Prediction and Analysis of Injury Severity in Traffic System using Data Mining Techniques

Dheeraj Khera
Research Scholar,
Punjabi University,
Patiala

Williamjeet Singh
Assistant Professor,
Dept of Computer Engineering,
Punjabi University, Patiala

ABSTRACT
Road traffic is an essential part to life, but the repeated road accidents bring severe bodily harm and loss of property. Road Traffic Accidents (RTAs) are considered as major public health concern, resulting in 1.2 million deaths and 50 million injuries worldwide each year as per estimation. The want of study is to scrutinize the performance of different taxonomy methods using WEKA and TANAGRA tool on Traffic Injury Severity Dataset. This paper presents results comparison of three supervised data mining algorithms using various performance criteria. The performance is evaluated by the algorithms Naive bayes, ID3 and Random tree. Comparison of Performance of data mining algorithm based on Error rate, Computing time, precision value and accuracy. The comparison of the model using WEKA experimenter showed that Naive Bayes outperforms Random tree and ID3 algorithms with an accuracy of 50.7%, 45.07 and 25.35% respectively and comparison of the model using TANAGRA experimenter showed that Random tree outperforms Naive Bayes and ID3 algorithms with an accuracy of 92.95%, 67.6% and 57.74% respectively. In the end, we have to conclude that TANAGRA tool is the best data mining tools as compare to the WEKA.

Keywords
Road Traffic Accidents, Data Mining, Naive Bayes, ID3, Random tree, Weka, Tanagra, Accuracy Measure

1. INTRODUCTION
Road traffic is an crucial part to life, but the numerous road accidents carry serious bodily harm and loss of property. Each side of road traffic accidents contains a large amount of information and data is the most common form of the most important information records. Data mining has been defined as the nontrivial extraction of previously unknown, implicit and potentially useful information from data .Via mining the data of road traffic accident, we can analysis accident distinctiveness in multi-angles, multi-level and more comprehensive, and discover potential. It is the science of extracting useful information from large databases. It is one of the responsibilities in the process of knowledge discovery from the database. [5]. There are two primary goals of data mining tend to be prediction and description. Prediction involves some variables or fields in the data set to calculate unknown or future values of other variables of interest. On the other hand Description focuses on finding patterns describing the data that can be interpreted by humans. The endeavor of this study is to investigate the performance of different classification methods using WEKA and TANAGRA focuses on Traffic Injury Severity Dataset. Along with some of the free data mining tools accessible these days, paper deals with the use of the categorization technique can be used. Tools on which classification technique has been implemented are Tanagra and WEKA (Waikato Environment for Knowledge Learning). Choices of classifier used for this purpose are Naïve Bayes, ID3 and Random tree. The paper is classified as follows: Section 2 describes the Literature Survey; Section 3 describes Data Mining Tasks. Section 4 describes Methods & Material; it includes the training data set explanation, supervised learning algorithms, Accuracy measures, Weka and Tanagra. Section 5 gives detailed outcome of the experiment and the proportional results of the tools used. Section 6 gives result analysis and finally Section 7 gives the conclusion and future work.

2. LITERATURE SURVEY
The costs of fatalities and injuries due to traffic accidents have a great impact on society. In recent years, researchers have paid a great attention at determining the factors that significantly affect driver injury severity in traffic accidents. The author in [1] identifies most important factors which affect injury severity by using classification & regression tree. The crash data from the records of the Information and Technology Department of the Iranian Traffic Police from 2006 to 2008 was used to study hundreds of drivers who were involved in traffic crashes on the main two-lane two-way rural roads of Iran. The results indicated that seat belt is the most important factor associated with injury severity of traffic crashes and not using it significantly increases the probability of being injured or killed. The author in [3] presents a random forest & rough set theory to identify the factors significantly influencing single vehicle crash severity. Fifty-nine records of single-vehicle crashes were extracted from the road traffic accident data between January 2004 and May 2008 in Beijing. The results shows that cause factors of Single vehicle crashes are lighting conditions, vehicle type, driving experience, whether wearing the seat belt or not that affect the severity of a SVC are significant factors. The author in [4] presents a decision tree technique which predicts causes of accidents and accident prone locations on highways. Using WEKA software to analyze accident data collected on Lagos-Ibadan road, it was found that decision tree can accurately predict the causes of accident and accident prone locations along the road. The author in [8] predicts Traffic accident duration of incident and driver information system. In this, actual traffic incident data was used to study the prediction problem of traffic incident duration by the method of neural network. 660 sets of actual traffic incident data from a freeway management center were used to train a neural network model, and 170 sets of incident data in the same data collection, which are different from training data, were used to test the prediction effect of the model. The result shows that the incident duration is statically predicted. In practice, as time goes and incident information gradually increases, the prediction result of incident duration should be dynamically updated for improving prediction accuracy. Table1 shows a sample of different data mining techniques used in traffic injury severity.

