Using Data Mining Technique to Predict Cause of Accident and Accident Prone Locations on Highways

Dipo T. Akomolafe¹,*, Akinbola Olutayo²

¹Dept. of Mathematical Sciences, Ondo State University of Science and Technology, Okitipupa, Nigeria
²Dept. of Computer Science, Joseph Ayo Babalola University, Ikeji Arakeji, Osun State

Abstract

Road accident is a special case of trauma that constitutes a major cause of disability, untimely death and loss of loved ones as well as family bread winners. Therefore, predicting the likelihood of road accident on highways with particular emphasis on Lagos – Ibadan express road, Nigeria in order to prevent accident is very important. Various attempts had been made to identify the cause(s) of accidents on highways using different techniques and systems and to reduce accident on the roads but the rate of accident keep on increasing. In this study, the various techniques used to analyse the causes of accidents along this route and the effects of accidents were examined. A technique of using data mining tool to predict the likely occurrence of accident on highways, the likely cause of the accident and accident prone locations was proposed using Lagos – Ibadan highway as a case study. WEKA software was used to analyse accident data gathered along this road. The results showed that causes of accidents, specific time/condition that could trigger accident and accident prone areas could be effectively identified.

Keywords

Data Mining, Decision Tree, Accident, WEKA, Data Modelling, Id3 Algorithm, Id3 Tree, Functional Tree Algorithm

1. Introduction

Road accident is a special case of trauma that constitutes a major cause of disability and untimely death. It has been estimated that over 300,000 persons die and 10 to 15 million persons are injured every year in road accidents throughout the world. Statistics have also shown that mortality in road accidents is very high among young adults that constitute the major part of the work force. In actual fact, accidents kill faster than AIDS and it gives no preparatory time to its victims. In order to combat this problem, various road safety strategies have been proposed and used. These methods mainly involve conscious planning, design and operations on roads. One important feature of this method is the identification and treatment of accident prone locations commonly called black spots; black spots are not the only cause of accidents on the highway. Also various organizations such as Police High Way Patrol, Vehicle Inspection Officer (VIO), Federal Road Safety Commission (FRSC) among others are charged with the responsibility of maintaining safety thereby reducing road accidents. However, lack of good forecasting techniques has been a major hindrance to these organizations in achieving their objectives.

* Corresponding author:
dtakomolafe@yahoo.com (Dipo T. Akomolafe)
Published online at http://journal.sapub.org/database
Copyright © 2012 Scientific & Academic Publishing. All Rights Reserved

It is against this background that Decision Tree is being proposed to model data from road accident database to determine causes of accidents and accident prone locations using historical data collected from Ibadan-Lagos express road as reference point.

2. Objective

The primary objective of this research is to use data mining technique; decision tree to predict causes of accident and accident prone locations on highways using data collected on Lagos – Ibadan express way.

3. Methods

3.1. Data Mining

Data Mining is an interactive process of discovering valid and novel, useful and understandable patterns or models in large database (Han, Mannila and Smyth, 2001). Data Mining, according to Han, Mannila and Smyth (2001) is a process that uses a variety of data analysis tools to discover patterns and relationships in data that may be used to make a valid prediction. Data mining uses advances in the field of Artificial Intelligence (AI) and Statistical techniques. Therefore, decision tree is being used in this research

3.2. Decision Trees
Decision Trees have emerged as a powerful technique for modelling general input / output relationships. They are tree-shaped structures that represent a series of roles that lead to sets of decisions. They generate rules for the classification of a dataset and a logical model represented as a binary (two-way split) tree that shows how the value of a target variable can be predicted by using the values of a set predictor variables. Decision trees, which are considered in a regression analysis problem, are called regression trees. Thus, the decision tree represents a logic model of regularities of the researched phenomenon.

3.3. Accidents along Lagos - Ibadan Express Way

Lagos to Ibadan Express road is one of the busiest roads in Africa. This is because Lagos was the capital of Nigeria until the seat of government moved to the Federal Capital Territory Abuja and also the headquarters of many national institutions while Ibadan is said to be the largest city in black Africa. The traffic along this route is very heavy because it is a gateway linkage of the heavy traffic going from the Northern, Eastern and Majority of Western states. Fig 3.1 shows the frequency of accidents between the distances of 1 and 40km from Ibadan to Lagos between January 2002 and December 2003. The statistics shows that having a means of predicting likely location of accident base on some input values is essential to advice on dangerous locations.

Table 3.1. Record of Accidents along Lagos Ibadan between year 2002 and 2003

| S/NO | Month       | No of Accident |
|------|-------------|----------------|
| 1    | Jan 2002    | 6              |
| 2    | Feb 2002    | 11             |
| 3    | March 2002  | 10             |
| 4    | April 2002  | 18             |
| 5    | May 2002    | 14             |
| 6    | June 2002   | 4              |
| 7    | July 2002   | 6              |
| 8    | August 2002 | 1              |
| 9    | September 2002 | 9          |
| 10   | October 2002| 6              |
| 11   | Nov. 2002   | 4              |
| 12   | December 2002| 5            |
| 13   | Jan 2002    | 5              |
| 14   | Feb 2003    | 5              |
| 15   | March 2003  | 4              |
| 16   | April 2003  | 7              |
| 17   | May 2003    | 2              |
| 18   | June 2003   | 1              |
| 19   | July 2003   | 4              |
| 20   | August 2003 | 5              |
| 21   | September 2003 | 8         |
| 22   | October 2003| 5              |
| 23   | Nov. 2003   | 5              |
| 24   | December 2003| 6           |

3.4. Process of Data Mining

The process of data mining consists of three steps which are:

3.4.1. Data Preparation

This includes; Data collection, Data cleaning and Data transformation.

