

# A Model of Interactive Traffic Management System and Traffic Data Analysis

Md. Safayat Karim

Department of Computer Science and Engineering  
American International University - Bangladesh  
Dhaka, Bangladesh

Md. Bariul Islam

Department of Computer Science and Engineering  
American International University - Bangladesh  
Dhaka, Bangladesh

Sajib Hasan

Department of Computer Science and Engineering  
American International University - Bangladesh  
Dhaka, Bangladesh

**Abstract:-** The Study introduces a model of interactive traffic management system with a goal of controlling traffic flow, vehicle tracking, traffic control, law enforcement and automated toll collection. The intelligent transport system is one of the principal components of modern urbanization and the current systems are not automated and human error-related. To achieve these goals fifty vehicles seven days data have been analyzed through statistical tools. Data used for this study were collected from Prohori, Onnorokom Electronics Limited. All results was focused on only law enforcement and all analysis results was done using a data mining software named WEKA workbench. Several algorithms were used to find out gaps between engine on time and idle time, detect vehicle which violated rules and classify them for law enforcement and predict vehicle route for next day.

**Keywords:-** Interactive traffic management; Intelligent transport system Smart traffic management; Big data; Traffic signal control; GPS tracking system; Weka.

## I. INTRODUCTION

Transport and the toll road network are the backbone of the two systems used by public services. As such, planning and tracking for a powerful transportation device could be very vital in constructing and renovating the financial system and protection of a location. However, need for road tour attempts to amplify as populace will increase whilst new buildings have now not stored tempo. Traffic regulations may take regions in urban areas, waters, seas, air, or even underground. For example, an annual total of 40 per cent of the community spends at least 1 hour per day on the street in the site visitors jam Despite the importance of tourist signing mechanisms, a recent country wide document card gave tourists low grades signing structures for missing [Add Tarn off report information] While there are many explanations for this misjudgment, we accept that there are three key contributing factors. First, there may be a lack of high total, first-rate references to identify suitable exercise.

Because of various operations, public transport systems, including railways and metros, are constantly needed in mega- cities to boom their resistance to extreme circumstances. Tokyo, for example, which is Japan's biggest mega-metropolis, would host the 2020 Summer Olympics as a means to justify huge human movements around the vast region surrounding Tokyo. Very effective inland earthquakes in the Tokyo metropolitan area are also expected to occur.

The foremost truth which stimulated us to steer this research is that during many cities of the world, sign system remains based totally on timer. The timer manner has a drawback that even if there's less visitors on a road, green signal continues to be proven to the street until its timer price falls in to 0 even as site visitors on any other road that is greater, faces pink signal at that point which causes queue and time loss to commuters. Most of the existing structures are not automatic and are biased to human mistakes. The predominant purpose of this paper seems to be to generate a higher street network gadget in town for smoother switching of site visitors to boom a metropolis or a rustic's overall productivity.

Followings are the Goals that we would like to achieve through our proposed system. Goals are listed as paragraph headings.

*Goal One (Vehicle Tracking):* Our first goal is to be able to track owner's vehicle and interacts with traffic management system. Vehicle trackers by using GPS are designed to allow vehicle owners to track the location of owner's property over a cellular network. If owner lost their car than they can easily find their car. In Interactive traffic management system, it needs to ensure that the updates are accurate and frequent in real time.

*Goal two (Traffic Control):* Our second aim is to manipulate the site visitors via an operator. The operators can reveal the site visitor's system. The whole device is prepared to verify the safe and green motion of motors along a guideway or separation infrastructure via imparting OK, correct, well timed statistics to the operator. Here operator will be close by police stations or whose are liable for visitors' control.

*Goal Three (Traffic Flow):* Our third goal is Traffic Flow control. By using real time vehicle traffic data, the system can give the best route to travel one place to another place. How much time do they need to reach the place or predict how much time need to cross a place or if two routes are recommended than which route will be the best route.

*Goal Four (Establishing Law Enforcement):* Our fourth goal is with a purpose to set up Traffic law enforcement to the complete enforcement chain, from detection of a violation through to the penalty. Enforcement of police regulation refers to the actual paintings of tracing a site visitors regulation violation, apprehending the culprit, and securing the proof wanted for his prosecution.

*Goal Five (Toll Collection):* Our Final purpose is collection of toll invoice using Interactive Traffic machine. Collecting the toll manually is the best shape of toll series. A computerized toll collection gadget contains and keeps along with transponders which can be provided for sale to the public in sealed programs and which can be pre approved for a predetermined quantity of pay as you go toll credit. When the automobile movements via the toll gate than it has indicate the RFID reader that it has crossed the clearing.

## II. LITERATURE REVIEW

Modern age advances the development and implementation of the concept of clever, along with smart network and smart city, dramatically rapidly. Following with the aid of this chapter, the attached self-sufficient automobile machine creates a accurate solution for the improved traffic protection and flexibility through nation-of-hazard Wi-Fi communications and self-sustaining riding strategies [3]. Road accidents are attributed to human factors as a contributing element [7]. Humans have restricted statistics processing abilities and should depend upon 3 fallible capabilities: belief, attention, and memory. Drivers commit errors because the situation exceeds the limitations of 1 or extra of those three functions [2]. The main functions of making a stronger traffic management network consist of linking visitor indicators and visitor control centers with GIS allowed the city's digital street map to use critical data analytically computational electricity as a key module[7].

