Diwasi: Intelligent GPS based Accident Prevention System

V.Y. Attigala	D.H. Hettiarachchi	R.S.D.S. Jayasekara	V.N. Vithana
Sri Lanka Institute of Information Technology, Sri Lanka	Sri Lanka Institute of Information Technology,Sri Lanka	Sri Lanka Institute of Information Technology,Sri Lanka	Sri Lanka Institute of Information Technology,Sri Lanka

ABSTRACT

In a demanding environment, accidents are unavoidable disasters. This has caused accidents to fall under the top 10 of the world's death rate. When considering Sri Lanka the accidents rate are in vast amount compared to world. The majority of accidents are caused by unawareness of critical zones rather than human errors. The research is focused on finding patterns within each accident with the aid of data mining techniques. The main objective was to minimize accidents caused by unawareness of the drivers. When considering the accident management systems available in the market, most of them are automobile integrated systems which are quite expensive and can hardly be affordable to ordinary people. In order to prevent this, the research team has created an accident management mobile app which is not integrated to the automobile and inexpensive. The application is built on a Global Position System (GPS) and is based on mobile phone platform considering the cost and the modern trend of using smart phones all over the world. The Data mining model which was built by the Research team, are effective mostly on B Conditions mid-range (20km-50km) roads in Sri Lankan Road conditions. This research on the accident management system mainly focus researchers to study further concerning vehicle independent systems and involve in implementing such inventive systems. Usage of the system will have an impact on considerable reduction of road accidents occurrences, due to the predictions of probable accidents that can happen in critical areas, which previously the driver was unaware of.

General Terms

Data Mining, GPS, Accident, Patterns, Intelligence, Microsoft Clustering Algorithm, Microsoft Decision Tree Algorithm, CART Method

Keywords

Diwasi, Mobile Based Applications, Vehicle Independent, Real time Data mining, 297B Accident Report Application.

1. INTRODUCTION

Data mining and warehousing is a concept use to discover patterns and trends in a large data base such as a data warehouse [1]. The research filed area covers the above mention technology.

The main approach is based on young and fast-growing field of data mining. The system analyses the accident data recorded in Sri Lanka police central database and identify accident pattern, and the system is intelligent for predicting situation, when vehicle reaches the same situation which is as same as the situation predicted by the system, it will be warned by pushing alerts. The development team decides to contribute Global Position System (GPS) based mobile phone

platform considering the cost and the modern trend of using smart phone all over the world.

Accidents are unavoidable incidents which contain many causes. The following statistical data represent how it will influence the world community. These statistical data are taken from World Health Organization [2].

Table 1. Accident Statistical Data

Country	Year
Australia	2010
Pakistan	2010
India	2010
United State of America	2010
Sri Lanka	2010

When consider about Sri Lanka, as a unit cell of the world, also faces a major influence of accidents by considering the following statistical data. The following static, Ministry of Internal Transport represent the criticalness of the situation [3].

Table 2. Analysis Of Data (Year 2013 -2014)

Overview	Total	%
Deaths	24001	6
Fatal Accidents	22368	6
Minor Accidents	127466	34
Critical Accidents	56387	14
Damage Only	173807	45

Considering the above factors, these statistical data represent a criticalness in an accidents and how accidents affect the society.

Further studding about road accident the research team has undergone through related researches and the authors of the article has represent that road accidents can be categories into four groups by their nature [4].

- Equipment Failure
- Roadway Design
- Poor Maintenance
- Driver Behaviors

From the above categories Texas A&M Transportation University has expressed that road design play the major role in accidents [5]. The research team expect to research in the area which is highlighted by Texas A&M Transportation University.

In this documentation the research team presents the core of the research with the aid of literature reviews done by fellow researches statistical analysis and data diagrams. And also the planning and the progress of the research are shown by Gantt chart and work distribution plan.

2. BACKGROUND

The research team deeply analyses the literature reviews of the past researches done by others. The following field has been discovered which is associated with the team's research area.