3.4.2. Data Modeling

This research considers the data of accident record between the first 40km from Ibadan to Lagos. The data were organized into a relational database.

The unknown causes in Table 3.2 may include other factors such as Law enforcement agent problems, attitude of...
other road users, inadequate traffic road signs, traffic congestion and general vehicle conditions.

The sample data used covered the period of 24 Months, that is, January 2002 to December 2003 as indicated in Fig. 3.1.

The output variable is the location and the locations can be divided into three distinct regions tagged regions A, B and C, meaning we have three outputs. Where

First location 1 – 10km is Region A or location 1, Above10km – 20km is region B or Location 2 and above 20km is region C or Location 3.

The data sample used covered a period of twenty four Months starting from January 2002 to December 2003. The data were collected by Akomolafe (2004) and this is presented in Table.3.3.

3.4.3. Deployment

In this stage, new sets are applied to the model selected in the previous stage to generate predictions or estimates of the expected outcome.

Table 3.2. showing variables given both continuous and categorical values

| S/N | Variable    | Description          | Value | Type    |
|-----|-------------|----------------------|-------|---------|
| 1.  | Vehicle Type| Small cars           | 1     | categorical |
|     |             | Heavy Vehicle        | 2     | categorical |
| 2.  | Time of the day| Morning              | 1     | Categorical |
|     |             | Afternoon            | 2     | Categorical |
|     |             | Evening              | 3     | Categorical |
|     |             | Night / Midnight     | 4     | Categorical |
| 3.  | Season      | Wet                  | 1     | Categorical |
|     |             | Dry                  | 2     | Categorical |
| 4.  | Causes      | Wrong Overtaking     | A     | Categorical |
|     |             | Careless Driving     | B     | Categorical |
|     |             | Loss of Control      | C     | Categorical |
|     |             | Tyre Burst           | D     | Categorical |
|     |             | Over Speeding        | E     | Categorical |
|     |             | Obstruction          | F     | Categorical |
|     |             | Pushed by another vehicle | G | Categorical |
|     |             | Broken Shaft         | H     | Categorical |
|     |             | Broken Spring        | I     | Categorical |
|     |             | Brake Failure        | J     | Categorical |
|     |             | Road problem         | K     | Categorical |
|     |             | Unknown Causes       | L     | Categorical |
|     |             | Robbery Attack       | M     | Categorical |

Table 3.3. Sample Data collected from FRSC (Akomolafe O.P 2004)