C. Tominski said that 3-D icons have been added to a map to represent separation temporal information[5]. In general, the evaluation topics in a multidimensional model of facts are a hard and fast numerical measure. Each numerical degree is decided by means of a difficult and fast dimensional approach. As an example, the dimensions of velocity, quantity, and occupancy are in a visitor 's knowledge, with its aspects being space and time. The dimensions of nature are Hierarchical[3]. It is worth highlighting that visualization techniques such as graphs and parallel coordinates have meanwhile, collectively with the RNA Visualization Module of Source Fire, located their way into industrial merchandising[4]. The length scales can

be classified as an example by 'Week,' 'Month,' 'Season,' or 'Year.' Likewise, distance dimensions can be divided by 'Station,'

'County,' 'Freeway,' or 'Site.' Because of the size and hierarchy, the dimensions can be aggregated into great combinations.

For example, the volumes of weekly Web page visitors can be analyzed for a given dual carriageway and a specified month [3].

Rajkumar said that using a standard tracking tool, together with surveillance cameras at a street pass, excessive route, and tunnel entry, location-based statistics can be collected. The device will not now best document the video data when a car passes but can also acquire their car, pace and path statistics. These surveillance information allows analysts to investigate exciting normal spatio temporal styles or motive for the coincidence of the traveler, and many modern methods have been proposed for visible surveillance information analytic. Harald and Matthias, for example, designed a tool known as AITVS is that could be attentive to the growing state of affairs in street tunnel surveillance[13].

In addition, basic traveler data is also set in time along with traveling places, parking location and distance based on large virtual presentations hooked up at metropolis facilities to guide drivers closer to their vacation spot [8].

M. Rath said that mobile sellers are versatile mobile programs that enable extends the range and connectivity between laptop work spaces, customers and databases, and passive components and controls. The benefit of being used as a reference for some use in the future, such that in any uncertain period they become self-reliant and re-configurable. Application-based cell agent performance is substantially impaired when in line with the official's behavior of handling the amount of complicated assignments. The task jacket throughout cautiously choosing the scripting language easily implementable for chatting between structures[6].

The use of the Internet of Things ( IOT) and Big Data[10] has been provided throughout [9] to exploit gadget through lower value real time clever web users. M. V. Moreno said that it provides advanced service by using indicators of website visitors to replace statistics related to website visitors at any location of activities within the community. This method creates low-charge sensors with the capacity of automotive device detection packages in a language of every 500 meters and a thousand meters in the middle of the lane. The IoT tools are now used to easily and similarly capture website-related online data sent to Big Reality analytics centers[10] for processing. The congested website tourist system is handled using a cellular algorithm solely based on the visitor level and interchange approach to the website being examined. Monitoring and regulating on-road air skills policies have been proposed [11] Fully built agent-based model consisting of urban road networks

and measuring the index of real-time and estimated air pollutants in exceptional route stages and delivering endorsed road user techniques. Such methods include reducing the range of vehicle usage in most segments of polluted avenues, lowering the range of pollutants on roads with increasing automotive drift. For this purpose, data sets are collected from a region of noteworthy tracking gadget, available street network statistics [14].

Based on the model of visitors to the derived website, and the SPH simulation, Y. Zhang increases the simulation system of an interactive visible website traveler in which fully numerical solutions based on SPH are provided. This machine completes the use of Microsoft Visual C++ on a Windows 10 PC pc with Intel Xeon E5-2670 CPU and 32 G reminiscence on board. Under Qt software program development framework the vacationer's simulator was advanced with the simulation primarily focused on OpenGL [12]. M. Rath suggested a new online traveler website to handle the use of a specific generation of vehicles under VANET configuration with a covered attempt to solve regular online site visitors related issues inside a high-volume tourist gateway[1].

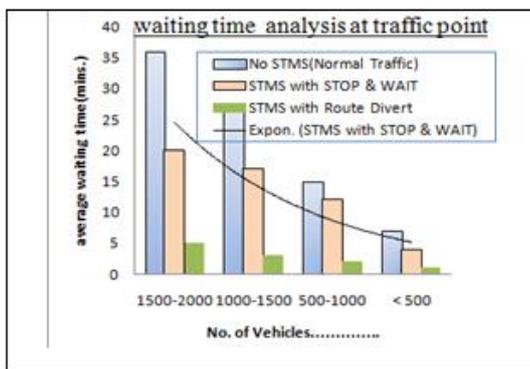


Fig 1:- Average waiting time analysis [6].

This form of community in a smart town helps and allows the groups that it touches. In a clever city, transport services can also be built alongside side facts and conversation technology alongside a likely rigid shipping community and exclusive infrastructures. However, with the aid of simple infrastructures such as governance to build and implement the smart technique, it is possible to perceive an exquisite effect on the well-being of the city's human beings, their varying positions as human beings, people, drivers, business owners. Intelligent Transport Systems (ITS) valuable tool to shipping services metropolis gadget. Considering the cutting-edge transformation taxonomy, referred to by the Western European Road Directors Group (WERD) as the 'Major Change' and integrating a move from pure avenue introduction to road community enhancement to street network customer guide, Intelligent Transport Systems (ITS) are used at all levels: strategic, tactical, and operational. These structures provide the quit user with more secure and extra-inexperienced transport structures in the form of useful real-time road records for both avenue customers and operators, allowing the preceding user to optimize their course plans based on modern conditions in the street community and the latter to

adapt avenue offers to select this optimized use of the avenue[14].

D. Singh, C. Vishnu and C. K. Mohan said that the use of the centralized or allocated fusion and the system studying statistics processing can be implemented with intelligence. Additionally, despite rapidly evolving situations, the man or woman nodes must be able to cooperate and render alternatives [7]. Information fusion is one of the main approaches for analyzing, handing over and analyzing the enormous quantities of records obtained from these networks using Machine Learning [7].

