Road Accident Analysis in Kerala and Location Based Severity Level Classification Using Decision Tree Algorithm

Deva Hema Devaraj(Assistant Professor)
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Deepanjali Chandrasekaran
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Bhuvanashree Pandian
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Bharath Syamlal Binitha
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Muli Dipikhasre
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Rahul Anil
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Abstract- Road accident analysis requires contributory factors that affect accidents. The objective is to analyze the accidents in Kerala and to find the severity level of accidents in the places of Kerala. Year wise, day wise and district wise analysis have been done at Kerala. Number of accidents at different road patterns and various contributory factors that affect the accident have been identified from the historical data in Kerala. The severity level of accidents has been classified as severe, Medium and low using decision tree algorithms. The accuracy of the proposed model is 92.5%. The performance of the proposed model is compared with existing algorithms and it outperforms the existing algorithms. The proposed model helps to identify the location that has a higher severity level of accidents in Kerala

Keywords: Road accident, decision tree, contributory factors, severity level.

I. INTRODUCTION

An accident is an unplanned situation that may have inconvenient or undesirable effects. It is an unpleasant event which causes injury or death. There are many traffic accidents that could happen at any time or any place. In India, accident analysis over the years shows that nearly 130,000 people have died on roads. India is constantly being on the top world's death rate list, mainly due to poor road condition and violation of rules by drivers. These are due to the Every year around 1.35 million people die in road accidents, and around 3700 people lose their life on the road. Nearly all road traffic deaths have occurred for unsafe road users like pedestrians, cyclists, and motorcyclists. Road accidents are the dominant cause of death in many countries. To reduce road accidents, it requires strong decisions by government, non-governmental organizations and international companies. At the time of September 2018 the crashes, damages and deaths were around 41,210, 46,150 and 4,700 and in the next 12 months it has been reduced to 31,500, 34,800 and 3,240. The reduction is around 30 percent for the crashes, damages and deaths. Accuracy of classification algorithms helps to find the better results for analysis of the roadway accidents. Taking the instance of Kerala in excess of 50 lakhs vehicles are running on the road. The roadway accidents will be one of the significant reasons for death. The primary goals of this paper are to analyze the contributory factors for road accidents and to classify the severity level of accidents in Kerala.

II. RELATED WORKS

The link between the accidents cases and severity of accidents has been analyzed. The outcome of the analysis shows the cause of the accident, the environmental issues , citizen responsibility, vehicle type, and accident time. [1] Log-linear model, driver characteristics, pedestrian characteristics, road traffic, vehicle typologies have been considered in the project so this gives a clear idea of what is affecting the accidents in the school areas.[2]

It is practically not possible to improve the road safety facilities at all junction points, therefore black spot identification method is analyzed and considered to identify the accident-prone zone. It is an efficient method to decrease road accidents by examines the cause of black spots. The Bayesian network was the best model to effectively identify road accident black spots. Accident occurring spot is known as black spot.[3] Implemented safety measures for highway transportation has more priority in transportation safety production. Proposed method for the estimation of crash rate and relative risk value is based on driver and passenger traits.[4]

A statistical analysis of serious traffic accidents based on the direct cause, accident type, accident responsibility, etc is used to provide support to strengthen the transportation industry management system. Many road safety acts were amended by the government authorities to reduce the severity in lane accidents. The Accident Prediction based on data mining helps us to analyze the different types of circumstantial information with slight complications [5]. Suggestions made for safe driving are based on rules and regulations, classification model, and clusters obtained.[6] Use of an Android Application called Crash prediction system(CPS) which analyzes the user's behavior and predicts whether the user is fit to ride or not.[7] Common problem in the crash analysis data is heterogeneous in nature. During the analysis of the collected data, heterogeneity has to be reviewed as some of the correlated data may be unsolved. Cluster Analysis is a key data mining technique that helps us to achieve the initial task like segregating the different types of road accidents. Regression Analysis is the most acceptable technique in the accident analysis, as the link between the crashes and factors affect them.[8]