- Intelligent speed adaptationAccording to the research paper [5], [6], when implemented motor vehicle would not be able to exceed speed limits and more advanced version would also be able to give whether predictions and road conditions to users

- GPS based accident prevention system [7],[8]In ordered to prevent mining accidents happening each year due to equipment collision, research have suggest a system that would sense the proximity of vehicles and warn them about potential collisions.

- Identify the pattern of accident [9],[10],[11]In Korea research has been conducted by incorporating data mining techniques to identify patterns of traffic accidents and to build up classifications of accident severity.

- Embedded system based accident prevention system [12], [13], [14]Research is being done to use Distributed vision-based analysis to detect occupant's posture, and features from multiple cameras to prevent accidents in the home.

The research team expects to invent a system using mobile platform. Which will be affordable for the developing countries like Sri Lanka.

The above sub topics are further analyzed by the research team. The conducted literature review is mention below.

3. METHODOLOGY

The Prototyping Model is a systems development method (SDM) in which a prototype is built, tested, and then reworked as necessary until an acceptable prototype is finally achieved from which the complete system or product can now be developed.

The research team has identified that the requirement are going to be complex. Because new algorithms are to be developed. And there not very certain. Since the research team will need to learn new technologies and language components that are related to centralized web service and database management systems and data warehouse, the academic research reviewing methodology will be more compatible with prototype methodology the research team are going to research in prototype development methodology.

The works to be done in the prototype methodology describe below.

- **Concept**
  In this phase the research team are researching about what research topics need to be done for the comprehensive design and analysis project. The team will gather past research information from the senior students then the team will finalized the research topics, finding the questions and research objectives and finally submit in the project registration form as first milestone.

- **Feasibility Analysis**
  In the feasibility analysis stage the team will consider about technology analysis, resources analysis and stakeholder analysis. In the technology analysis stage the team will analysis deeply about the technology risk. That the research team will face when to in the research. In the resource analysis stage the development team consider about finding internal or external resources which will support the research. In the stakeholder analysis the research team gather information stakeholders and their preferences.

- **Planning**
  In the planning phase the research team mainly consider about planning the architecture, cost management and the implementation planning. In the planning architecture stage the research team will focus about which platform will be used, how many platforms will be used and how to interconnect between platforms. Then planning the data warehouse is the next major task in platform architecture. In cost management section the research team will plan about how managing the budget which will concern about hardware and software implementation. In the implementation planning stage the research team will focus about delivering academic submissions such as prototype presentation, proposal presentation and SRS presentation.

- **Gathering Requirements**
  In the gathering requirement phase the research team focus about collecting database inputs, analyzing the research articles, reading the literature reviews and finding technical resources. The above mention details will help the research team to identify what are the requirement of the research.

- **Analyzing**
  In the analyzing phase the research team will focus about analyzing data mining algorithms which has been done previously and analyze the collected data warehouse to find the data distribution which can widely express about the relation between accident patterns.

- **Designing**
  In the designing phase the research team will more focus about development of the three application, the mobile application, web service application development and the desktop application development. In addition to that pattern recognition system design and test plan designing is also concern. The stage is a recursive stage in the prototype methodology .until its satisfy the objectives.

- **Implementation**
  The research team will focus about academic perspective of product deliveries. In the prototype demonstration the research team will expect to deliver 60% of the working prototype. The implementation phase is separated in to two segments as phase one and the phase two. In phase one the team will deliver versions of the product which will be compatible with prototype demonstration and mid review demonstration. In the phase two the research team focus about hardware implementation and the software implementation and expect to publish the product to the mobile platform market. As a milestone final report submission and final presentation and viva can be considered. The approved prototype design is implemented in this stage

- Testing and Maintaining.In this phase the research team will be conducting the test plan and focus about maintaining the solution and finally analyzing the user feedback for the further use.