Oliveira and Camponogara [9] adopted a multiple agent method. Even though the proposed set of rules for making choices produced genuine results, they depend on the convergence of the selections of the agent. This role imposes regulations on the version's complexity, since the variables worried in a sensible version are closely coupled due to the reality. Srinivasan et al.[15] developed an artificial neural community to model the dynamics of the visitors using pattern identification. Despite having better implications than the conventional technique of adaptation, the approach is focused entirely on discovering patterns in human behaviour, which is especially stochastic. G. B. Castro [16] previously suggested a multi layer approach for providing the disbursed controllers with an international perspective of the city community [9].

Instead of trying to version human behave, our approach is focused on how to react to the current state of the website visitors based on the implications of a semaphore movement. In order to achieve this, our manipulated version allows us to incorporate understanding of the main elements of the dynamics of web page visitors as diagnosed by Prikryl [15].

Zhang, K., Qu, B. and Chi, X said that the GPS-GPRS primarily based tracking machine is a system that uses the Global Positioning System to decide the precise area of a car to which the device is connected. When a big wide variety of items or automobiles were accelerated all around the floor, the proprietor of enterprise desires to keep hint for gasoline saving, protection reason and many others. A monitoring gadget is required to discover the vicinity of any object at any given time and the distance travelled. Also, the want for a tracking machine in person car is used to control any type of robbery on the grounds that police can use monitoring reports to discover a stolen vehicle location [16].

The WEKA workbench a information mining software program software, shown in recently runs on Sun workstations under X home windows, and offers get entry to device studying gear written in a spread of programming languages (C, C++, and LISP). It isn't always an unmarried coding software program, but as an alternative a fixed of tools sure collectively thru a common purchaser interface. WEKA is a software program application workbench for system learning that is meant to useful resource inside the software of machine getting to know techniques to a spread

of actual-global issues. In order to maintain layout independence, information is transformed to an intermediate illustration called ARFF (Attribute Relation File Format). ARFF documents contain blocks describing members of the family and their attributes, collectively with all the times of the relation—and there are often very masses of those. They are saved as simple textual content for ease of manipulation [17]. This software program helps the numerous preferred records mining responsibilities, extra in particular, statistics pre processing, clustering, magnificence, regression, visualization, and characteristic preference.

### III. METHODOLOGY

This segment provides the proposed Interactive Traffic Management System and all the diverse actors that represent it and records-set we are the use of to locate our effects and analysis. Here we present a layered structure that depicts the functionalities of our interactive traffic control gadget showcases all the exclusive entities which it accommodates. After that we're going to discuss our facts, we got from prohori.com a vehicle tracking corporation of Bangladesh.

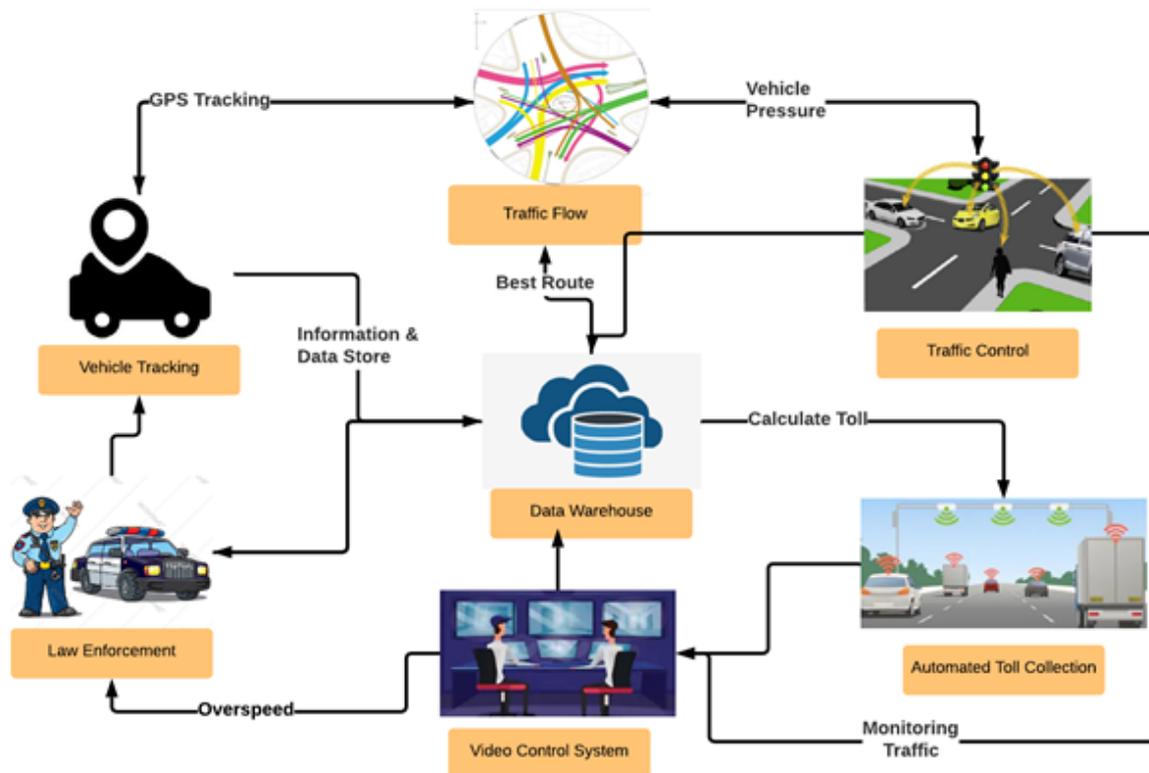


Fig 2:- Diagram of Proposed System.

#### A. Vehicle Tracking:

The system uses an On-Vehicle Module creates of GPS receiver and GSM modem; the device resides within the car to be traced. To hold the movement of the car Google Maps used for mapping the vicinity. The GSM modem consists of the GPS location and sends it to the server using GPRS. Research paintings have been made in the place of object-primarily based system ranging from GSM primarily based place willpower to GPS primarily based location determination. The integration of GPS and GSM become first created the usage of SMS as a method of transmitting GPS coordinates. The adding of GPRS technique to transmit vicinity regions to a far-flung server helps the tracking of item remotely the use of any computer connected to the web. By using the internet, the clients can browse music on digital map using motive designed web software on internet site[16].