Table 1 Summary of the Pertinent Literature

Author	Objective	Data Mining Techniques	Accuracy	WORK DONE
Ali et.al (2010)	To identify Most important factors which affect injury severity	Classification & Regression tree	72.49%	Cause factors: seat belt, improper overtaking & speed.
Chaozhong Wu, Ming Ma, Hu Lei, Xinping Yan(2009)	To identify the factors significantly influencing single vehicle crash severity.	Random Forest Roughset theory	0.73% 0.54%	Cause factors: Lighting Conditions, vehicle type, driving experience, wearing belt or not. The efficiency of attribute reduction is not high.
DipoT. Akomolafe, Akinbola Olutayo (2012)	To predict causes of accidents and accident prone locations	Decision tree: Id3, Functional tree	77.70% 70.27%	Decision tree predict causes of accidents and accident prone locations accurately
Liping et.al (2010)	To predict Traffic accident duration of incident and driver information system	Artificial neural Networks	85.35%	The incident duration is predicted in practice, as time goes & incident information gradually increases.
S.Krishnaveni, Dr. M.Hemalatha (2011)	To predict severity of injury using data mining techniques & compare algorithm performance.	Naïve Bayes AdaBoostM1 Meta Classifier Part J48, Random Forest	84.66% 84.64% 84.64% 85.18% 88.25%	Random Forest Outperforms than classification algorithms
S.Shanthi, R.Geetha Ramani (2012)	Significance of data mining classification algorithms in predicting the factors which influence road traffic accident.	Classfication techniques: C4.5,ID3, CS-CRT,CR-T,CS MC4, Naive Bayes, Random forest	99.73%	Random tree classifier using Arc-x4 Meta Classifier outperforms & also improves accuracy.
Tibebe et.al (2013)	To Explore the possible application of data mining technology for developing a classification model	Classification & Regression tree	87.47%	Results shows that developed model could classify accidents with in a reasonable accuracy.

The author in [9] used various data mining techniques. The aim of this paper is to detect the causes of accidents. The dataset for the study contains traffic accident records of the year 2008 produced by the transport department of government of Hong Kong and investigates the performance of Naive Bayes, J48, AdaBoostM1, PART and Random Forest classifiers for predicting classification accuracy. The result shows that Random forest Outperforms than other classification algorithms. The author in [10] talks about the Significance of data mining classification algorithms in predicting the factors which influence road traffic accidents specific to injury severity. It compares the performance of classification algorithms viz. C4.5, CR-T, ID3, CS-CRT, CS-MC4, Naïve Bayes and Random Tree, applied to modeling the injury severity that occurred during road traffic accidents. The author in [11] used to explore the possible application of data mining technology for developing a classification model. This research is focused on developing adaptive regression trees to build a decision support system to handle road traffic accident analysis for Addis Ababa city traffic office. The study focused on injury severity levels resulting from an accident using real data obtained from the Addis Ababa traffic office. The results show that decision tree shows 87.47% accuracy.

3. DATA MINING TASKS

The cycle of data and knowledge mining comprises various analysis steps, each step focusing on a different aspect or task. It proposes the following categorization of data mining tasks [7].

3.1.1 Description and Summarization

At the commencement of each data analysis are the hope and the need to get an overview on the data to see general trends as well as extreme values rather rapidly. Typically, getting the overview will at the same time point the analyst towards particular features, data quality problems and additional required background information. . It is important to broadcast with the data to get an idea what the data might be able to tell you where boundaries will be and which further analyses steps might be suitable. Summary tables, simple univariate descriptive statistics and simple graphics are tremendously valuable tools to achieve this task.