After the stage structuring as above the followings has been conducted,

The research team used a specific police division Gampaha, which contains 12 police stations accidents data, to develop the data warehouse for the system. In Sri Lanka Accidents are reported in a format documentation called 297B, which contains detailed descriptions of the accidents and casualties. In this research paper the research team collected the data from the 297B form.

The team received the data in a Microsoft Access Database format. In the first stage, the team converted the Microsoft Access database format into Microsoft SQL server database format using the data import wizard in the SQL server. The team had to create relationships between tables in order to develop the desktop application which will be used to enter the data into the system by the police department. The database contained 7180 tuples of accidents data. The cleaning process of the data warehouse proceeded after the creating of data warehouse relationship. The following views has been implemented after cleaning the database. The views contains data where no null values are included to the data mining algorithm.

- Filtered Basic Details

- Filtered Basic Details with Traffic elements

In the above views Details are included where critical factors are not null. From a completely cleaned database the research can continue to develop few mining models which included CART methods in preview research papers. The following factors were discovered to determine the lowest km post using Microsoft decision tree, this processes involves with the output of the Microsoft clustering algorithm.

- Road no

- Time Accident

- Light Condition

- Day of Week

- Weather

- Urban/Rural

The above is sorted according to the importance in descending order. The above models is stored on a virtual private server .the database server run on a job schedule to process and update the mining model, when the new datasets are inserted into the database .using the above methodology the more inserting data into the system will make the system more intelligent.

- Mapping the Traditional coordinate system into the GPS coordinates.

  In Sri Lanka Accidents are mapped into the documents using the traditional coordinates system which is using the nearest lower km post defined by the road development authority. When implementing a GPS based accident prevention system. The research team used 3 1:500000mm maps to cover Gampaha police divisions which

indicates the km post by km post. The team has to use Google map to manually trace the km post in the map into the GPS coordination and store the results to the database table.in this method mapping the 7180 would be efficient. Due to every accidents are mapped on to a road coordination.

- Mobile Application Navigational and alerting mechanism.[15]

  The main objective is to eliminate the usage of mobile phone while driving and driving away the attention from the mobile application from the driver while driving. The research team uses a mechanism of switching off the display of the mobile application. Using the location change listener in android application.

In the server side development stage, the main objective was to implement the business intelligent solution in a cloud server. The system was implemented on a dedicated virtual private server. The reason to choose the above reason was to run DMX queries which was used to design the CART Mining model (Microsoft Decision Tree).The SQL Server 2014 was used to run the Analysis services, which support DMX queries.

The research team has used a separate Query which was XMALA query which was used to run and train the mining model automatically the algorithm used to train the model.

The next task was to implement the intelligent automation periodically. The research team used a SQL Server tool, SQL Agent which schedule job in the server.in order to get data from mining model the T-SQL server need to access analysis services to transmit data from the mining model.to obtain the goal the research team has used a linked server, which is used to communicate between two or more remote server. The linked server will link the database engine in the SQL server and the analysis server, using the above mechanism the research team was able to get the data from the mining model. The implementation of the linked server was mentioned below

```
<Process
xmlns="http://schemas.microsoft.com/analy
sisservices/2003/engine">
<Object>
<DatabaseID>2 Find Useable Input Decision
Tree</DatabaseID>
</Object>
<Type>ProcessFull</Type>
<WriteBackTableCreation>UseExisting</Writ
eBackTableCreation>
</Process>GO
```

The SQL server needed a specific permission and mechanism to access a remote linked server. The team used a proxy account and the proxy account was given a Credential, then the team created an impersonator account to access the services remotely, in that manner, external application can access the data mining model.

- Mobile Application special security feature

  An application is only shown when the car is stationary. When the car is moving at a speed of over about 8 km/h no picture is shown and No visual media available while driving appears on the application screen, although the audio alert is heard during this time. The picture is

shown again as soon as the car's speed falls below about 8 km/h.