#### B. Traffic Control:

GIS enabled virtual street map. Smart Traffic Management System: Traffic controller uses heavy website users to track and monitor the duration of the pre described rush time on the lane. By using this process, the video monitoring system checks the increased online traffic on the website via video virtual camera and while the variety of motors, through particular, may increase a pre-computed scaling factor, notifies the STMS traffic control in price with an alarm indicating "restricted reached visitors to the website" and prevents any additional car from entering that direction.

#### C. Traffic Flow:

Different data analytics tools obtain traffic enforcement social justice and use of Data collection below on time. Additionally, primary traveller information along with travelling locations, parking spaces and length is also displayed on a time-based basis on large virtual display

installation at city centers front factors to direct travelers inside their holiday spot [17].

**D. Law Enforcement:**

For law enforcement we are growing a guideline detecting the ones cars that have surpassed velocity of 60 km/h. From our records-set 50 automobile 7 days trip document. We also can come across the motors whose are befall coincidence with the aid of video tracking device. We additionally come across whose bikers are not sporting helmet.

**E. Automated Toll Collection:**

Automatic toll collection is a communication device that collects automatically the fee or toll paid to use the motors for the use of tollways, Highway lanes, pedestrian overpasses and toll bridges. In this gadget car owner will pay online toll while the car went in front of the digital

video camera will detect the wide variety of that vehicle and the device will allow the vehicle to bypass.

**F. Data Acquisition:**

Just after an automobile starts moving, the tool connected with the automobile is activated automatically. The sensors connected with the micro-controller begins sensing information. The velocity sensor starts off evolved sensing the speed of the car on occasion, and the vibration sensor senses the vibration of the vehicle. The GPS tracker additionally starts off evolved tracking the coordinates of the modern region of the vehicle. Traffic facts speak to the record units generated and accumulated by sensors in the cars or video display units of visitors mounted along the roadways. Explanations of traffic data along with cars Gps location, human mobility cell station or GSM destinations, and surveillance gadget video / photo / counting data.

C	D	E	F	G	H	I
Avg. Speed(km/h)	Max Speed(km/h)	Travel Distance(km)	Engine On Time	Engine Idle Time	Start longitude and latitude	End longitude and latitude
5.65	79	307.23	20 h 25 min	9 h 42 min (48%)	23.7097316,90.4318977	23.7097316,90.4318977
2.52	69.39	159.44	21 h 9 min	9 h 30 min (45%)	23.7132955,90.4252148	23.7132955,90.4252148
4.76	74	213.73	18 h 33 min	9 h 21 min (50%)	23.8672301,90.3968501	23.8672301,90.3968501
1.79	54	87.07	13 h 26 min	9 h 14 min (69%)	23.692279,90.4961275	23.692279,90.4961275
4.92	51	330.42	23 h 12 min	9 h 13 min (40%)	24.0487739,90.3842595	24.0487739,90.3842595
7.44	55.06	157.81	2 h 35 min	8 min (5%)	23.7335043,90.4030965	23.7335043,90.4030965
0	0	0	8 min	8 min (100%)	23.874709,90.5028489	23.874709,90.5028489
5.36	57	378.42	23 h 27 min	8 h 59 min (38%)	24.2196042,90.4081632	24.2196042,90.4081632
0	0.46	0.01	8 h 54 min	8 h 54 min (100%)	23.697334,90.5061494	23.697334,90.5061494
4	58	241.51	21 h 34 min	8 h 47 min (41%)	23.7231094,90.5115913	23.851703,90.004829
0	3.91	0.06	8 h 41 min	8 h 40 min (100%)	23.7924197,90.3887075	23.7924197,90.3887075
4.91	63	277.72	19 h 17 min	8 h 4 min (42%)	23.7229814,90.4111007	23.7229814,90.4111007
6	52	227.13	16 h 21 min	8 h (49%)	23.8186024,90.3569288	23.8186024,90.3569288
4.3	56	237.22	17 h 45 min	7 h 59 min (45%)	23.7238317,90.4074531	23.7238317,90.4074531
1.77	56	87.7	15 h 35 min	7 h 48 min (50%)	23.7816378,90.5627474	23.7816378,90.5627474
5.93	71	386.38	23 h 34 min	7 h 48 min (33%)	23.6752365,90.410536	23.6752365,90.410536
5.81	52	372.22	22 h 34 min	7 h 45 min (34%)	23.833784,90.4685941	23.833784,90.4685941
3.73	59	250.43	23 h 10 min	7 h 44 min (33%)	23.7853594,90.4140049	23.7853594,90.4140049
6.27	63	402.22	23 h 18 min	7 h 29 min (32%)	23.8452305,90.3687496	23.8452305,90.3687496
6.08	62	374.97	21 h 14 min	7 h 17 min (34%)	23.748438,90.3781068	24.098788,90.127700
4.78	54	193.46	15 h 33 min	6 h 54 min (44%)	23.7238317,90.4074531	23.7238317,90.4074531
0	58	0	16 h 32 min	6 h 5 min (37%)	23.8555738,90.516048	23.8555738,90.516048

Fig 3:- Data set of 50 vehicles 7 days trip report.

For the working purpose we gather a dataset from prohor.Com. In this dataset we get 7 days facts of 50vehicles each day. In this dataset we get which wide variety of time engine on and stale by a motive force ,Average pace of a vehicle, Maximum speed of a car , Total Travel distance of a car , Engine on time via a driver , Engine Idle time via a driver. From this given record we plot data of the Start longitude and range and End longitude and range randomly through thinking about their travel distance. Using this data set we are able to discover the statistics which vehicle is violated visitors' guidelines and which vehicle is going for walks maximum instances in street and maximum speed of a vehicle. We can use records mining for find out the result for which automobile is violated the visitor's policies for velocity and hole among the engine on time and engine idle time.