The analysis results in the preventive and risk factors of motorcycle accidents. There are different injury patterns during the crashes like the general pattern.(head injuries, Lower-extremity injuries). Road crashes, Street transport frequently speaks to the most serious danger that we are presented to during an ordinary working day.[9]

As indicated by the information from dataset, the death cases normally occur during weekend days between 6:00 a.m. to 6:00 p.m. that includes (HGVs) Heavy Goods Vehicles. Bus accidents happen frequently on weekend days. On weekend days, during early morning and late evening usage of buses are more during the peak hours. A few examinations show that organization vehicles regularly are associated with accidents at high speeds.[10-11]

From the analyzed traffic accident dataset, data mining techniques helps to decrease the death rate. The database used for the road safety purpose helps to drop down the fatality rate by performing road safety events and creating awareness programs throughout the district. The information collected from the database gives the brief description about the accident like road conditions, gender, person involved, health issues and other information, that will be convenient for the evaluation and for fetching additional data for proper conclusion of the case.[12] The International Road Traffic and Accident Database (IRTAD), GLOBESAFE, site for ARC networks are the best assets to gather accident information. For the Investigation purpose, Information are taken from the internet to design and create the map. It helps to characterize the data as well as gives cautioning sound or video. Information from the dataset (Year:1988 to 1993) results describes that the risk of accident for drivers was around 9.5 fatalities per 100 million man hours and it was nearly 3 for other professions.[13]

Some of the classification techniques help to foretell the severity level of lane accidents. Naive Bayes classifier, Decision Tree classifier, AdaBoostM1 Meta classifier, PART Rule classifier, and Random Forest are looked at for characterizing the severity level road accidents. The end-product reveals that Random Forest technique beats the other four models.[14] Study of severity levels in accident helps to scrutinize the link between the severity level, bunch of parameters which include driver details, vehicle details, road pattern and cause of accidents. The impact of different factors on seriousness of the injury like day, time, speed limit, traffic details, climatic conditions, and driver details also used in the investigation.[15–17]

The study of data helped to evaluate the attributes of road accidents seriousness information and approached most common data used for the analysis. The logistic regression model helps to evaluate the autonomous commitment of accident and vehicle characteristics encouraged by the drivers. The variables, for example, driver's age, gender (female), liquor concentration are more prominent than 0.30, drivers without seatbelts, drivers health issues, poor quality of driving, speed limit more than 65 to 70 mph, and more established vehicles were related to greater chances of accident. When the wheel size is increased by 25 cm helps to decrease the death rate.[18]

Few models help to recognize major cause of accident that improves the chances of severity in accidents. The three levels of severity are minor, grievous, or dead injuries. They have been divided into three classes by their severity levels – for example deaths, severe and minor injuries, and they used negative binomial (NB) models on every class of road setbacks independently, bringing about three independent univariate models.[19]

In summary, Year wise, day wise and district wise analysis have been done at Kerala. Numbers of accidents at different road patterns and various contributory factors that affect the accident have been identified from the historical data in Kerala. The severity level of accidents has been classified as severe, Medium and low using decision tree algorithms. The performance of the proposed model outperforms the existing algorithms. The proposed model helps to identify the location that has a higher severity level of accident in Kerala.

III. PROBLEM STATEMENT

There are various problems in real time for the prevention of chance of accidents in the location. Accidents are one of the serious issues faced by people in this modern era. The main reasons for this are the negligence and carelessness of people. Even pedestrians are also facing severe injuries and the present system is not effective. The present road conditions provide a major aid for accidents to occur. Moreover, people are unaware of the speed limit and the accident-prone areas while they are traveling. If they are made aware of these things the accidents and problems can be reduced to a certain limit. The study of the accident analysis at Kerala points out the severity of accidents increasing day by day on road accidents of Kerala and year wise, day wise, district wise analysis and road pattern based analysis. Identifying severity level of locations in Kerala helps to find the accident prone areas for the driver and people.