3.1.2 Descriptive Modeling

Descriptive modeling try to find models for the data. The aim of this model is to illustrate not to predict models. Even these models are used in the setting of unconfirmed learning. . In segmentation analysis, the user typically sets the number of groups in advance and tries to partition all cases in homogeneous subgroups. While for some cases it might be difficult to decide to which group they belong, we assume that the resulting groups are clear-cut and carry an inherent meaning Assortments of methods of descriptive modeling are density estimation, smoothing, data segmentation, and clustering. The most broadly used method of clustering is k-means clustering. The interpretation behind cluster analysis is the hypothesis that the data set contains natural clusters which, when discovered, can be characterized and labeled.

3.1.3 Predictive Modeling

The plan of this task is to construct a model that will consent the value of one variable to be predicted from the known values of other variables. Many models are typically built to foresee the behavior of new cases and to broaden the knowledge to matter that are new or not yet as widely understood. The nature of the target variable determines the type of model: classification model, if it is a discrete variable or regression model, if it is a incessant one. Predictive modeling falls into the category of supervised learning; hence, one variable is plainly labeled as target variable and will be explained as a function of the other variables.

3.1.4 Discovering Patterns and Rules

The region of the preceding tasks has been much within the statistical tradition in describing functional relationships between explanatory variables and target variables. There are situations where these functional relationships are tough to accomplish in a consequential way. So Association Rules are a method originating from market basket scrutiny to extract patterns of common behavior.

3.1.5 Retrieving Similar Objects

Having found an appealing article or picture, it is a common desire to find similar objects rapidly. The World Wide Web contains a massive amount of information in electronic journal articles, electronic catalogs and private and commercial homepages. Based on key words and indexed meta-information search engines are providing us with this preferred information. They can not only work on text documents, but to a certain extent also on images.

4. METHODS AND MATERIAL

To classify Traffic Injury Severity data set with high accuracy and efficiency, supervised learning algorithms viz. ID3, Clustering, C4.5 and Naive Bayes are used. In this paper TANAGRA and WEKA data mining tools are used for modeling Traffic Injury Severity data. These are open source data mining software mainly used for academic and research purposes. It proposes several data mining methods from exploratory data analysis, statistical learning, machine learning and database.

4.1 Training Dataset Description

This dataset contains 71 instances and 13 attributes as two are continuous attributes and eleven are discrete attributes. Among these records 13 attributes (Driver Sex, Driver Age, Vehicle Type, Driver Vehicle Relation, Vehicle Insurance, Accident Time, Accident Location, Weather Condition, Road Condition, Road Orientation, Light Condition, Hospital Type, and Accident Severity) affected the accidents. Complete description of variables is shown in table 2.

Table 2. Complete Description of Dataset

S. No	Attribute Name	Data Type	Description
1	Driver Sex	Text	The sex of the driver causing the accident.
2	Driver Age	Number	The Age of the driver causing the accident
3	Vehicle Type	Text	The type of vehicle in the accident
4	Driver Vehicle Relation	Text	The owner of the vehicle (government, Hired Driver, other)
5	Vehicle Insurance	Text	What type of Insurance the Vehicle has? (Third party ,Full Insurance, No)
6	Accident Time	Text	The time of the accident.
7	Accident Location	Text	The location of the accident.
8	Road Orientation	Text	How the road is oriented
9	Road Condition	Text	Is the condition of the surface of the road is dry, muddy or wet.
10	Weather Condition	Text	The weather condition
11	Light Condition	Text	The light condition
12	Hospital Type	Text	The type of hospital (Govt ,Private)
13	Accident Severity	Text	The severity of the accident

4.2 Supervised Learning Algorithms

The supervised learning algorithms used for classifying the Traffic Injure Severity data are as follows:

4.2.1 Naive Bayes

A Naive Bayesian classifier is a simple probabilistic classifier based on applying Bayesian theorem (from Bayesian statistics) with strong (naive) independence assumptions. It is one of the frequently used methods for supervised learning. It provides an efficient way of handling any number of attributes or classes which is purely based on probabilistic theory. Bayesian classification provides practical learning algorithms and prior knowledge on observed data.[6]