This speed limit is defined by leading auto mobile manufacture called Volvo in Sweden.

4. RESULTS

In Sri Lanka the Roads are categorized into two main categories based on the infrastructure of the road by Road authorization department of Sri Lanka. Therefore the Hypothetical Database selected by the research team for creation of the mining model, which was given by the police Department also contains two categories, "A" condition roads and "B" Condition roads.

The roads in the database was portioned in the following manner.

- "A" Condition Road
  - Length of the road less than 50km
  - Length of the road greater than 100km
- "B" Condition Road
  - Length of the road less than 5km
  - Length of the road greater than 5km & less than 10km
  - Length of the road greater than 10km & less than 20km
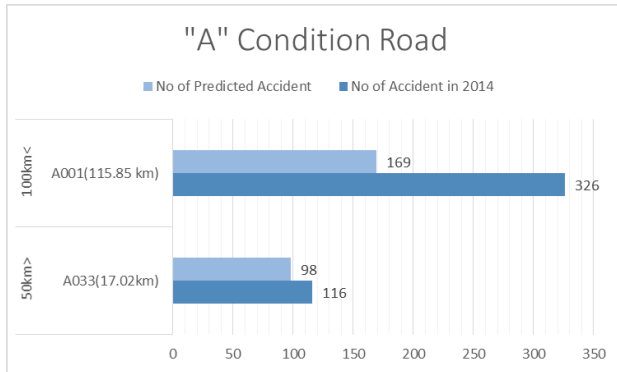 - Length of the road greater than 20km

Table 3. Prediction Results

| Accident Category				No of Accident in 2014	No of Predicted Accident
Class of Accident wise			Fetal	123	88
			Grievous	356	289
			Non - Grievous	625	383
			Damage only	517	397
Day of week wise			Sunday	225	188
			Monday	217	208
			Tuesday	217	193
			Wednesday	234	149
			Thursday	242	137
			Friday	234	133
			Saturday	252	149
Road No wise	A Condition Road	50km >	A033(17.02km)	116	98
		100km <	A001(115.85 km)	326	169
	B Condition Road	5km>	B433 (3.22 km)	2	2
			B601(4.40 km)	22	21
			B108(5.79 km)	9	9
			B179(4.10 km)	7	5
			B226(2.82 km)	2	2
			B161(4.42 km)	17	6
		5km< <10km	B443(8.62 km)	3	2
			B361(6.84 km)	23	7
			B012(8.05 km)	2	2
			B602(8.30 km)	13	1
			B290(8.21 km)	6	6
			B453(9.25 km)	6	2
			B090(8.61 km)	2	2
			B434(7.08 km)	17	0
			B237(8.05 km)	3	3
			B444(7.45 km)	3	3
			B063(5.63 km)	4	3
		10km< <20km	B479(12.87 km)	29	10
			B225(12.87 km)	4	4
			B095(12.07 km)	39	27
			B472(12.87 km)	63	22
			B430(14.88 km)	22	21
			B183(10.22 km)	17	16
			B289(12.63 km)	17	16
			B456(10.86 km)	13	13
			B603(12.00 km)	7	1
			B292(10.06 km)	4	3
			B604(14.60 km)	17	15
			B288(13.27 km)	49	42
			B321(14.67 km)	5	1
		20km<	B214(27.76 km)	63	5
			B129(36.60 km)	1	1
			B146(24.14 km)	116	67
			B111(26.95 km)	90	51
			B445(32.18 km)	69	38
			B208(22.32 km)	91	21

Using the spastically and the graphical analysis of the above data, the research team can show the efficiency of the mining model in accident predictions. Therefore more effective roads conditions for the system is "B" condition roads. But the additions of the databases in the whole country Sri Lanka the mining model of the system can be applied to the A condition Roads.