IV. METHODOLOGY

4.1 Analysis of road accidents in Kerala

Fatal accidents occurred in Kerala include heavy vehicles, Buses, and also multi vehicle accidents. High speed and road pattern conditions increase the seriousness of accidents. This study evaluates the contributory elements liable for the event of accidents at the most frequent locations. In Kerala, the most significant method of transport yet unexpectedly passing from road traffic accidents (RTA) have been currently portrayed in this state as a hidden epidemic which influences all areas of society. It is discovered that, of all absolute reported accidents, 13 percent are death rate and a number of fatal accidents are happening each year in which countless street users lose their lives in every accident. In this paper , Analysis of data has been done based on road patterns, causes of accidents, No of cases at each district of Kerala, gender wise and day/time analysis.

4.2 Classification model

Location based Accident severity level classification model has been developed using decision tree algorithm. The major groups of data or the future data to be predicted can be done by the classification and prediction models. These models are the two types of data analysis model. Classification model is a data preprocessing method which used to predict class label for the given input data. For the training dataset, classification model helps to provide the result from the given input data and helps to foretell the category of the new data.

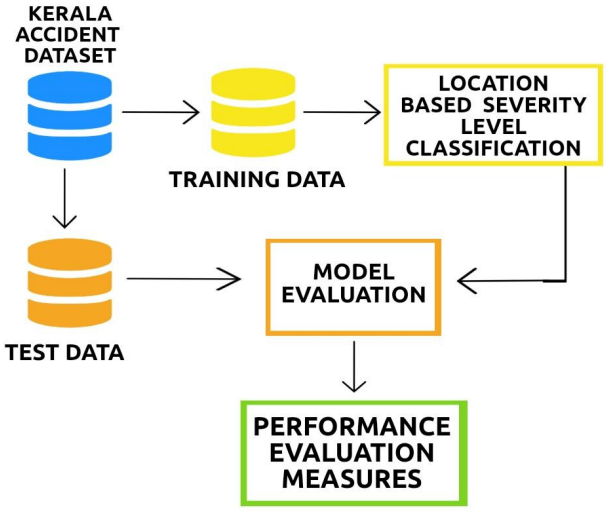

Figure 1. Accident severity level classification framework

4.2.1 Decision Tree Algorithm

A decision tree is a prediction modeling method, which is a decision-supporting tool. It works as a tree-like structure and predicts the accurate values. The process of accident severity level classification has been explained in Fig.1.The decision tree is a non-parametric method of classification model. This method can handle huge volumes of data with good accuracy values. It selects the attribute and according to that it measures the Information Gain, Gain Ratio, Gini index.

It is a supervised learning algorithm of machine learning algorithm. This is a diagrammatic representation of data by which we would obtain the statically probability. Every branch of the decision tree algorithms helps to represent the possible outcome. The tree included in this algorithm will be described by two elements : decision nodes and leaves. This provides a suitable manner to represent these algorithms with conditional control statements. Based on the target elements there are possibilities of two varieties of decision trees namely: categorical variable and continuous variable decision tree. This method is a famous and a simple classification model to explain and follow. The main objective of this model is to generate a training set which helps to discover the target variable value, with the help of decision tree algorithms laws deduced from the original data. For foretelling the label of the class from the given dataset, we have to go ahead from the root of the decision tree and have to differentiate root values from values of dataset parameters. With the help of the differentiation we need to follow the branch of the tree similar to the given value and have to move to the next branch. All the attributes are divided into two types of variables (target variables and feature variables). Target variables are the dependent variable, feature variables are the independent variables. The death, grievous, minor parameters are selected and used in the test data for the classification. The whole dataset have split up into two sets: Training set and Test set. Selection of parameters from the dataset can be done when the given dataset includes K attributes, we have to decide which parameter will accommodate across various levels of branches in decision tree or at root of the decision tree as this process is very much tangled. We can't resolve the problem when we select any node randomly as when we follow random technique this may lead to inadequate results with low precision value. To resolve this issue some researchers gave a few proposals namely: Information Gain or IG, Chi-Square, Gini index (Gini impurity), Gain Ratio (Split Ratio), Entropy, Variance Reduction.