- Let **X** be a data sample : class label is unknown
- Let H be a hypothesis that X belongs to class C
- Classification is to determine P(H|**X**), (i.e., posteriori probability): the probability that the hypothesis holds given the observed data sample **X**
- P(H) (prior probability): the initial probability
- P(**X**): probability that sample data is observed
- P(**X**|H) (likelihood): the probability of observing the sample **X**, given that the hypothesis holds
- training data **X**, posteriori probability of a hypothesis H, P(H|**X**), follows the Bayesian theorem we can write

$$p(C \mid F1....Fn) = \frac{p(C)p(F1.....Fn \mid C)}{p(F1.....Fn)}$$

4.2.2 ID3

ID3 is a simple decision learning algorithm, developed by J.Ross Qunilan. It accepts only categorical data for building a model. The basic idea of ID3 is to construct a decision tree by employing a top down greedy search through the given sets of training data to test each attribute at every node. It uses statistical property known as information gain to select which attribute to test at each node in the tree. Information gain measures how well a given attribute separates the training samples according to their classification.[6]

4.2.3 Random Tree

Random Tree is a supervised Classifier developed by Brieman. It is an ensemble learning algorithm that generates many individual learners. It employs a bagging idea to produce a random set of data for constructing a decision tree. In standard tree each node is split using the best split among all variables. In a random forest, each node is split using the best among the subset of predicators randomly chosen at that node. [6]

4.3 Accuracy Measures

Accuracy measure represents how far the set of tuples are being classified correctly.TP refers to positive tuples and TN refers to negative tuples classified by the basic classifiers. Similarly FP refers to positive tuples and FN refers to negative tuples which is being incorrectly classified by the classifiers. The accuracy measures used here are precision and recall. The precision is the measure of accuracy provided that a specific class has been predicted. Recall is the percentage of positive labeled instances that were predicted as positive [13]. A distinguished confusion matrix was obtained to calculate sensitivity, specificity and accuracy. Confusion matrix is a matrix representation of the classification results (Table 3).

Table 3. Confusion Matrix

	Predicted Positive	**Predicted Negative**
Predicted Positive	TP	FN
Predicted Negative	FP	TN

Precision=TP/TP+FP

Recall=TP/TP+FN

Accuracy=TP+TN/TP+TN+FP+FN

4.4 Weka

Weka is a Waikato Environment for Knowledge Analysis. Weka is a collection of machine learning algorithms for data mining tasks and well suited for developing new machine learning schemes. Weka is java based software capable of working under various operating systems. These algorithms can either be applied directly to a dataset or can be called from your own java code. Weka is probably the most successful open source data mining software which has inspired by the development of other programs with more sophisticated graphical user interface and better visualization methods [2][7]. In Weka datasets should be formatted to the ARFF format. The Weka Explorer will use these automatically if it does not recognize a given file as an ARFF file the Preprocess panel has facilities for importing data from a database, a CSV file, etc. and for preprocessing this data using a filtering algorithm. These filters can be used to transform the data and make it possible to delete instances and attributes according to specific criteria.

4.5 Tanagra

Tanagra is open source data mining software used for academic and research purposes. It proposes various data mining methods like exploratory data analysis, statistical learning and for databases area also. The main purpose of the Tanagra is to give researchers and students easy to use data mining software. It is also used to propose to researchers an architecture allowing them to easily add their own data mining methods, to compare their performances. This software has free access to source code. In this way, Tanagra can be considered as a pedagogical tool for learning programming techniques [12].

4.6 Comparison

Data mining is the process that helps to make use of the data in various databases and find new patterns in it. Table 4 depicts the result chart of the data mining tool comparison is given below.

Table 4. Comparison of WEKA and Tanagra

	WEKA	**TANAGRA**
Platform Supported	Cross Platform	Windows
Mode of Software	Open Source	Open Source
Partioning of dataset into training and testing data	have with limited partioning abilities	have with limited partioning abilities
Descriptor Scaling	Fail(can not save parameters for scaling to apply to future datasets)	Fail(can not save parameters for scaling to apply to future datasets)

Descriptor Selection	have the facility but not part of Knowledge flow	have wrapper methods that is valid only for logistic regression
Applications	Machine Learning	Machine Learning, Multivariate Analysis