| SNO | DATE      | TYPE | TIME | SEASON | CAUSE       | LOCATION | REG. NO |
|-----|-----------|------|------|--------|-------------|----------|---------|
| 1   | 6.1.2002  | 2    | 2    | 1      | 2           | 31       | XG 506 LND |
| 2   | 7.1.2002  | 2    | 1    | 1      | 1           | 14       | XC 720 ACD |
| 3   | 11.1.2002 | 1    | 1    | 1      | 1           | 14       | AM 713 LND |
| 4   | 12.1.2002 | 2    | 1    | 1      | 2           | 27       | XE 905 JJJ |
| 5   | 19.1.2002 | 1    | 2    | 1      | 3           | 27       | AA 559 LAF |
| 6   | 30.01.02  | 3    | 3    | 1      | 2           | 12       | AA 156 NWD |
| 7   | 03.02.02  | 2    | 2    | 1      | 2           | 35       | XF 635 JJJ |
| 8   | 05.02.02  | 2    | 1    | 1      | 2           | 10       | XE 141 AKD |
| 9   | 05.02.02  | 2    | 3    | 1      | 2           | 14       | XE 124 AKD |
| 10  | 06.02.02  | 2    | 3    | 1      | 2           | 31       | XE 124 AKD |
| 11  | 11.02.02  | 1    | 1    | 1      | 3           | 5        | AG 276 LAR |
| 12  | 14.02.02  | 1    | 1    | 1      | 2           | 14       |        |
| 13  | 18.02.02  | 2    | 1    | 1      | 2           | 18       |        |
| 14  | 21.02.02  | 2    | 1    | 1      | 2           | 19       | XD 249 SMK |
| 15  | 21.02.02  | 3    | 2    | 1      | 2           | 19       | XC 361 KTU |
| 16  | 24.02.02  | 2    | 1    | 1      | 2           | 18       | XE 716 SMK |
| 17  | 27.02.02  | 2    | 3    | 1      | 2           | 35       | XC 307 SGM |
| 18  | 03.03.02  | 2    | 1    | 1      | 2           | 16       | XE 807 NSR |
| 19  | 05.03.02  | 1    | 2    | 1      | 2           | 10       | XE 348 AKP |
| 20  | 07.03.02  | 2    | 1    | 1      | 2           | 2        | OY 2270 JB |
| 21  | 07.03.02  | 2    | 1    | 1      | 2           | 13       | AP 820 LSD |
| 22  | 07.03.02  | 3    | 2    | 1      | 2           | 18       | XE 322 APP |
| 23  | 19.03.02  | 2    | 2    | 1      | 2           | 19       | XC 993 AGL |
|   |    | 3 | 1 | 2 | 2 | LA 1804 RF |
|---|----|---|---|---|---|----------------|
| 24 | 19.03.02 |   | 1 | 2 | 14 | AM 343 FST |
| 25 | 30.03.02 |   | 1 | 2 | 14 | KC 461 ABA |
| 26 | 31.03.02 |   | 1 | 2 | 14 | BS 142 KJA |
| 27 | 31.03.02 |   | 1 | 2 | 22 | AA 807 EGB |
| 28 | 01.04.02 |   | 1 | 2 | 22 | BX 527 GGE |
| 29 | 01.04.02 |   | 1 | 2 | 30 | AG 787 GNN |
| 30 | 01.04.02 |   | 1 | 2 | 35 | AU 725 MAP |
| 31 | 02.04.02 |   | 1 | 2 | 27 | XG 399 APP |
| 32 | 02.04.02 |   | 1 | 2 | 15 | CY 65 EKY |
| 33 | 04.04.02 |   | 1 | 2 | 17 | AJ 21 AGG |
| 34 | 04.04.02 |   | 1 | 2 | 6  | AW 45 FST |
| 35 | 05.04.02 |   | 1 | 2 | 30 | XB 855 AKD |
| 36 | 06.04.02 |   | 1 | 2 | 13 | AL 567 YAB |
| 37 | 07.04.02 |   | 1 | 2 | 12.5 | XA 787 WWP |
| 38 | 09.04.02 |   | 1 | 2 | 11 | XB 791 GNN |
| 39 | 13.04.02 |   | 1 | 2 | 11 | XA 127 AFN |
| 40 | 13.04.02 |   | 1 | 2 | 11 | AH 202 AKN |
| 41 | 13.04.02 |   | 1 | 2 | 15 | RA 01 KRJ |
| 42 | 22.04.02 |   | 1 | 2 | 11 | BB 731 KJA |
| 43 | 22.04.02 |   | 1 | 2 | 27 | AU 739 JJ |
| 44 | 27.04.02 |   | 1 | 2 | 14 | AE 316 FST |
| 45 | 28.04.02 |   | 1 | 2 | 12 | AZ 824 AAA |
| 46 | 03.04.02 |   | 1 | 2 | 20 | AA 654 GBY |
| 47 | 5.8.2002   |   | 1 | 2 | 30 | XF 65 JJ |
| 48 | 5.8.2002   |   | 1 | 2 | 35 | DM 207 AAA |
| 49 | 5.10.2002 |   | 1 | 2 | 35 | BL 86 AAA |
| 50 | 5.10.2002 |   | 1 | 2 | 35 | BR 608 LSR |
| 51 | 5.11.2002 |   | 1 | 2 | 26 | XB 606 APP |
| 52 | 5.13.2002 |   | 1 | 2 | 26 | XA 616 YLW |
| 53 | 5.13.2002 |   | 1 | 2 | 26 | BM 566 GGE |
| 54 | 5.14.2002 |   | 1 | 2 | 15 | XC 348 AKD |
| 55 | 5.15.2002 |   | 1 | 2 | 19 | OY 2077 JB |
| 56 | 5.15.2002 |   | 1 | 2 | 14 | AJ 101 NND |
| 57 | 5.20.2002 |   | 1 | 2 | 26 | AU 682 ABC |
| 58 | 5.21.2002 |   | 1 | 2 | 24 | XG 719 FST |
| 59 | 5.25.2002 |   | 1 | 2 | 12 | AV 70 LSR |
| 60 | 6.2.2002  |   | 1 | 2 | 12 | AZ 191 MUS |
| 61 | 6.3.2002  |   | 1 | 2 | 16 | AQ 742 YYY |
| 62 | 6.15.2002 |   | 1 | 2 | 12 | XA 682 YRE |
| 63 | 6.16.2002 |   | 1 | 2 | 21 | AL 885 AKN |
| 64 | 6.16.2002 |   | 1 | 2 | 21 | XE 751 SMK |
| 65 | 7.15.2002 |   | 1 | 2 | 12 | XB 266 KNK |
| 66 | 7.19.2002 |   | 1 | 2 | 12 | AE 232 SM |
| 67 | 7.20.2002 |   | 1 | 2 | 12 | XA 940 KHI |
| 68 | 8.8.2002  |   | 1 | 2 | 4  | AX 94 JJ |
| 69 | 9.20.2002 |   | 1 | 2 | 7  | XC 768 BDJ |
| 70 | 9.20.2002 |   | 1 | 2 | 29 | BL 254 SMK |
| 71 | 9.21.2002 |   | 1 | 2 | 16 | AP 647 AKR |