Information Gain: Reduction of Entropy or transformation of the dataset which is frequently used in the training set of the decision tree model. Information Gain is measured by the difference between the entropy of the given dataset before and after the transformation of the dataset. It is defined as in Eq.(1).

$$IG = Entropy\ (parent)\text{-}[Weightedaverage] \times Entropy\ (children) \quad (1)$$

Gini Index: This is the cost function that helps to estimate the splitting of the provided dataset. This is also known as Gini impurity. It is measured by the volume of probability of the particular attribute which is distributed incorrectly during the random selection. The degree varies from 0 to 1. It is defined as in Eq. (2).

$$Gini = 1 - \sum(P(x=k))^2 \quad (2)$$

Gain Ratio: This was found byRoss Quinlan and also known as split ratio. When choosing an attribute from the dataset while taking the number and size of branches into account this helps to reduce the partiality in attributes which have more than one value. While considering theintrinsic information about the split this Split INFO rectifies the information gain. It is defined as in Eq.(3).

$$GainRATIO_{split} = GAIN_{split}/SplitINFO \quad (3)$$

4.2.2 Support Vector Machines

It is a supervised machine learning algorithm that will be useful in classification or regression methods. SVM provides the highest accuracy, used in a variety of applications like face detection, classification of hand-writing, gene detection, etc. SVM also constructs hyperplane to separate the various classes; it also generates optimal hyperplane to minimize the error. Hyperplane is the decision plane which segregates the object with different classes. Different kernels like linear, Radial Basis Function, and polynomial are used in the SVM algorithm, which helps to convert the input data into required form.

4.2.3 Naive Bayes Classifier

This classification method is based on Bayes theorem. This technique helps to build easily, as we have a large volume of data. Naive Bayes algorithm was utilized to create the model which can predict accident seriousness. Two significant advantage of this model are its capacity to readily accommodate fragmented data, or data received at various points as expected, the two of which are characteristic of the incident management process. Although incident duration prediction stays a troublesome and complex issue, the Naive Bayesian classifier is exhibited to give an easier, more adaptable, and more helpful methodology than regression can give, without sacrificing precision. It is defined as in Eq.(4).

$$P(x|y) = P(y|x) \times P(x) / P(y) \quad (4)$$

P(x)- Class Prior Probability
P(y)- Predictor Prior Probability
P(y|x)- Likelihood
P(x|y)- Posterior Probability

Most significant contemplations while picking a calculation is dimensions of the training data (accumulate a decent measure of information to get dependable forecasts). Accuracy and/or Interpretability of the output, Speed or Training time, Linearity, Number of features.

4.3 Data Set

Dataset is the most superior part when you work on an analysis problem. Drafting the data is the first step and the Kerala dataset has been collected from online[20]. The data which has been collected is based on the factors of accidents which help to analyze the problem statement. In this paper, we use accidents occurring at different road types, cause of accidents, gender based accident analysis, severity of the accident cases.

4.4 Data Transformation

All the missing values in the dataset were removed. Every value in the chosen attribute is converted into a proper formal value. Preprocessing of data includes normalization of variables. The min-max normalization is used for normalization which helps to scale the data value between zero (0) and one (1). It is defined as in Eq. (5).

$$y_i = [x_i - min(x)]/[max(x) - min(x)] \quad (5)$$

x_i :i^{th} data point , min: minimum and max: maximum

4.5 Performance Evaluation

Estimation of performance value is based on the confusion matrix. A confusion matrix is in the form of a table which provides the performance of a classification model. This is made on the test data which has the true value.