Our facts analysis turned into finished with the aid of Weka device studying software program. Initially records were stored in excel file. Our information is converted into

an intermediate representation known as ARFF (Attribute Relation File Format) to maintain independence from the layout. Weka's data mining system supports numerous common mining responsibilities in statistics, particularly pre possessing information, clustering , classification, regression, visualization, and selection of functions. So, we used classifier for our analysis of the results.

**IV. RESULT & ANALYSIS**

**A. Relation between Max speed and travel distance.**

The Graph of a device of 7 days displays the analyzed data in a graph format. The currently displayed data name is displayed on the graph section. Displayed data can be changed in the statistics section. The Y axis of the graph represents values for the speed (km/h), while the X axis always represents day with total travel distance. We plotted two values one is average speed of a device and max speed of the device. The blue line is for maximum speed and red line for average speed. For 7 days each line has two dots.



Fig 4:- Graph of max speed and travel distance.

In this graph a vehicle traveled 330 km in day one. Its average speed was 4.92 and maximum speed was 51 km/h. In the second day same car traveled 264 km and its max speed was 63.62 which was the maximum speed of seven day. So, in day 3 and day 6 this vehicle traveled less than 100 km and its max speed was 45 and 39. In day 4 this vehicle travels 150 km and max speed was 50 km/h. That means from this graph we can say that if travel distance of this vehicle is more than 100 km than it travels 50 km/h max speed. Using same procedure we can predict maximum speed of a vehicle for any day.

**B. Detecting traffic rules breaker vehicles**

As we mentioned in our methodology part. Our data set is taken from Prohori.com a well-known vehicle tracing company of Bangladesh. They gave us 50 vehicle’s 7 days trip report. So, from this data set we separated the values of max speed and average speed of 7 days and created a new data set. From new data set we have found 350 instances to conduct out data analysis. From this data set those devices max speed is more than or equal of 60 km/h are violated rules. So, we can detect them and can get a list of violated rules devices so that we can give penalty. Of course, this is one of our goal law Enforcement.

Device Number	Day	Max speed	Avg Speed	Class
1	Day 1 MS(km/h)	79	5.65	Overspeed
2	Day 1 MS(km/h)	69.39	2.52	Overspeed
3	Day 1 MS(km/h)	74	4.76	Overspeed
4	Day 1 MS(km/h)	54	1.79	Fine
5	Day 1 MS(km/h)	51	4.92	Fine
6	Day 1 MS(km/h)	55.06	7.44	Fine
7	Day 1 MS(km/h)	0	0	Fine
8	Day 1 MS(km/h)	57	5.36	Fine
9	Day 1 MS(km/h)	0.46	0	Fine
10	Day 1 MS(km/h)	58	4	Fine
11	Day 1 MS(km/h)	3.91	0	Fine
12	Day 1 MS(km/h)	63	4.91	Overspeed
13	Day 1 MS(km/h)	52	6	Fine
14	Day 1 MS(km/h)	56	4.3	Fine
15	Day 1 MS(km/h)	56	1.77	Fine
16	Day 1 MS(km/h)	71	5.93	Overspeed
17	Day 1 MS(km/h)	52	5.81	Fine
18	Day 1 MS(km/h)	59	3.73	Fine
19	Day 1 MS(km/h)	63	6.27	Overspeed
20	Day 1 MS(km/h)	62	6.08	Overspeed
21	Day 1 MS(km/h)	54	4.78	Fine
22	Day 1 MS(km/h)	58	0	Fine
23	Day 1 MS(km/h)	65	7.98	Overspeed

Fig 5:- Data-set of 50 vehicles 7 days max speed and average speed.

In this data set we classify by 2 classes. Those devices are more than or equal to 60 km/h are over speed class and bellow 60 km/h are Fine class. Here in result and analysis part different kind of algorithm will be discussed which we apply to our data set and find some result. Regarding that result which is best for finding over-speed vehicles and by which algorithm we can find those gaps. Here we try to explain those analyses.

➤ **Analysis with KStar Algorithm:**

K star is classification for lazy learning. This approach specifically intended for cluster studies [15]. To find the best violated rules vehicles we can apply KStar algorithm. The class of a test occasion depends on the class of those preparation occurrences like it, as dictated by some similitude work.

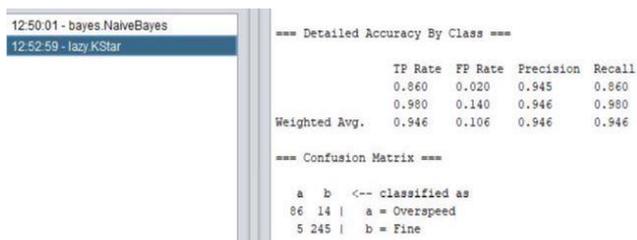


Fig 6:- K-star Algorithm

From K-star algorithm 331 instance or 94. 57% data are correctly classified from 350 vehicles. And only 19 instance or 5.5429% are incorrectly classified. From confusion matrix over speed is denoted by a and Fine is denoted by b. From this algorithm whose cars are violated the rules and found violated is 86 instances and 14 instances are break the 60km/h rules, but this algorithm shows fine. On the other hand, fine vehicles are 250 vehicles. From these vehicles 5 of them are found fine but the algorithm shows these vehicles are violated the rules and 245 instances are find and classify as fine.