| 72 | 9.21.2002 |   | 1 | 2 | 18 | NC 253 GGE |
| 73 | 9.21.2002 |   | 1 | 2 | 10 | LA 979 BG |
| 74 | 9.22.2002 |   | 1 | 2 | 16 | XU 510 GGE |
| 75 | 9.22.2002 |   | 1 | 2 | 12 | AA 05 MHA |
| 76 | 9.27.2002 |   | 1 | 2 | 13 | AE 869 MUS |
| 77 | 10.1.2002 |   | 1 | 2 | 15 | XB 888 AKR |
| 78 | 10.2.2002 |   | 1 | 2 | 7  | XD 168 BDJ |
| 79 | 10.2.2002 |   | 1 | 2 | 6  | AA 342 LES |
| 80 | 10.2.2002 |   | 1 | 2 | 5  | BX 877 KJA |
| 81 | 10.2.2002 |   | 1 | 2 | 12 | XC 637 RKJ |
| 82 | 11.10.2002|   | 1 | 2 | 11 | XC 937 SMK |
| 83 | 11.10.2002|   | 1 | 2 | 12 | AA 466 KNK |
| No. | Date       | Code | Code | Code | Code | Code     |
|-----|------------|------|------|------|------|----------|
| 87  | 2.12.2004  | 2    | 1    | 1    | 2    | 14       |
| 88  | 12.7.2002  | 3    | 2    | 1    | 2    | 1        |
| 89  | 12.10.2002 | 2    | 3    | 1    | 3    | 13       |
| 90  | 12.11.2002 | 2    | 2    | 1    | 2    | 16       |
| 91  | 12.12.2002 | 1    | 1    | 1    | 2    | 14       |
| 92  | 23.01.2002 | 1    | 3    | 1    | 1    | 16       |
| 93  | 18.01.03   | 1    | 3    | 1    | 1    | 18       |
| 94  | 27.01.03   | 2    | 2    | 1    | 2    | 8        |
| 95  | 29.01.03   | 3    | 4    | 1    | 2    | 12       |
| 96  | 29.01.03   | 2    | 1    | 1    | 2    | 14       |
| 97  | 02.02.03   | 1    | 1    | 1    | 2    | 18       |
| 98  | 12.02.03   | 1    | 2    | 1    | 1    | 18       |
| 99  | 12.02.03   | 2    | 2    | 1    | 2    | 18       |
| 100 | 12.02.03   | 1    | 3    | 1    | 1    | 12       |
| 101 | 17.02.03   | 2    | 3    | 1    | 2    | 11       |
| 102 | 05.03.03   | 1    | 2    | 1    | 2    | 6        |
| 103 | 19.03.03   | 2    | 1    | 1    | 2    | 12       |
| 104 | 28.03.03   | 3    | 1    | 1    | 1    | 12       |
| 105 | 31.03.03   | 2    | 3    | 1    | 2    | 13       |
| 106 | 05.04.03   | 2    | 2    | 2    | 3    | 11.5     |
| 107 | 06.04.03   | 1    | 1    | 2    | 3    | 12       |
| 108 | 06.04.03   | 1    | 1    | 2    | 2    | 12       |
| 109 | 14.04.03   | 1    | 1    | 2    | 2    | 28       |
| 110 | 24.04.03   | 1    | 2    | 2    | 2    | 7        |
| 111 | 24.04.03   | 3    | 2    | 2    | 2    | 9        |
| 112 | 30.04.03   | 3    | 3    | 2    | 1    | 16       |
| 113 | 10.05.03   | 1    | 2    | 40    |
| 114 | 16.05.03   | 1    | 3    | 2    | 20    |
| 115 | 02.06.03   | 1    | 1    | 2    | 1    | 8        |
| 116 | 20.07.03   | 2    | 1    | 2    | 2    | 27       |
| 117 | 26.07.03   | 1    | 2    | 2    | 2    | 9        |
| 118 | 28.07.03   | 2    | 2    | 2    | 2    | 13       |
| 119 | 28.07.03   | 2    | 2    | 2    | 1    | 18       |
| 120 | 02.08.03   | 1    | 1    | 2    | 2    | 13       |
| 121 | 02.08.03   | 1    | 1    | 2    | 1    | 8        |
| 122 | 09.08.03   | 1    | 1    | 2    | 1    | 19       |
| 123 | 16.08.03   | 2    | 2    | 2    | 2    | 2        |
| 124 | 31.08.03   | 1    | 1    | 2    | 1    | 14       |
| 125 | 01.09.03   | 3    | 2    | 2    | 1    | 8        |
| 126 | 08.09.03   | 1    | 2    | 17    |
| 127 | 14.09.03   | 2    |      |      |      |          |
| 128 | 16.09.03   | 1    | 2    | 2    | 2    | 6        |
| 129 | 21.09.03   | 2    | 1    | 2    | 2    | 31       |
| 130 | 24.09.03   | 2    | 2    | 2    | 1    | 18       |
| 131 | 28.09.03   | 2    | 1    | 2    | 2    | 14       |
| 132 | 28.09.03   | 2    | 3    | 2    | 2    | 13       |
| 133 | 06.10.03   | 1    | 2    | 2    | 2    | 11       |
| 134 | 14.10.03   | 2    | 2    | 2    | 2    | 12       |
| 135 | 18.10.03   | 2    | 3    | 2    | 2    | 28       |
| 136 | 19.10.03   | 1    | 2    | 2    | 2    | 22       |
| 137 | 20.10.03   | 2    | 2    | 2    | 2    | 27       |
| 138 | 01.11.03   | 3    | 1    | 1    | 1    | 9        |
| 139 | 02.11.03   | 2    | 2    | 1    | 2    | 18       |
| 140 | 25.11.03   | 1    | 1    | 1    | 3    | 24       |
| 141 | 27.11.03   | 1    | 1    | 1    | 2    | 18       |
| 142 | 27.11.03   | 2    | 2    | 1    | 2    | 13       |
| 143 | 06.12.03   | 2    | 1    | 1    | 2    | 13       |
| 144 | 09.12.03   | 3    | 3    | 1    | 1    | 13       |
| 145 | 13.12.03   | 2    | 1    | 1    | 1    | 7        |
| 146 | 22.12.03   | 1    | 1    | 1    | 1    | 11       |
| 147 | 24.12.03   | 1    | 1    | 3    | 12    |
| 148 | 24.12.03   | 2    | 2    | 1    | 2    | 13       |
4. Results