- Accuracy: The total number of correct predictions divided by the total number of samples is known as accuracy.
- Precision: The correct identification of actual positives of the value is known as precision.
- Recall or True Positive Rate: The number of true positives divided by total number of true positives and false negatives is known as recall.
- F1 Score: The measure that gives the single measure of combined precision and recall is known as F1 Score.
- Misclassification Rate or Error rate: The ratio between total number of misclassified samples and total number of samples known as Error rate.

V. RESULT AND DISCUSSION

Year wise and district wise analysis have been done at Kerala. Number of accidents at different road patterns and various contributory factors that affects the accident has been identified from the historical data in Kerala. The severity level of accident has been classified as severe, Medium and low using decision tree algorithm.

5.1 Data analysis in Kerala

This study says that the road accident problem in highways is due to the number of vehicular traffic, and mixed traffic along with the pedestrian, Loss of life and property occurs due to the traffic accident.

5.1.1 Analysis of road patterns

The Fig.2 shows the accidents occurred in different road types. Ordinary roads comprise all the roads within the city. Every year 3% of accidents are being increased. We must build up various road safety protection systems and enhance the safety control of ordinary roads as they are in worst conditions without proper safety solutions. NH refers to national highways, SH refers to state high way and OR refers to ordinary road.

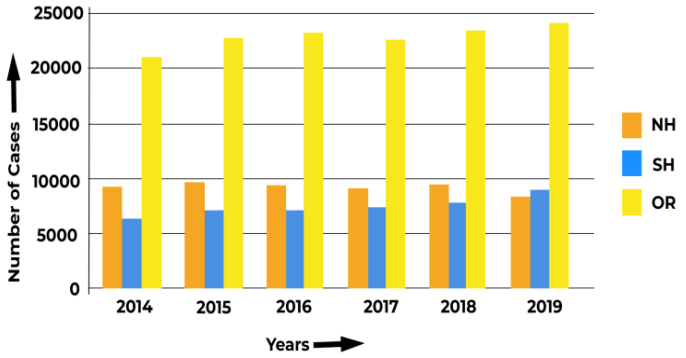

Figure 2. Number of accidents versus Road patterns

5.1.2 Analysis of Accident occurred at different districts

The Fig.3 shows the number of accidents occurred in various districts (14) of Kerala. Malappuram is the most populated district of Kerala, the population is 41,10,956.The second populated district is Trivandrum (33,07,284 people), followed by Ernakulum district with 32,79,860 people. As per the data, the number of accidents happening at Ernakulum stays higher. The second highest number of accidents occurred at Trivandrum.

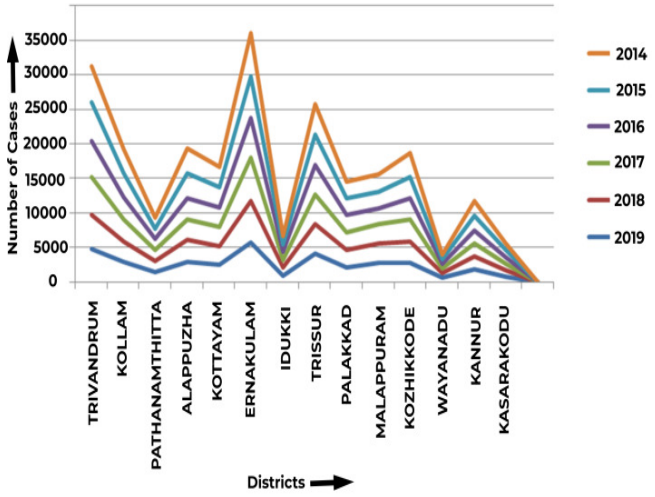

Figure 3. Number of accidents versus Districts at Kerala

5.1.3 Analysis of accidents occurred at day/night accident analysis

The Fig.4 shows the distribution of traffic accidents based on day/night time. Every year accidents have increased, more number of accidents occurred during the day time (peak hours). Night time accidents mainly occurred in hilly areas due to driver's fault or environmental situations. Over-speeding, violation of rules, drivers health issues, drunken driving, poor road patterns are being the major factors of lane accidents during the daytime.