5. EXPERIMENTAL RESULTS

The experimental results of basic classifiers are discussed in this section using the data mining tools Tanagra and Weka. Traffic Injury Severity data contains Accident Severity which represents the severity of the Accident. The four kinds of accident severity are Injury, Property loss, Slightly Injury and Fatality. The data source for this research consumption is collected from various people that have been suffered from accident by answering the questionnaire. The data stored is in Excel file format. It stores partial road accident records of year 2008-12 that occurred in the city. To classify them correctly from the training data set the error rates and accuracy using classifiers are evaluated. In this study, all data is considered as instances and features in the data are known as attributes. The simulation results are partitioned into several sub items for easier analysis and evaluation. Different performance matrix like TP rate, FP rate, and Precision, Recall, F-measure and ROC area are presented in numeric value during training and testing phase. The summary of those results by running the techniques in WEKA is reported in Table 5. Then we run the algorithms in Tanagra and results are reported through Tables 6. Finally Table 7 shows the Accuracy in percentage and Table 8 shows the Error rates for different classification Comparing two tools.

Table 5. Performance measuring in training and test data set using WEKA

Classifier name	Accuracy	Error Rate	Time (seconds)
Naive Bayes	50.7	0.4089	0.02
ID3	25.35	0.592	0
Random tree	45.07	0.506	0.02

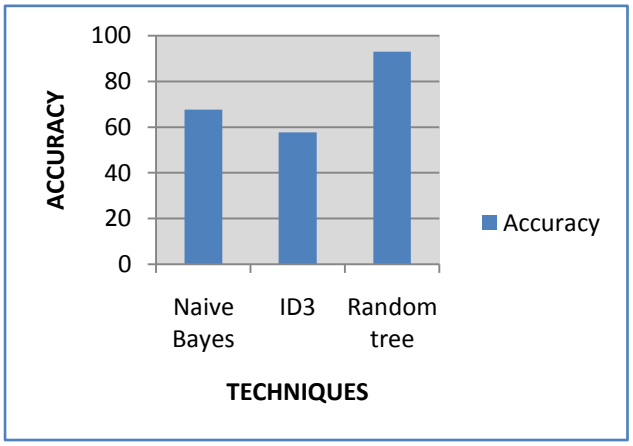

Fig 1. Performance measuring in training and test data using WEKA

From Fig 1. it could be observed that the Naive bayes classifier has greater accuracy when compared with other Basic classifiers in predicting injury severity. The error rates and accuracies of each classifier are listed in Table 3. Among them Naive bayes sounds better with 0.4089 error rate and 50.7% accuracy. From this result we can infer that out of all classifiers Naive bayes suits best for predicting injury severity.

Table 6. Performance measuring in training and test data set using TANAGRA

Classifier name	Accuracy	Error Rate	Time (seconds)
Naive Bayes	67.6	0.3239	0.01
ID3	57.74	0.4225	0.04
Random tree	92.95	0.0704	0.03

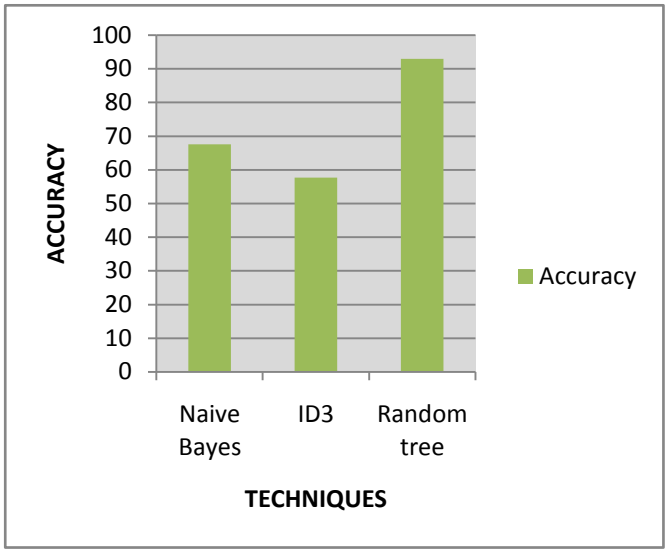

Fig 2. Performance measuring in training and test data using TANAGRA

From Fig 2. it could be observed that the Random tree classifier has greater accuracy when compared with other Basic classifiers in predicting injury severity. The error rates and accuracies of each classifier are listed in Table 4. Among them Random tree sounds better with 0.0704 error rate and 92.95% accuracy. From this result we can infer that out of all classifiers Random tree suits best for predicting injury severity.