4.1. Analysis

The major step required to obtain result of the research was carried out by analysing the data using WEKA. WEKA is a collection of machine learning algorithms and data processing tools. It contains various tools for data pre-processing, classification, regression, clustering, association rules and visualization. There are many learning algorithms implemented in WEKA including Bayesian classifier, Trees, Rules, Functions, Lazy classifiers and miscellaneous classifiers. The algorithms can be applied directly to a data set. WEKA is also data mining software developed in JAVA it has a GUI chooser from which any one of the four major WEKA applications can be selected. For the purpose of this study, the Explorer application was used.

The Explorer window of WEKA has six tabs. The first tab is pre-process that enables the formatted data to be loaded into WEKA environment. Once the data has been loaded, the preprocess panel shows a variety of information as shown in figure 4.3 below.

![WEKA Explorer GUI Chooser](image1.png)

![WEKA Explorer](image2.png)

Figure 4.1. WEKA GUI chooser

Figure 4.2. WEKA Explorer
4.1. Weka Classifiers

There are several classifiers available in WEKA but Function Tree and Id3 were used in this study in case of Decision Tree. Prism Rule based learner was generated using WEKA. Attribute importance analysis was carried out to rank the attribute by significance using information gain. Finally, correlation based feature subset selection (cfs) and consistency subset selection (COE) filter algorithm were used to rank and select the attribute that are most useful. The F-measure and the AUC which are well known measures of probability tree learning was used as evaluation metrics for model generated by WEKA classifiers.

Several numbers of setups of decision tree algorithms have been experimented and the best result obtained is reported as the data set. Each class was trained with entropy of fit measure, the prior class probabilities parameter was set to equal, the stopping option for pruning was misclassification error, the minimum n per node was set to 5, the fraction of objects was 0.05, the maximum number of nodes was 100, surrogates was 5, 10 fold cross-validation was used, and generated comprehensive results.

The best decision tree result was obtained with Id3 with 115 correctly classified instances and 33 incorrectly classified instances which represents 77.70% and 22.29% respectively.

Mean absolute error was 0.1835 and Root mean squared error was 0.3029.

The tree and rules generated with Id3 algorithm are given thus:

4.2. Id3 Tree

```
TYREBURST = TRUE
  | SEASON = WET
  |   | TYPE = HEAVY VEHICLE
  |   |   | TIME = EVENING: LOCATION2
  |   |   | TIME = AFTERNOON: LOCATION2
  |   |   | TIME = MORNING: LOCATION2
  |   |   | TIME = NIGHT: null
  |   | TYPE = SMALL CAR: LOCATION2
  |   | TYPE = MOTORCYCLE: null
  | SEASON = DRY
  |   | TIME = EVENING
  |   |   | TYPE = HEAVY VEHICLE: LOCATION2
  |   |   | TYPE = SMALL CAR: LOCATION3
  |   |   | TYPE = MOTORCYCLE: null
  |   | TIME = AFTERNOON
  |   |   | TYPE = HEAVY VEHICLE: LOCATION2
  |   |   | TYPE = SMALL CAR: LOCATION2
  |   |   | TYPE = MOTORCYCLE: null
  |   | TIME = MORNING
  |   |   | TYPE = HEAVY VEHICLE: LOCATION3
  |   |   | TYPE = SMALL CAR: LOCATION3
  |   |   | TIME = NIGHT: null
  |   | TYREBURST = FALSE
  |   | TIME = EVENING
```

```
OVERSPEEDING = FALSE: LOCATION2
  | OVERSPEEDING = TRUE
  |   | TYPE = HEAVY VEHICLE: LOCATION2
  |   | TYPE = SMALL CAR: LOCATION2
  |   | TYPE = MOTORCYCLE: null
  | TIME = AFTERNOON
  | LOSS-OF-CONTROL = FALSE
  | OVERSPEEDING = FALSE
  | BRAKE-FAILURE = FALSE
  |   | TYPE = HEAVY VEHICLE
  |   |   | WRONG-OVERTAKING = FALSE
  |   | BROKEN-SHAFT = FALSE: LOCATION1
  | BROKEN-SHAFT = TRUE: LOCATION3
  | WRONG-OVERTAKING = TRUE: LOCATION2
  |   | TYPE = SMALL CAR
  |   | SEASON = WET: LOCATION3
  |   | SEASON = DRY:
  |   | CARELESSDRIVING = FALSE: LOCATION3
  |   | CARELESSDRIVING = TRUE: LOCATION2
  |   | TYPE = MOTORCYCLE: LOCATION3
  |   | BRAKE-FAILURE = TRUE
  |   | TYPE = HEAVY VEHICLE: LOCATION1
  |   | TYPE = SMALL CAR: LOCATION1
  |   | TYPE = MOTORCYCLE: LOCATION2
  |   | OVERSPEEDING = TRUE
  |   | TYPE = HEAVY VEHICLE: LOCATION2
  |   | TYPE = SMALL CAR
  |   | SEASON = WET: LOCATION2
  |   | SEASON = DRY: LOCATION1
  |   | TYPE = MOTORCYCLE: LOCATION1
  | TIME = MORNING
  | SEASON = WET
  | OVERSPEEDING = FALSE
  | TYPE = HEAVY VEHICLE
  |   | WRONG-OVERTAKING = FALSE
  |   | CARELESSDRIVING = FALSE: LOCATION1
  |   | CARELESSDRIVING = TRUE: LOCATION2
  | WRONG-OVERTAKING = TRUE: LOCATION1
  | TYPE = SMALL CAR
  | CARELESSDRIVING = FALSE
  | LOSS-OF-CONTROL = FALSE: LOCATION3
  | LOSS-OF-CONTROL = TRUE: LOCATION3
```
ATION2  
|       |       |       |       | CARELESSDRIVING = TRUE: LOCATION3
TION1  
|       |       |       | TYPE = MOTOCYCLE: LOCATION2
|       | OVERSPEEDING = TRUE: LOCATION2
| SEASON = DRY
|       | BROKEN-SHAFT = FALSE
|       | TYPE = HEAVY VEHICLE
|       | CARELESSDRIVING = FALSE
|       | LOSS-OF-CONTROL = FALSE
|       | BROKEN-SPRING = FALSE
|       | OVERSPEEDING = FALSE: LOCATION2