➤ *Analysis with J-Rip Algorithm:*

JRip is one of the popular algorithms that get to know the method. JRip continues by considering all the instances

of a decision found in the school records as a class, and having a set of rules that cowl all the individuals in that class.

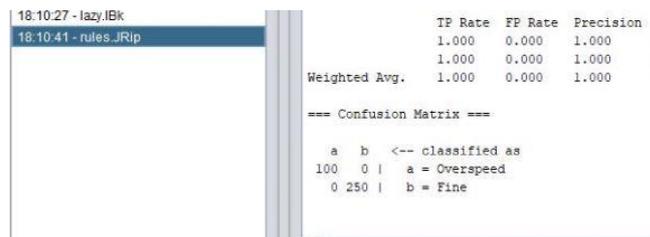


Fig 7:- JRip Algorithm

From J-Rip algorithm all data are classified correctly. From 350 instances 350 instances are correctly classified. From confusion matrix over speed is denoted by a and Fine is denoted by b. From this algorithm whose cars are violated the rules and found violated is 100 instances. On the other hand, fine vehicles are 250 vehicles and these vehicles are found fine. So far it is the best result.

Same as k-star algorithm and jRip algorithm we applied K nearest (IBK) and Random tree algorithm. Here are the comparison.

Algorithm Name	Correctly classified (%)	Incorrectly classified (%)
Kstar	94.574	5.5429
Random tree	99.43	.57
JRip	100	0
K Nearest (IBK)	85.714	13.286

Table 1:- Different Algorithms instances.

Among all the algorithm (Kstar, IBK, Jrip, Random tree) J-Rip algorithm gives us the best result. All the instances are classified correctly in JRip algorithm. So, if we apply j-Rip algorithm to find out the over-speed vehicles and normal speed vehicles we will able to find them correctly.

C. *Prediction For the Next day:*

Prediction for the next day we use another data-set which have 50 vehicles 7 days max speed. So we sum all of this and took a average value. We created a rules so that we can predict the data that if average value is more than or equal to 55 km/h than that vehicle can violate the rules of 60 km/h.

B	C	D	E	F	G	H	I	J	K
Day 1 MS(km/h)	Day 2 MS(km/h)	Day 3 MS(km/h)	Day 4 MS(km/h)	Day 5 MS(km/h)	Day 6 MS(km/h)	Day 7 MS(km/h)	Total of 7 day	Average value	Class
79	0	55.89	81	64	49.56	124	453.45	64.77857143	Overspeed
69.39	51.37	58	0	3	27.08	60	268.84	38.40571429	Normal
74	59.01	54	47	14.85	52.14	0	301	43	Normal
54	50.24	50.21	56	50.74	23.5	69	353.69	50.52714286	Normal
51	62.63	45	60.69	67	39.23	57	382.55	54.65	Normal
55.06	58.15	61	0	81	0	57.69	312.9	44.7	Normal
0	49.28	51.99	56.13	48	47.6	56	309	44.14285714	Normal
57	53.75	53	51.32	57.54	51.73	50	374.34	53.47714286	Normal
0.46	85.14	55	51	50	58.93	61	361.53	51.64714286	Normal
58	57.45	64	77	45	65.31	81	447.76	63.96571429	Overspeed
3.91	53.8	55	55	60	49.78	66	343.49	49.07	Normal
63	60.62	62	53.76	52	48.1	48.36	387.84	55.40571429	Overspeed
52	68.23	57.84	49.78	101	52.34	79	460.19	65.74142857	Overspeed
56	57.52	59.13	48.84	51.8	61.08	70	404.37	57.76714286	Overspeed
56	67.21	62.02	61.15	56	56.99	52	411.37	58.76714286	Overspeed
71	44.8	45	52.86	50	52.99	50	366.65	52.37857143	Normal
52	62.93	67.65	58.06	68	63.34	49	420.98	60.14	Overspeed
59	59.12	76.93	58	68	58.25	47.95	427.25	61.03571429	Overspeed
63	50.63	49.72	54.45	53	55.17	60.39	386.36	55.19428571	Overspeed
62	72.64	51	52.45	30	49.6	49.34	367.03	52.43285714	Normal
54	64.33	64	53.21	52.15	66.06	49.67	403.42	57.63142857	Overspeed
58	52.6	63	52	55.58	77.19	47.04	405.41	57.91571429	Overspeed
65	41.71	54	65.37	87	51.74	51	410.77	58.61	Overspeed

Fig 8:- Data-set of Average Speed of 50 days vehicle

If the value is more than or equal to 55 km/h than it will over-speed else it will be a normal vehicle. We can apply some algorithms to find out which algorithm gives us the most correctly classified instances and less incorrectly classified instances.

Algorithm name	Correctly classified instances	Incorrectly classified instances
Naive Bayes	47	3
KStar	45	5
Random Tree	47	3
Lazy	42	8
JRip	49	1

Table 2:- Algorithm’s classified instances

From this table we can easily pick JRip algorithm as it gives the most Correctly classified instances. So, here we discussed about JRip algorithm.