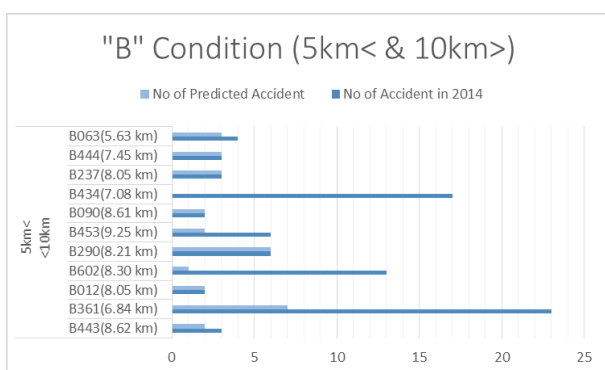1. Statistical Analysis of "A" Condition Road

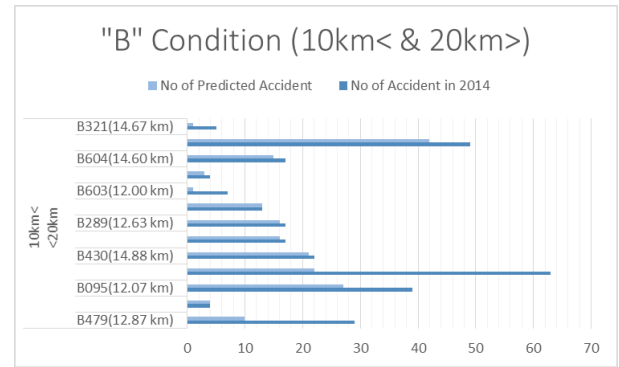2. Statistical Analysis of "B" Condition Road

    a. Length of the road less than 50km
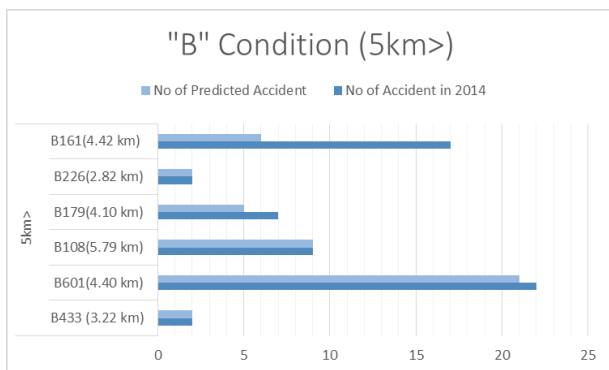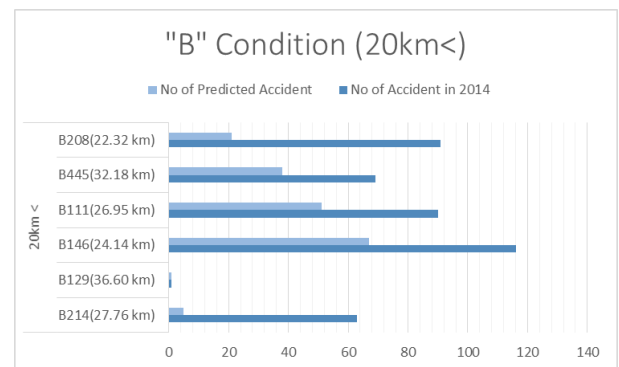

    b. Length of the road greater than 5km & less than 10km

    c. Length of the road greater than 10km & less than 20km

    d. Length of the road greater than 20km

Using the about graphical representation, For the hypothetical location Gampaha District the systems predict with the delicate accuracy in mid-range B conditions road between 1-5 km in length. Compared with the actual database of the 2014 given by the police department, Gampaha district.