Rule 3 If TREE-OBSTRUCTION = TRUE
and TIME = EVENING then LOCATION3

Rule 4 If TYREBURST = TRUE
and TIME = MORNING
and TYPE = SMALL CAR
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PULLED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION3

Rule 5 If TYPE = MOTOCYCLE
and CARELESSDRIVING = TRUE then LOCATION3

Rule 6 If ROAD-PROBLEM = TRUE
and TYPE = SMALL CAR
and TIME = AFTERNOON
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and TYREBURST = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PULLED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION3

Rule 7 If TYREBURST = TRUE
and SEASON = DRY
and TIME = MORNING
and TYPE = HEAVY VEHICLE
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PULLED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION3

Rule 8 If UNKNOWN-CAUSES = TRUE
Rule 9 If \(\text{TYREBURST} = \text{TRUE}\) 
\(\text{and TYPE} = \text{HAEVY VEHICLE}\) 
\(\text{and TIME} = \text{AFTERNOON}\) 
\(\text{and SEASON} = \text{DRY}\) then LOCATION3

Rule 10 If \(\text{TIME} = \text{MORNING}\) 
\(\text{and OVERSPEEDING} = \text{TRUE}\) 
\(\text{and TYPE} = \text{SMALL CAR}\) 
\(\text{and SEASON} = \text{DRY}\) 
\(\text{and \text{TYREBURST} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{FALSE}}\) 
\(\text{and \text{LOSS-OF-CONTROL} = \text{FALSE}}\) 
\(\text{and \text{OVERSPEEDING} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3

Rule 11 If \(\text{TIME} = \text{MORNING}\) 
\(\text{and \text{LOSS-OF-CONTROL} = \text{TRUE}}\) 
\(\text{and \text{TYPE} = \text{HAEVY VEHICLE}}\) 
\(\text{and \text{SEASON} = \text{DRY}}\) 
\(\text{and \text{WROUNG-OVERTAKING} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{TRUE}}\) 
\(\text{and \text{TYREBURST} = \text{FALSE}}\) 
\(\text{and \text{OVERSPEEDING} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3

Rule 12 If \(\text{TYREBURST} = \text{TRUE}\) 
\(\text{and \text{TYPE} = \text{HAEVY VEHICLE}}\) 
\(\text{and \text{TIME} = \text{AFTERNOON}}\) 
\(\text{and \text{SEASON} = \text{DRY}}\) 
\(\text{and \text{WROUNG-OVERTAKING} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{FALSE}}\) 
\(\text{and \text{OVERSPEEDING} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3

Rule 13 If \(\text{TIME} = \text{MORNING}\) 
\(\text{and \text{LOSS-OF-CONTROL} = \text{TRUE}}\) 
\(\text{and \text{TYPE} = \text{HAEVY VEHICLE}}\) 
\(\text{and \text{SEASON} = \text{DRY}}\) 
\(\text{and \text{WROUNG-OVERTAKING} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{FALSE}}\) 
\(\text{and \text{TYREBURST} = \text{FALSE}}\) 
\(\text{and \text{OVERSPEEDING} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3

Rule 14 If \(\text{UNKNOWN-CAUSES} = \text{TRUE}\) 
\(\text{and \text{TYPE} = \text{SMALL CAR}}\) 
\(\text{and \text{TIME} = \text{MORNING}}\) 
\(\text{and \text{SEASON} = \text{WET}}\) 
\(\text{and \text{WROUNG-OVERTAKING} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{FALSE}}\) 
\(\text{and \text{LOSS-OF-CONTROL} = \text{FALSE}}\) 
\(\text{and \text{TYREBURST} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3

Rule 15 If \(\text{TYREBURST} = \text{TRUE}\) 
\(\text{and \text{TYPE} = \text{HAEVY VEHICLE}}\) 
\(\text{and \text{SEASON} = \text{WET}}\) 
\(\text{and \text{TIME} = \text{EVENING}}\) 
\(\text{and \text{WROUNG-OVERTAKING} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{FALSE}}\) 
\(\text{and \text{OVERSPEEDING} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3