Table 7. Accuracy in percentage for different classification comparing two tools

Tool	Naive Bayes	ID3	Random Tree
WEKA	50.7	25.35	45.07
TANAGRA	67.6	57.74	92.95

Table 8. Error rates for different classification comparing tools

Tool	Naive Bayes	ID3	Random Tree
WEKA	0.4089	0.592	0.506
TANAGRA	0.3239	0.4225	0.0704

6. ANALYSIS

In this study, we examine the performance of different classification methods that could generate accuracy and some error to diagnosis the data set. To predict Accident Severity, various Classification models are built using Naive Bayes, ID3 and Random Tree. Decision trees are easy to build and understand can manage both continuous and categorical variables and can perform classification as well as regression.

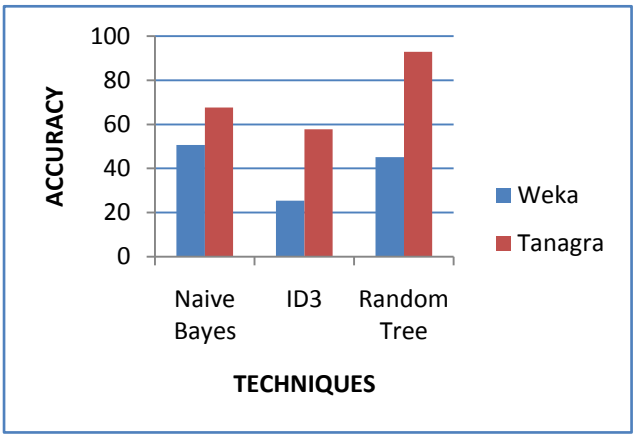

Fig 3. Compare Accuracy among Weka and Tanagra tools

Fig 3. shows the comparison of accuracy among Weka and Tanagra tools. The statistics shows that having a means of predicting likely accuracy of different techniques base on some input values. It is evident from the line graph that value of Naive Bayes in Tanagra tool is less than the Weka Tool. We can clearly see the highest accuracy is 92.95% belongs to Random Tree and lowest accuracy is 45.07%. The total time required to build the model is also a crucial parameter in comparing the classification algorithm. Classification models are generated on the basis of the training data whose independent variables and target variables are known, to be applied for the new dataset whose objective is the prediction of the target variable.

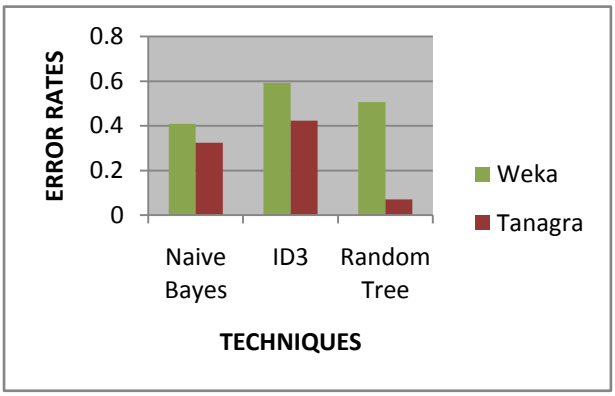

Fig 4. Compare Error Rate among Weka and Tanagra tools

Based on Fig 4, we can compare errors among different techniques in Weka and Tanagra. We clearly find out that Random tree is the best in Tanagra rather than weka, second best is the naive bayes in Tanagra rather than weka. An algorithm which has a lower error rate will be preferred as it has more powerful classification capability. It further reveals that Random tree in Tanagra tool is best comparatively other classifiers cause 92.95% accuracy and error rate is 0.0704 that also lowest error rate as compare to other classifiers.