The research team used two modes to give the warning message to the user

- Navigational Mode

- Operational Mode

When the mobile application user launch the application he/she has the following option to select the mode when he/she swipe the screen to the left

If the user select the operational mode the map interface will be loaded into the system. Then the system will continuously monitor the GPS Coordinate changes (Location Awareness).the system will continuously communicate with the Server to get the predicted warning zone from the mining model. Then the system will give the user the warning message within the 5 km from the predicted accident zone by a voice command alert
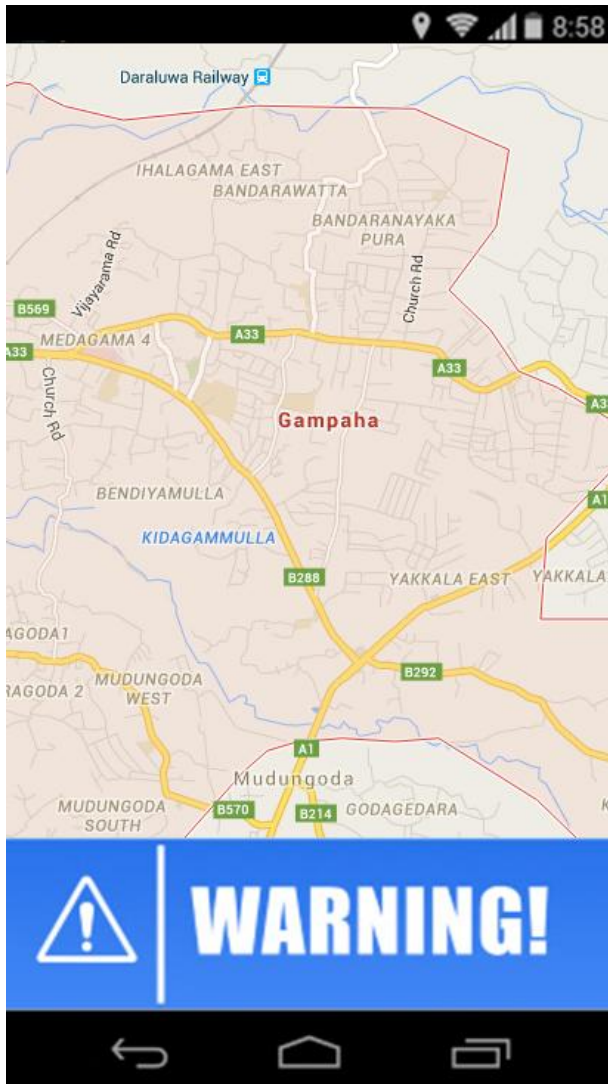
Figure 1: Mobile App Interface, Pushing Alert

5. DISCUSSION

After a period of time, if the accident rate of a critical zone drops the system will try to mark the zone as none critical zone. Therefore the System has to be updated and recreate the model using clustering and regression to keep the system updated.

The research is mainly based on the data mining concept therefore the accuracy of the mining model ,the development team uses are playing a major role ,when considering, giving warning message, to the end users based on the prediction. the results.as mentioned in the above methodology section, the two data mining and regression algorithm Microsoft Clustering algorithm and the Algorithm representing (according to the literature review), CART methodology Microsoft Decision Tree, the application of the above two algorithm and their accuracy is discussed in this section.

The research needed authentic data for the discussion of the accuracy of the system. The research team selected a hypothetical area, Gampaha District. The selection of the hypothetical area are based on the no of accident rate per year, which was taken from a survey conducted by the Traffic police Department of Sri Lanka. The specific database contained accidents data from the year 2002 to 2014 in the Gampaha district.

For the discussion of the prediction model, the research team divided the database into three segments as mentioned below

1. 2002-2013 section-the section is used for selecting most effective factors related to accidents using the Microsoft Clustering Algorithm

2. 2002-2013 section-the section is used for the training of the mining model Decision Tree.

3. 2014 section-the section is used to verify and compare the accuracy of the mining model prediction related to the authentic data emulating the real-time environments.

Using the above segmentation. The research team conducted a analytical comparison between two data segments. The above methodology was used to emulate the prediction of the real time scenario. The results are described in the below table under the following categories. The methodology will ensure that the accuracy of the mining model are in the standards which can be improved using addition of the complete database of Accidents in Sri Lankan police department. The positive results of the accuracy table will prove that the success rate of the system, which can be also say that application of the mobile platform along with the data mining to predict and prevent accidents is possible.