Rule 16 If \(\text{TIME} = \text{MORNING}\) 
\(\text{and \text{TYREBURST} = \text{TRUE}}\) 
\(\text{and \text{TYPE} = \text{HAEVY VEHICLE}}\) 
\(\text{and \text{SEASON} = \text{WET}}\) 
\(\text{and \text{WROUNG-OVERTAKING} = \text{FALSE}}\) 
\(\text{and \text{CARELESSDRIVING} = \text{FALSE}}\) 
\(\text{and \text{OVERSPEEDING} = \text{FALSE}}\) 
\(\text{and \text{TREE-OBSTRUCTION} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SHAFT} = \text{FALSE}}\) 
\(\text{and \text{BROKEN-SPRING} = \text{FALSE}}\) 
\(\text{and \text{BRAKE-FAILURE} = \text{FALSE}}\) 
\(\text{and \text{ROAD-PROBLEM} = \text{FALSE}}\) 
\(\text{and \text{UNKNOWN-CAUSES} = \text{FALSE}}\) 
\(\text{and \text{ROBBERY-ATTACK} = \text{FALSE}}\) then LOCATION3
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION3

Rule 17 If CARELESSDRIVING = TRUE
and TYPE = HAEVY VEHICLE
and SEASON = DRY then LOCATION3

Rule 18 If TIME = MORNING
and TYPE = SMALL CAR
and SEASON = DRY
and CARELESSDRIVING = FALSE
and WRONG-OVERTAKING = FALSE
and LOSS-OF-CONTROL = FALSE
and TREE-OBSTRUCTION = FALSE
and BRAKE-FAILURE = FALSE then LOCATION3

Rule 19 If TIME = NIGHT then LOCATION2

Rule 20 If WRONG-OVERTAKING = TRUE
and TYPE = SMALL CAR then LOCATION2

Rule 21 If TIME = EVENING
and CARELESSDRIVING = TRUE then LOCATION2

Rule 22 If TIME = EVENING
and UNKNOWN-CAUSES = TRUE then LOCATION2

Rule 23 If TIME = EVENING
and LOSS-OF-CONTROL = TRUE then LOCATION2

Rule 24 If TIME = EVENING
and ROBBERY-ATTACK = TRUE then LOCATION2

Rule 25 If TIME = EVENING
and TYPE = HAEVY VEHICLE
and SEASON = DRY then LOCATION2

Rule 26 If SEASON = WET
and TYPE = MOTOCYCLE then LOCATION2

Rule 27 If SEASON = WET
and OVERSPEEDING = TRUE
and TIME = MORNING then LOCATION2

Rule 28 If TYREBURST = TRUE
and SEASON = WET
and TYPE = SMALL CAR then LOCATION2

Rule 29 If TYREBURST = TRUE
and SEASON = WET
and TIME = MORNING
and TYPE = HAEVY VEHICLE
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE

and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 30 If TYPE = HAEVY VEHICLE
and ROBBERY-ATTACK = TRUE then LOCATION2

Rule 31 If TYPE = HAEVY VEHICLE
and OVERSPEEDING = TRUE
and TIME = AFTERNOON then LOCATION2

Rule 32 If TYREBURST = TRUE
and SEASON = WET
and TYPE = HAEVY VEHICLE
and ROBBERY-ATTACK = TRUE then LOCATION2

Rule 33 If TYREBURST = TRUE
and SEASON = WET
and TYPE = HAEVY VEHICLE
and TIME = AFTERNOON
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 34 If TYPE = HAEVY VEHICLE
and TIME = EVENING then LOCATION2

Rule 35 If TYPE = HAEVY VEHICLE
and OVERSPEEDING = TRUE
and TIME = MORNING
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and TYREBURST = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 36 If TYREBURST = TRUE
and TIME = AFTERNOON
and TYPE = SMALL CAR
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 37 If BRAKE-FAILURE = TRUE
and TYPE = MOTORCYCLE then LOCATION2

Rule 38 If WRONG-OVERTAKING = TRUE
and TIME = AFTERNOON then LOCATION2

Rule 39 If TREE-OBSTRUCTION = TRUE
and TIME = MORNING then LOCATION2

Rule 40 If BROKEN-SPRING = TRUE
and TYPE = HEAVY VEHICLE
and TIME = MORNING
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 41 If TYPE = HEAVY VEHICLE
and TYREBURST = TRUE
and TIME = AFTERNOON
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 42 If LOSS-OF-CONTROL = TRUE
and TIME = MORNING
and TYPE = SMALL CAR then LOCATION2

Rule 43 If UNKNOWN-CAUSES = TRUE
and TYPE = HEAVY VEHICLE
and SEASON = DRY then LOCATION2

Rule 44 If OVERSPEEDING = TRUE
and TIME = AFTERNOON
and SEASON = WET then LOCATION2

Rule 45 If TYPE = HEAVY VEHICLE
and LOSS-OF-CONTROL = TRUE
and TIME = MORNING
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 46 If SEASON = WET
and LOSS-OF-CONTROL = TRUE
and TIME = AFTERNOON
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 47 If CARELESSDRIVING = TRUE
and TIME = AFTERNOON
and TYPE = SMALL CAR then LOCATION2

Rule 48 If OVERSPEEDING = TRUE
and TIME = AFTERNOON
and TYPE = SMALL CAR
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and TYREBURST = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 49 If SEASON = WET
and TIME = EVENING
and TYPE = SMALL CAR
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and LOSS-OF-CONTROL = FALSE
and TYREBURST = FALSE
and OVERSPEEDING = TRUE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

Rule 50 If TYPE = HEAVY VEHICLE
and LOSS-OF-CONTROL = TRUE
and TIME = AFTERNOON
and SEASON = DRY
and WRONG-OVERTAKING = FALSE
and CARELESSDRIVING = FALSE
and TYREBURST = FALSE
and OVERSPEEDING = FALSE
and TREE-OBSTRUCTION = FALSE
and PUSHED-BY-A-CAR = FALSE
and BROKEN-SHAFT = FALSE
and BROKEN-SPRING = FALSE
and BRAKE-FAILURE = FALSE
and ROAD-PROBLEM = FALSE
and UNKNOWN-CAUSES = FALSE
and ROBBERY-ATTACK = FALSE then LOCATION2

5. Discussion

There are 50 rules generated from this tree. Rule 1-18 indicate the occurrence of accident in Location 3 and rule 19-50 also shows the occurrence of accident in location 2. This indicates that, location 2 has the highest number of road accident occurrence with Heavy-vehicle in the afternoon and during the dry season.