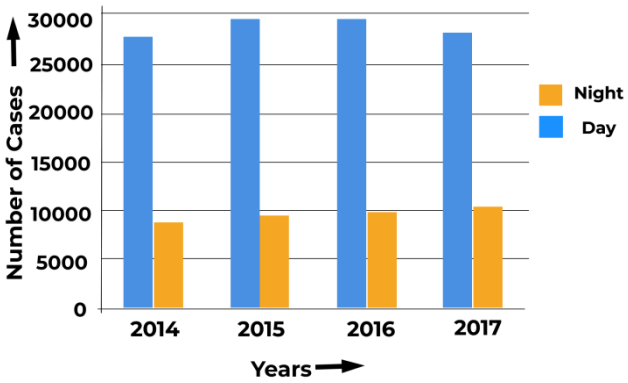

Figure 4. Number of Accidents versus Day/Night

5.1.4 Analysis of cause of accidents

95 percent of the accidents are due to the fault of the drivers. According to the analysis over speed is the major factor for accidents. The accidents are caused by drivers drowsiness, overload, speed, road rules violation, drunk and drive, roads and facilities. Distribution of cause of accidents is described in Fig.5. Other causes of accidents are lack of facilities, vehicle puncture, and mechanical failure and fault of passengers.

5.1.5 Analysis of accidents occurred in Trivandrum district (year 2014-2019)

The Fig.6 shows the number of crashes occurred in the Trivandrum district from the year 2014 to 2019. Every year there is a slight variation in the number of accident cases. According to the analysis number of accidents were higher at the year 2018 when compared to other years

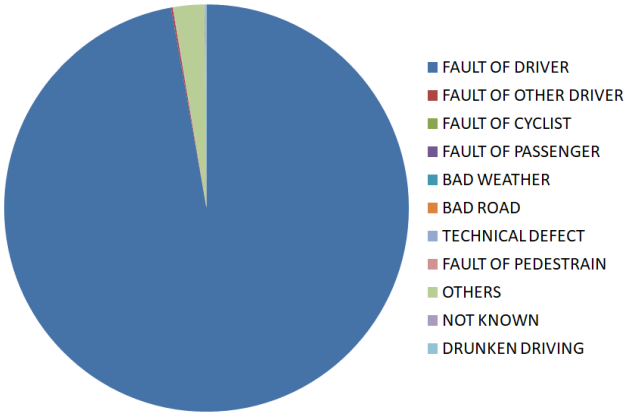

Figure 5. Distribution of causes of accidents

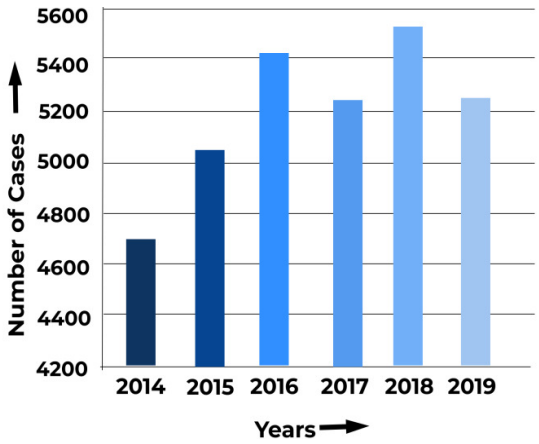

Figure 6. Number of accidents versus years (Trivandrum District)

5.1.6 Severity of Accident cases

The Fig.7 shows the severity level of roadway accidents between the years 2014 to 2019. 0 represents the minor cases, 1 represents the grievous cases, 2 represents the death cases. Identifying the cause of accidents that increase or decrease the risk of severe Injuries in the roadway accidents is one of the primary duties required to develop the safe and proper facilities on Kerala roads.

