaluation of effective solution, Transportation, 7(2), 167-191.

[43] Turner, S.M., Lomax, T.J. and Levinson, H.S. (1996) Measuring and estimating congestion using travel time-based procedures,Transportation Research Record: Journal of the Transportation Research Board, No. 1564, pp. 11-19.