7. CONCLUSION

The endeavor of this paper is to spot the causes of accidents. There are different data mining categorization algorithm that can be used for driver and road factors for car accidents and identify buried patterns in the accident data set. To accomplish these goals: The WEKA and TANAGRA data mining tools have been used to employ the Naive Bayes , ID3 and Random tree algorithms. The data source for this research expenditure is collected from various people that have been suffered from accident by answering the questionnaire. It stores fractional road accident records of year 2008-12 that occurred in the city. The total accident dataset obtained is 71. Model is also evaluated using accuracy and recall. The assessment of the model using WEKA experimenter showed that Naive Bayes outperforms Random tree and ID3 algorithms with an accuracy of 50.7%, 45.07 and 25.35% respectively and comparison of the model using TANAGRA experimenter showed that Random tree outperforms Naive Bayes and ID3 algorithms with an accuracy of 92.95%, 67.6% and 57.74% respectively. In the end, we have to conclude that TANAGRA tool is the greatest data mining tool in this research as compare to the WEKA. In future, we have to enlarge the classification accuracy of road traffic accidents types; data quality has to be improved. Another future effort is to test the applicability of other data mining techniques.

8. REFERENCES

[1] Ali et al.," A Data Mining Approach to identify key factors of traffic injury severity" Traffic & Transportation, Vol. 23, 2011, No. 1, 11-17.

[2] Bouckaert Remco, Eibe Frank, Mark Hall, Richard Kirkby, Peter Reutemann, and Alex Seewald, 2008.WEKA Manual for Version 3-6-0. University of Waikato, New Zealand.

[3] Brijesh Kumar Baradwaj, Saurabh Pal," Mining Educational Data to Analyze Students Performance" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011

[4] Chaozhong et.al.,"Severity Analyses of Single-Vehicle Crashes Based on Rough Set Theory" 2009 International Conference on Computational Intelligence and Natural Computing.

[5] DipoT.Akomolafe, Akinbola Olutayo," Using Data Mining Technique to Predict Cause of Accident and Accident Prone Locations on Highways" American Journal of Database Theory and Application 2012, 1(3): 26-38.

[6] Dissanayake, S., Lu, J.J." Factors influential in making an injury severity difference to older drivers involved in fixed object - passenger car crashes", Accident analysis and prevention, Vol. 34, No. 5, 2002, pp. 609-618.

[7] Elma kolce (cela), Neki Frasheri, "A Literature Review of Data Mining Techniques used in Healthcare Databases", ICT Innovations 2012 Web Proceedings - Poster Session.

[8] Han, Jiawei and Kamber, Micheline. (2006). Data Mining: concepts and Techniques. San Fransisco; Morgan kufman Publishers

[9] http://en.wikipedia.org/wiki/weka (machine learning)/ accessed on May 2014.

[10] Liping et al.," Traffic Incident Duration Prediction Based on Artificial Neural Network" 2010 International Conference on Intelligent Computation Technology and Automation.

[11] S.Krishnaveni, Dr. M.Hemalatha," A Perspective Analysis of Traffic Accident using Data Mining Techniques", International Journal of Computer Applications (0975 – 8887) Volume 23– No.7, June 2011.

[12] S.Shanthi, R.Geetha Ramani " Feature Relevance Analysis and Classification of Road Traffic Accident Data through Data Mining Techniques" Proceedings of the World Congress on Engineering and Computer Science 2012 Vol. I WCECS 2012, October 24-26, 2012, San Francisco, USA

[13] Tibebe Shah, Shawndra Hill (2013)," Mining Road Traffic Accident Data to Improve Safety: Role of Road-related Factors on Accident Severity in Ethiopia".

[14] Tanagra – a Free Data Mining Software for Teaching and Research, Available at: http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html.

[15] Trilok Chand Sharma, Manoj Jain,"WEKA Approach for Comparative Study of Classification Algorithm", International Journal of Advanced Research in computer and Communication Engineering Vol. 2, Issue 4, April 2013.

9. AUTHOR'S PROFILE

Dheeraj Khera completed the B-Tech. degree in Computer Science Engineering in 2011 Punjab Technical University, Jalandhar (Punjab). Presently he is working as Lecturer in Bhai Gurdas Global Polytechnic College, Rakhra and pursuing his M-Tech. in Computer Engineering Department, Punjabi University, Patiala (Punjab).

Williamjeet Singh received his B.Tech, M.Tech degree in Computer Science and Engineering in 2005 and 2007 respectively from Punjab Technical University, Jalandhar (Punjab) and Punjabi University, Patiala (Punjab). Presently he is working as Assistant Professor and pursuing his PhD from Computer Engineering Department, Punjabi University, Patiala. His current research interests are primarily in the area of Wireless networks, Algorithms and Data mining**.**