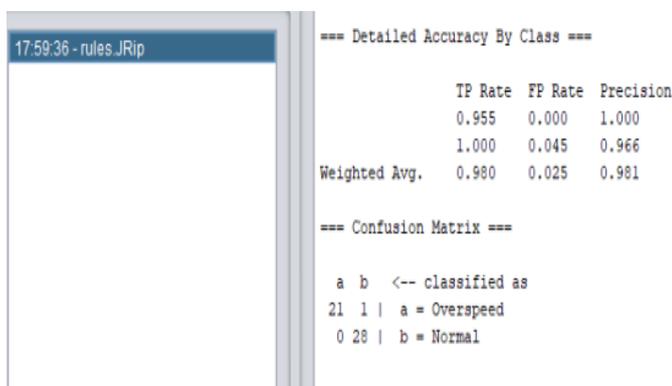


Fig 9:- JRip Algorithm

We can see the total number of instances from this data set result is 50. Properly classified instances are 49% or 98% and incorrectly classified instances are only 1% or 2%. From Confusion Matrix, out of 50 vehicle 22 over-speed devices are detected by JRip algorithm. Among them 21 vehicles are detected as over-speed as they actually violated the 60km/h rules. But this algorithm shows only one vehicle is normal but as this vehicle is violating the rules. On the other hand from 28 normal vehicles this algorithm detects 100% normal means these 28 vehicles are not violating the rules. 22 vehicles from all of them going to break the rules of 60 km/h and 28 vehicles will drive bellow 60 km/h. So, from this learning we can predict with a constant rules (ex: 55kmph/ 60km/h) that if we want to find out which vehicles may break that constant rule for 8<sup>th</sup> day we can do that easily.

Now we are trying to predict the data that is more accurate by doing action prediction against intention. Here we consider precision of different individuals. Precision is utilized with review, the percent of every single important report that is returned by the hunt. Here where precision

rate is high that has been classified the right instances against intention.

In this analysis, we applied JRip algorithm because this algorithm gives us the most correctly classified instances. From this we found some precision against our data set attributes. Here for places precision is given below,

Class	Precision
Over-speed	1.00
Normal	0.966

Table 3:- Precision against JRip Algorithm

We know that if precision rate is high then it can correctly classify the instances. It means for intention and actual action of speed is those vehicles average speed is normal today or bellow from threshold value. Definitely that vehicle’s average speed will be normal for the next day. Here from the table above precision rate of is over-speed is more than normal. That means those 22 vehicles from 50 vehicles are going to break the threshold value.

*D. Gap between engine on time and Idle time:*

Here we need to compare Gap between engine On-time and Idle time. We know that linear regression can be used to compare between two variables. Like when there is one dependent variable and one is independent variable then is can compare them. For that reason, we need numeric value. However, our data set is not numeric. Therefore, at first, we need to convert the data into numeric value for our experiment. Suppose 9-hour 33 minute It is not numeric value, so we write as 9.33 hour. For more specifically A hour is 60-minute means 60 minute is 100% a 30 minute is 50% and so that above 25 minute we coincided as an hour. So here 9-hour 33 minute we wrote as 10 hour. In this data set we have 350 data of 50 vehicle 7 day engine on time idle time and travel distance.

A	B	C	D
Device Number	Travel Distance(km)	Engine On Time (Hour)	Engine Idle Time (Hour)
1	307.23	21	10
2	159.44	22	10
3	213.73	19	9
4	87.07	14	9
5	330.42	23	9
6	157.81	3	0
7	0	0	0
8	378.42	24	9
9	0.01	9	8
10	241.51	22	9
11	0.06	9	9
12	277.72	19	9
13	227.13	16	8
14	237.22	18	8
15	87.7	16	8
16	386.38	24	8
17	372.22	23	8
18	250.43	23	8
19	402.22	23	8
20	374.97	21	7
21	193.46	16	7
22	0	17	7
23	562.52	24	7

Fig 10:- Data set of Engine on time and Idle time.

Now different kind of algorithm will be discussed which we apply to our data set and find some result. Regarding that result which is best for finding the gap and by which algorithm we can find those gaps. Here we try to explain those analyses.

➤ *K star Algorithm:*

```

20:16:36 - lazyKStar
KStar options : -B 20 -M a
Time taken to build model: 0 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient      0.9146
Mean absolute error        0.9899
Root mean squared error    2.1755
Relative absolute error    24.6391 %
Root relative squared error 40.3681 %
Total Number of Instances  350
    
```

Fig 11:- K-star Algorithm

The correlation coefficient is 0.9146 and Mean absolute error is 0.9899 It means that connection between dependent and independent variable is 0.9146 and there each prediction error is 0.9899 it means that it cannot predict that part.

➤ *Random Tree Algorithm:*

In this algorithm the correlation coefficient is 0.8545 and Mean absolute error is 1.1416 It means that connection between dependent and independent variable is 0.8545 and there each prediction error is 1.1416 it means that it cannot predict that part.

```

20:18:03 - trees.RandomTree
=== Cross-validation ===
=== Summary ===

Correlation coefficient      0.8545
Mean absolute error        1.2416
Root mean squared error    2.8835
Relative absolute error    30.9037 %
Root relative squared error 53.5056 %
Total Number of Instances  350
    
```

Fig 12:- Random Tree Algorithm

For, Engine on time and idle time data set among those algorithms we can find gap between dependent and independent variable. All the results are close enough. The results are given bellow.

Algorithm	Correlation coefficient	Mean absolute error
KStar	0.9146	0.9899
M5P	0.9124	0.9899
Random Tree	0.8545	1.1416
IBK	0.8773	0.998

Table 4:- Correlation coefficient and mean absolute error of Algorithms

From the result we can find correlation coefficient of K Star is higher than the other algorithms because correlation coefficient is high. For mean error is same. Therefore, with Lazy K Star algorithm tree we can find more accurate relational gap between them.

If we manage to find out the best algorithm for gap between engine on time and engine idle time. Then we can use results for further implementation. Such as if we managed to find out the longitude and latitude of a vehicle after a certain amount of time continuously than we can find out the traffic jam or the vehicle is in trouble or not by using the gap between engine on time and engine idle time. But for the lack of information of longitude and latitude we could not able to do this analysis.