6. CONCLUSIONS

The Diwasi intelligent GPS Based Accident prevention system is implemented to minimize unaware accidents based on the data mining concept and the mobile application with the GPS technology, with the aim of building a vehicle independent economical accident prevention system to the Sri Lankan road condition. The research team implemented the system using three platform architecture as described in the documentation. The research team used a regression and classification algorithm to create the mining model of the system. The system used a separate desktop application to enter the data into the system a specific server is implemented to mine the data periodically. The mobile application was used to give the warning to the user when a crucial accident area is reached. The system implemented with the mining model with the accuracy of 71.55%.the system is implemented using a hypothetical area Gampaha district.to expand the accuracy of the research the research need to be implemented using the databases of the All the police Division in Sri Lanka.

The research on the Accident Management System will be generally impact for researchers to study further regarding the Vehicle independent systems and contain in implementing such inventive systems. Using of the system will impact on a considerable reduction of road accidents occurrences due to the prediction of possible accidents happening in the critical areas which were unaware incidents for the vehicle drivers.

Using the results of mining model on the described environments described in the specification, the research team can come to a conclusion of the accuracy of the mining model under the following categories.

- Class of accidents

- Day of week wise

- Road No wise

Based on the results of the mining model .under the class of the accident, which describe the severity of the accident, the mining model predicted with the accuracy of 71.55% for the

fatal accidents, 81.18% for the grievous accidents. The research team can come to a conclusion for the accidents class that the mining model has a considerable accuracy to predict accidents which involves human life. Further, implementation of the system and distributing to the country will help to reduce accidents which costs human life with the considerable amount of the accuracy in the mining model.

Under the Day of week wise category Department of Motor Traffic in Sri Lanka, can be guided and use the Diwasi system mining model as a reference to plan the daily traffic routines based on the no of accidents involved in the day of the week. Having an accuracy of more than 56% of prediction distribution for the accident prediction for the day of the week will encourage the Traffic department to refer the mining model of Diwasi System as a reliable and real-time updating mining model, which will be updated in every 24 hours of interval.

In the final category, by analyzing the results the research team can specify the class of the road where the mining model is most efficient. the mining model is having a considerable accuracy when the length of the road is between 10 km to 20 km s. the research team can come to a conclusion that for the hypothetical area using the data set, the mining model is success full in the B Condition road between 10 km to 20 km.

The system is based on the mobile platform and the GPS technology, therefore efficiency that the prediction go to the end user depends on the following Technologies

- GPS

- Internet Bandwidth

The research is purely based on the dataset taken from the hypothetical area (Gampaha District) therefore. Accuracy is based on the quality of the data. Based on the quality of the data set the conclusion are subject to change in the research.

The mining model use a regression algorithm to predict the accidents. Therefore if an implementation of the system in a real life occurs, after a long period of time with the benefit of the system, accidents rate in a particular area may become lower in percentages. Therefore the no of accidents enters to the data will be reduced in the long term. As a side effect the mining model prediction can predict a lower possibility of accidents.

In order to address the drop in accuracy of the data mining model the maintaining team has to recreate the mining model based on the new factors affects to the accidents using the procedures described in the research documentation.

For the optimal usage of the system, the research team advises the responsible parties who willing to take the system to the commercial level implementation, to use the proposed Desktop application in each Department of motor traffic in Sri Lanka.in order to address the accuracy drop in period of time the recommended recreation interval of the mining model of the system is 2 years.

For the benefits of the future researchers who are interested in the research topic, the research team can focus the attention of the readers to the following aspects

- Vehicle Independent Accident Prevention Systems using Mobile Platform

- Application of Data mining to the Prevention of Accidents

- Future Studies on the Police Report 297B

The research Diwasi intelligent GPS Based Accident Prevention System showed that for the economical and reliable in real-time accident prevention system for Sri Lanka, would be vehicle independent accident prevention system which will be based on the mobile platform. This encourage the future researchers to develop their research on Vehicle Independent accident prevention system for a developing country.