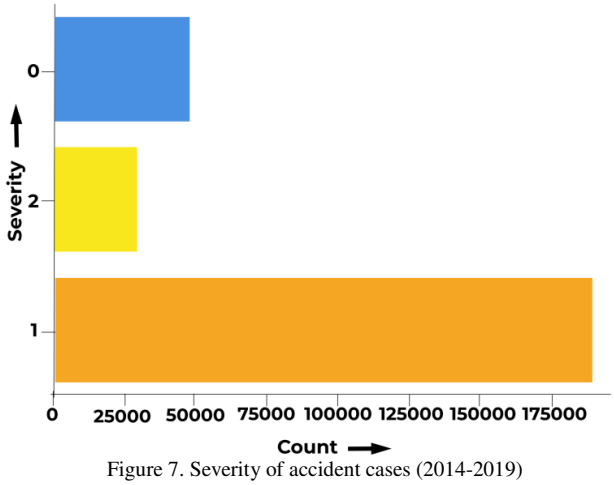

Figure 7. Severity of accident cases (2014-2019)

5.1.6 Severity of Accident cases

The Fig.7 shows the severity level of roadway accidents between the years 2014 to 2019. 0 represents the minor cases, 1 represents the grievous cases, 2 represents the death cases. Identifying the cause of accidents that increase or decrease the risk of severe Injuries in the roadway accidents is one of the primary duties required to develop the safe and proper facilities on Kerala roads

5.2 Severity Level Classification

Delays in detecting and giving consideration to those engaged with roadway accident increment the seriousness of accident cases. The treatment taken for the injuries after the accident has happened, it's an incredibly time sensitive case. Time delay has the effect among life and death. There must be an development in the admittance for care and also there must be proper emergency clinic centers for the post-accidents care, for example, we can achieve with the road safety authority preparing programs. Safe vehicles play a basic part in deflecting accidents and diminishing the probability of serious injury. The Table 1 shows that the decision tree algorithm has the most accuracy and the suitable classification model for the analysis of road accidents at Kerala.

Table 1 Accuracy for different classification models

ALGORITHMS	ACCURACY
Decision tree	**0.925**
Logistic Regression	**0.862**
Naive Bayes	**0.785**
Random Forest	**0.772**
Support Vector Machine	**0.704**

VI. CONCLUSION

In this paper, we have analyzed the causes of accidents and classified the severity of crashes in the 14 districts of Kerala. The Ernakulum and Trivandrum districts have been constantly higher in all the consecutive years from 2014 to 2019. We have considered the different types of road patterns (Ordinary Road, National Highway, and State Highway) and analyzed that ordinary roads have a major number of accidents. Numerous accidents occurred mostly at hilly areas and intersections of roads. The analysis result revealed that the cause of accidents is mainly due to the drivers fault. The fatal accidents don't completely rely on the driver's fault, other cases like vehicle problems, pedestrians or other vehicle issues can also lead to fatal accidents. We have used classification models to find the severity of all accidents. The accuracy of the proposed model is 92.5%. The estimation of the proposed model is made with existing algorithms and it outperforms the existing algorithms. The proposed model helps to identify the location that has a higher severity level of accidents in Kerala.

REFERENCES

[1] X. Xu, S. Yan, W. Yixuan, and M. Lin, "Researching on Traffic Accident Based on Relevance Analysis.," In: *2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*. pp. 629–632. *IEEE* (2019).

[2] Á. Briz-Redón, F. Martínez-Ruiz, and F. Montes, "Estimating the occurrence of traffic accidents near school locations: A case study from Valencia (Spain) including several approaches.," *Accident Analysis & Prevention*. Vol. 132, p. 105237, 2019.

[3] C. Zhang, Y. Shu, and L. Yan, "A Novel Identification Model for Road Traffic Accident Black Spots: A Case Study in Ningbo, China.," *IEEE Access*. Vol. 7, pp. 140197–140205, 2019.