## V. FUTURE WORK AND CONCLUSION

### A. Future Work:

In this study we try to established law enforcement rules which is one of our major goal. We try to find out the ways how we can detect rules breaker vehicles and predict for the next day that is the same vehicle violated the rules or not. We find out the gap between engine on time and idle time. But for more accurate result we need longitude and latitude than using best algorithm we can find out the traffic jam routes. We can also discover the nearby automobiles using longitude and range. Using average speed of the day and engine idle time we can find out how many hours or minute faced traffic jam. For that purpose, we need some specific location longitude and latitude as we mentioned earlier for example need a longitude and latitude data set 5-minute interval. So, from that data set those cars engine was idle for 10-15-minute mark that route as low traffic area if the vehicle is idle for 20-30-minute mark that route as heavy traffic area. From longitude and latitude, we can find which vehicle are traveled outside the city or which vehicles are used for public transport such as uber. For some user security issue and company rules and regulation tracking companies deny giving vehicles time to time longitude and latitude information.

### B. Conclusion:

Our reason right here is set up a interactive site visitors control machine. We understand that one of the areas of a Smart City in which extensive work can be seen is the Traffic Management System. We endorse a brand new version of Smart Traffic Management System that could interacts with us with the aid of idea the satisfactory course plan, automatic toll collection, a couple of path if any twist of fate occurs or more than one manner if there any traffic jam or caught of a car, Storing cars data and vehicles owners information, if any vehicle violate site visitors regulations or destroy guidelines that automobile can be effortlessly discover by using visitors video tracking device. We can also manipulate site visitors lighting fixtures and routes as we are able to offer first-class routes. If a vehicle should wait in a signal for a long term than opportunity routes are suggested mechanically for other motors. For analysis cause we gather a few vehicle tracing information from Prohori, Onnorokom Electronics. In methodology, we describe approximately our working technique. There we point out approximately our data set. Then by using doing analysis we've got some result. We discover the relation between most pace and total travel distance vehicle. We also discover that are the automobile violated the edge fee of maximum velocity and prediction

for tomorrow which gadgets may additionally ruin the guidelines. Through our research, we had been a hit in generating car list the ones are ruin the rules together with developing predictions. The device additionally deliver idea about activities of accidents and the way they may have an impact at the visitors flow of a town. Levels of precipitation, an prevalence of an twist of fate, the fee of fuel intake, % go with the flow of visitors and use of WEKA are some of the novel functions of our work. In destiny, we want to work extra with this.

## REFERENCES

- [1]. Montella, A., Ariab, M., D'Ambrosio, A., Galantea, F., Mauriello, F., Pernetto, M. "Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections in relation to different perceptual cues," *Accident Analysis and Prevention* 43, (2010): 2072-2084.
- [2]. A. Gregoriades, Nicosia and Cyprus, M. Pampaka. "Design of a 3D Interactive Simulator for Driver Behavior Analysis," *Proc SCSC Summer Computer Simulation Conference*, vol. 17, no. 6, 2013.
- [3]. C, Lu, A. P. Boedihardjo and J. Zheng. "AITVS: Advanced Interactive Traffic Visualization System", *IEEE 22nd International Conference on Data Engineering (ICDE'06)*, 2006.
- [4]. Sourcefire, Real-Time Network Awareness, Nov. 2005, [http:// www.sourcefire.com/products/rna.html](http://www.sourcefire.com/products/rna.html).
- [5]. C. Tominski, P. Schulze-Wollgast, and H. Schumann, "3D Information Visualization for Time Dependent Data on Maps," in *Proc. IV'05*, 2005, pp. 175–181.
- [6]. M. Rath. "Smart Traffic Management System for Traffic Control using Automated Mechanical and Electronic Devices", *IOP Conference Series: Materials Science and Engineering*, 2018.
- [7]. D. Singh, C. Vishnu and C. K. Mohan, "Visual Big Data Analytics for Traffic Monitoring in Smart City," *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Anaheim, CA, 2016, pp. 886- 891.
- [8]. N. Kumar, A. V. Vasilakos and J. J. P. C. Rodrigues, "A Multi-Tenant Cloud-Based DC Nano Grid for Self-Sustained Smart Buildings in Smart Cities," in *IEEE Communications Magazine*, vol. 55, no. 3, pp. 14-21, March 2017.
- [9]. P. Rizwan, K. Suresh and M. R. Babu, "Real-time smart traffic management system for smart cities by using Internet of Things and big data," *2016 International Conference on Emerging Technological Trends (ICETT)*, Kollam, 2016, pp. 1-7.
- [10]. M. V. Moreno et al., "Applicability of Big Data Techniques to Smart Cities Deployments," in *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 800-809, April 2017.
- [11]. A. El Fazziki, D. Benslimane, A. Sadiq, J. Ouarzazi and M. Sadgal, "An Agent Based Traffic Regulation System for the Roadside Air Quality Control," in *IEEE Access*, vol. 5, no., pp. 13192-13201, 2017.

- [12]. H. Piringer, M. Buchetics, R. Benedik, Alvis: Situation awareness in the surveillance of road tunnels, Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on, IEEE, 2012, pp. 153–162.
- [13]. Rajkumar, R.: The New Architecture of Smart Cities. The Urban Technologist (2012).
- [14]. Nakamura, E.F., Loureiro, A.A., Frery, A.C.: Information fusion for wireless sensor networks: methods, models, and classifications. ACM Comput. Surv. 39(3), Article 9 (2007).
- [15]. A. V. Solanki, "Data Mining Techniques Using WEKA classification for Sickle Cell Disease", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4), 2014.
- [16]. Zhang, K., Qu, B. and Chi, X., 2015. Design and Implementation of Browser based GPS/GPRS Vehicle Positioning and Tracking System. MATEC Web of Conferences, 31, p.17002.
- [17]. Kumar Janahan, S., R.M. Veeramanickam, M., Arun, S., Narayanan, K., Anandan, R. and Javed Parvez, S., 2018. IoT based smart traffic signal monitoring system using vehicles counts. International Journal of Engineering & Technology, 7(2.21), p.309.