The research team has showed that using the data mining concept, people can implement accident prevention systems. The system further encourage the Researchers to create system based on Data mining to prevent accidents.

The research team used the 297B Report of Sri Lankan police department for the creation of the data mining model With the accuracy of 71.5%.this encourage the researchers in Sri Lanka to research on 297B Report to improve the efficiency of the Accident Reporting with the aim of categorizing them into factors which can detect and predict a pattern of happening accidents. The ultimate aim would be a preventing an accidents which would save a human life to the world.

7. REFERENCES

[1] "The top 10 causes of death - World Health Organization" [Online]. Available: http://www.who.int/mediacentre/factsheets/fs310/en/. [Accessed: Jan. 23, 2015].

[2] "Road Accidents -The way they are taking place - Minister of internal transport" [Online]. Available: http://www.transport.gov.lk/web/index.php?option=com _content&view=article&id=279&Itemid=171&lang=en. [Accessed: Jan. 23, 2015].

[3] "The Parameters for the road accidents - World Health Organization" [Online]. Available: http://www.who.int/mediacentre/factsheets/fs358/en/. [Accessed: Jan. 23, 2015].

[4] United Nations Human Settlements Program, Enhancing Urban Safety and Security: Global Report on Human Settlements 2007. London, Sterling, VA: UN.

[5] "Data mining of tree-based models to analyse freeway accident frequency,", Jan. 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S00224 37505000708. [Accessed: May. 2,2015].

[6] "FREEWAY INCIDENT FREQUENCY ANALYSIS BASED ON CART METHOD," vnunet.com, Feb, 2013. [Online]. Available: http://www.google.com/US5280555. [Accessed: May. 2,2015].

[7] J. Geralds, "Sega Ends Production of Dreamcast," vnunet.com, para. 2, Jan. 31, 2001. [Online]. Available: http://nl1.vnunet.com/news/1116995. [Accessed: May. 2,2015].

[8] "Older drivers and accidents: A meta analysis and data miningapplication on traffic accident data," Jan. 2005. [Online]. Available: www.sciencedirect.com/science/article/pii/S0957417405 000771. [Accessed: May. 2,2015].

[9] "Mapping patterns of pedestrian fatal accidents in Israel," vnunet.com,, Feb, 2013. [Online]. Available:

http://www.sciencedirect.com/science/article/pii/S00014
57510003994. [Accessed: May. 2,2015].

[10] " Using data mining techniques to road safety improvement in Spanish roads,", Jan. 31, 2001. [Online]. Available: www.sciencedirect.com/science/article/pii/S1877042814 062752. [Accessed: May. 2,2015].

[11] "An approach to accidents modeling based on compounds road environments,", Jan. 2010. [Online]. Available: www.ncbi.nlm.nih.gov/pubmed/23376544. [Accessed: May. 2,2015].

[12] "Prediction of Road Accident Severity Using the Ordered Probity Model," vnunet.com, para. 2, Feb, 2013. [Online]. Available: www.sciencedirect.com/science/article/pii/S2352146514 002701. [Accessed: May. 2,2015].

[13] "Distributed vision-based accident management for assisted living" [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-540-73035-4_21#page-2. [Accessed: Jan. 23, 2015].

[14] IEEExplore, "The RUNES Middleware for Networked Embedded Systems and its Application in a Disaster Management Scenario" [Online]. Available: http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4 144751&url=http%3A%2F%2Fieeexplore.ieee.org%2Fx pls%2Fabs_all.jsp%3Farnumber%3D4144751. [Accessed: Jan. 23, 2015].

[15] "Google Developer, [Online]. Available: https://developers.google.com/maps/. [Accessed: May. 2,2015].Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.