[4] T. Jian, L. Zhi-qiang, X. Jian-feng, and G. Hong-yu, "Cause Analysis and Countermeasures of fatal Traffic Accidents on Road Passenger Transportation Based on Typical Cases.," In: *2019 5th International Conference on Transportation Information and Safety (ICTIS)*. pp. 951–955. *IEEE* (2019).

[5] Y.-K. Ki and D.-Y. Lee, "A Traffic Accident Recording and Reporting Model at Intersections.," *IEEE Transactions on Intelligent Transportation Systems*. Vol. 8, No. 2, pp. 188–194, 2007.

[6] L. Li, S. Shrestha, and G. Hu, "Analysis of road traffic fatal accidents using data mining techniques.," In: *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*. pp. 363–370. *IEEE* (2017).

[7] E. Grant, P.M. Salmon, N.J. Stevens, N. Goode, and G.J. Read, "Back to the future: What do accident causation models tell us about accident prediction?," *Safety Science*. Vol. 104, pp. 99–109, 2018.

[8] S. Kumar and D. Toshniwal, "A data mining framework to analyze road accident data.," *Journal of Big Data*. Vol. 2, No. 1, p. 26, 2015.

[9] T.-O. Nævestad, R.O. Phillips, and B. Elvebakk, "Traffic accidents triggered by drivers at work – A survey and analysis of contributing factors.," *Transportation Research Part F: Traffic Psychology and Behaviour*. Vol. 34, pp. 94–107, 2015.

[10] D. Deva Hema, S. Nandhini, S. VisnuDharsini, and J. ShivaNandhini, "Intelligent Speed Control in Motor Bikes for Accident Prevention Using Internet of Things.," *Journal of Computational and Theoretical Nanoscience*. Vol. 17, No. 4, pp. 1600–1605, 2020.

[11] S. Newnam, B. Watson, and W. Murray, "Factors predicting intentions to speed in a work and personal vehicle.," *Transportation Research Part F: Traffic Psychology and Behaviour*. Vol. 7, No. 4–5, pp. 287–300, 2004.

[12] F. Fan and Y.-B. Qian, "Analysis of Factors Affecting the Severity of Car Accidents at iNtersections Based on Cumulative Logistic Model.," In: *2017 International Conference on Computer Systems, Electronics and Control (ICCSEC)*. pp. 310–313. *IEEE* (2017).

[13] G. Yannis, A. Dragomanovits, A. Laiou, et al., "Use of Accident Prediction Models in Road Safety Management – An International Inquiry.," *Transportation Research Procedia*. Vol. 14, pp. 4257–4266, 2016.

[14] S. Krishnaveni and M. Hemalatha, "A Perspective Analysis of Traffic Accident using Data Mining Techniques.," *International Journal of Computer Applications*. Vol. 23, No. 7, pp. 40–48, 2011.

[15] D. Delen, R. Sharda, and M. Bessonov, "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks.," *Accident Analysis & Prevention*. Vol. 38, No. 3, pp. 434–444, 2006.

[16] H. Manner and L. Wünsch-Ziegler, "Analyzing the severity of accidents on the German Autobahn.," *Accident Analysis & Prevention*. Vol. 57, pp. 40–48, 2013.

[17] P.T. Savolainen, F.L. Mannering, D. Lord, and M.A. Quddus, "The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives.," *Accident Analysis & Prevention*. Vol. 43, No. 5, pp. 1666–1676, 2011.

[18] M. Bédard, G.H. Guyatt, M.J. Stones, and J.P. Hirdes, "The independent contribution of driver, crash, and vehicle characteristics to driver fatalities.," *Accident Analysis & Prevention*. Vol. 34, No. 6, pp. 717–727, 2002.

[19] R.C. Gray, M.A. Quddus, and A. Evans, "Injury severity analysis of accidents involving young male drivers in Great Britain.," *Journal of Safety Research*. Vol. 39, No. 5, pp. 483–495, 2008.

[20] Kerala Government, "Official Webportal for Kerala Police," https://old.keralapolice.gov.in/public-information/crime-statistics/road-accident.