Rule 41 is the best one that can be used for prediction. The rule says that, Tyre bust is the cause of road accident with heavy vehicle within location 2 in the day time and during the dry season.

Decision Tree Performance Analysis on Iđ3

Table 5.1. Detailed Accuracy By class

| Class     | TP rate | FT rate | Precision | Recall | F-measure | Roc Area |
|-----------|---------|---------|-----------|--------|-----------|----------|
| Location (3) | 0.688   | 0.069   | 0.733     | 0.688  | 0.71      | 0.942    |
| Location (2) | 0.897   | 0.361   | 0.78      | 0.897  | 0.834     | 0.888    |
| Location (1) | 0.517   | 0.025   | 0.833     | 0.517  | 0.638     | 0.95     |
| Weighted Avg. | 0.777   | 0.232   | 0.78      | 0.777  | 0.769     | 0.912    |

Table 5.2. Confusion matrix Predicted category

| Actual category | Location (3) | Location (2) | Location (1) |
|----------------|--------------|--------------|--------------|
| Location (3)   | 22           | 10           | 0            |
| Location (2)   | 6            | 78           | 3            |
| Location (1)   | 2            | 12           | 15           |

Decision Tree performance Analysis on Function Tree (FT)

Table 5.3. Detailed Accuracy By Class

| Class     | TP rate | FT rate | Precision | Recall | F-measure | Roc Area |
|-----------|---------|---------|-----------|--------|-----------|----------|
| Location (3) | 0.625   | 0.086   | 0.667     | 0.625  | 0.645     | 0.869    |
| Location (2) | 0.77    | 0.361   | 0.753     | 0.77   | 0.761     | 0.736    |
| Location (1) | 0.586   | 0.101   | 0.586     | 0.586  | 0.586     | 0.832    |
| Weighted Avg. | 0.703   | 0.25    | 0.702     | 0.703  | 0.702     | 0.783    |

Table 5.4. Confusion Matrix Predicted category

| Actual category | Location (3) | Location (2) | Location (1) |
|----------------|--------------|--------------|--------------|
| Location (3)   | 20           | 12           | 0            |
| Location (2)   | 8            | 67           | 12           |
| Location (1)   | 2            | 10           | 17           |
6. Conclusions

Using WEKA software to analyze accident data collected on Lagos-Ibadan road, it was found that decision tree can accurately predict the cause(s) of accident and accident prone locations along the road and other roads if relevant data are gathered and analyzed as in this case.

In Decision Tree Performance analysis, the dataset were experimented with two algorithms; Id3 and FT (function tree). For Id3 algorithm, there were 115 correctly classified instances and 33 incorrectly classified instances which represent 77.70% and 22.29% respectively. Mean absolute error was 0.1835 and Root mean squared error was 0.3029.

Also for functional tree algorithm (FT), total number of tree size was 5 with 105 correctly classified instances representing 70.27% and 44 incorrectly classified instances representing 29.73%.

From the detailed accuracy by class and confusion matrix, Id3 attained accuracy rate of 0.777 and FT attained accuracy rate of 0.703.

REFERENCES

[1] Akomolafe et al (2009) “Enhancing road monitoring and safety through the use of geographic technology” International Journal of Physical Sciences Vol. 4 (5), pp. 343-348

[2] Akomolafe, O.P. (2004); predicting possibilities of Road Accidents occurring, using Neural Network. M. Sc. Thesis, Department of Computer Science, University of Ibadan

[3] Abdalla, I.M., Robert, R., Derek, B. and McGuicagan, D.R.D., (1987) An investigation into the relationships between area social characteristics and road accident casualties. Accid. Anal. Prev. 29 5, pp. 583-593, 1997

[4] Gelfand, S.G., Ravishanker, C.S., and Delp, E.J. (1991) An iterative Growing and Pruning Algorithm for Classification Tree Design, PAMI(13), No. 2, February 1991, pp. 163-174

[5] Han J. and Kamber M. (2001) Data mining Concepts and Techniques Morgan Kaufman, Academic Press

[6] Han J. and Kamber M. (2001) Data mining Concepts and Techniques Morgan Kaufman, Academic Press

[7] Hand, D., Mannila, H., & Smyth, P., (2001) Principles of data Mining. The MIT Press, 2001

[8] Kim, K., Nitz, L., Richardson, J., & Li, L., (1995) Personal and Behavioral Predictors of Automobile Crash and Injury Severity. Accident Analysis and Prevention, Vol. 27, No. 4, 1995, pp. 469-481

[9] Martin, P. G., Crandall, J. R., & Pilkey, W. D.,(2000) Injury Trends of Passenger Car Drivers in the USA Accident Analysis and Prevention, Vol. 32, 2000, pp. 541-557

[10] Ossenbruggen, P.J., pendharkar, J. and Ivan, J., (2001) Roadway safety in rural and small urbanized areas. Accid. Anal. Prev. 334, pp. 485-498, 2001