**Machine Learning Research**
2017; 2(1): 26-34
http://www.sciencepublishinggroup.com/j/mlr
doi: 10.11648/j.mlr.20170201.14

SciencePG
Science Publishing Group

# A Decision Tree Algorithm Based System for Predicting Crime in the University

**Adewale Opeoluwa Ogunde[1, *], Gabriel Opeyemi Ogunleye[2], Oluwaleke Oreoluwa[1]**

[1]Department of Computer Science, Redeemer's University, Ede, Nigeria

[2]Department of Computer Science, Federal University, Oye-Ekiti, Nigeria

## Email address:

ogundea@run.edu.ng (A. O. Ogunde)

[*]Corresponding author

## To cite this article:

Adewale Opeoluwa Ogunde, Gabriel Opeyemi Ogunleye, Oluwaleke Oreoluwa. A Decision Tree Algorithm Based System for Predicting Crime in the University. *Machine Learning Research.* Vol. 2, No. 1, 2017, pp. 26-34. doi: 10.11648/j.mlr.20170201.14

**Received**: January 27, 2017; **Accepted**: February 13, 2017; **Published**: March 2, 2017

**Abstract:** CRIME is one of the major problems encountered in any society and universities together with other higher institutions of learning are not exceptions. Thus, there is an urgent need for security agents and agencies to battle and eradicate crime. The Directorate of Students and Services Development (DSSD) are responsible for investigating and detecting criminals of any crime committed within the Redeemer's University. DSSD faces major challenges when it comes to detecting the real perpetrators of several crimes. An improvement in their strategy can produce positive results and high success rates, which is the basic objective of this project. Several methods have been applied to solve similar problems in the literature but none was tailored to solving the problem in Redeemer's University and other universities. This work therefore applied classification rule mining method to develop a system for detecting crimes in universities. Past data for both crimes and criminals were collected from DSSD. In order to develop and test the proposed model, the data was pre-processed to get clean and accurate data. The Iterative Dichotomiser 3 (ID3) decision tree algorithm obtained from WEKA mining software was used to analyze and train the data. The model obtained was then used to develop a system that showed the hidden relationships between the crime-related data, in form of decision trees. This result was then used as a knowledge base for the development of the crime prediction system. The developed system could effectively predict a list of possible suspects by simply analyzing data retrieved from the crime scene with already existing data in the database. This system has all the potentials of helping the students' affairs department and security apparatus of any university and other institutions to quickly detect either the real or possible perpetrators of crimes in the system.

**Keywords:** Crime, Classification Rules, Data Mining, Decision Trees, ID3, Prediction, University

## 1. Introduction

Crime is an act usually deemed socially harmful, specifically defined, prohibited and punishable under criminal law. Crime rates are rapidly increasing and changing. Crime prediction is very important in any university. At Redeemer's University, crime detection among students is very crucial and especially to the Directorate of Students and Services Development (DSSD) – It is their sole responsibility to enforce the law, find and apprehend irresponsible students, reduce and curtail any and every form of indiscipline. Some of the crimes in this situation include: burglary, sexual harassment (and/or rape), abuse of drugs, alcoholism, homosexuality, misuse and abuse of school properties, disobedience of school rules and regulations, stealing and many other crimes included in the Redeemer's University Code of Conduct/Rules and Regulations. After a crime is committed, it becomes the duty of the Directorate of Students and Services Development (DSSD) to forecast the potential suspects of that crime, perform a series of investigations, apprehend and then prosecute the real criminal.

According to [15], data mining is the extraction of concealed prescient data from substantial databases. It is a new innovation with awesome potential to help organizations concentrate on the most imperative data in their data warehouses. Data mining techniques foresee future patterns and practices, permitting organizations to make proactive, learning driven choices. Data mining tends to work well when we have a large amount of data. Grouping guideline is among the normally connected data mining procedures, which allows us to arrange or foresee estimations of target variables from estimations of attributes of variables [16]. The main objective of data mining is prediction. Therefore, here data mining will be utilized for prediction of potential suspects by simply analyzing data retrieved from the crime scene with already existing data in the database. A more accurate system which is computerized will help reduce all these limitations. The ID3 decision tree learning algorithm will be utilized to mine the data and the rules generated will be utilized as a model for anticipating conceivable suspects from past records.

The existing system used by the Directorate of Students and Services Development (DSSD) are manual systems that require the initiative of the security to determine and detect crimes and relationship between crimes and suspects. Securities face difficulty in effectively predicting and detecting crimes and suspects. This may result in unsolved cases and the arrest and incarceration of innocent students. The existing system used by the Directorate of Students and Services Development requires more man power, it is time consuming, needs manual calculations and it is imperfect. Therefore, a new system that could efficiently solve the highlighted problems becomes inevitable, which is the major focus of this work.

## 2. Data Mining

The removal of several hidden patterns from a large database which appears useful is called data mining [15]. It is a strong technology with substantial potential to help and knowing the world and explain natural phenomenon. Over the years, people have been gathering and analyzing data organizations concentrate on the most valuable information in their database. However, gradually new technologies have begun to play a vital role to handle the storage, analysis and processing of data. Specially, the advent of computer technology has revolutionized the way in which data are managed. This new method of searching through the data as well as the keen interest to learn from data has brought disciplines like that of data mining. Some researchers also noted that data mining is valuable to discover relevant and useful information from huge data stored in the database through building computer programs that sort through the database automatically, seeking meaningful patterns [18]. The opportunity for the application of data mining has increased tremendously as databases grew extremely and new machine with searching capabilities evolved.

Looking into this data has introduced various theories, observations, and approaches that will help in understanding the law and knowing the natural world [2].

### *2.1. The Data Mining Process*

Data mining is not all about the use of software [17]. It is a process that involves series of steps to transform data prior to mining, evaluate and interpret modelling results. It is the process of discovering relevant patterns in a large amount of data that can describe or tell us about past events in a way that the modelled results can be employed to predict the future [15].

According to [2], the most frequently used data mining steps are; identifying the source of the data, preparing data for analysis, building and training a computer model and evaluating the computer model.

### *2.2. Classification Rule Mining and Decision Trees*

Classification is a commonly used data mining technique; it uses a large population of records for classification to create a model from a set of pre-classified examples. This technique makes use of decision tree or neural network-based classification algorithms. The most common task in data mining is to build models to predict the class of an object based on its attributes. Classification trees can have binary or unary branches. Most times the tree structure in classification trees have binary branches, when we split the data in two ways it will result in a better separation. The classification process of data involves learning and classification.

In learning, the classification algorithm analyzes the data to be trained. In classification, the accuracy of the classification rules estimates the test data. The classifier-training algorithm makes use of pre-classified examples to determine the set of parameters needed for proper discrimination [14]. Classification is also called supervised learning; supervised learning is a process where network can learn from target vector containing the desired output from a pair of input vector. The prediction next time would be closer to the correct answer if the learning algorithm can take the differences between the correct output and the prediction of neural networks.

### *2.3. How Decision Trees Work*

According to [2], decision tree can be used to classify and predict the dividing records in the database to small sets in relation to values of other fields. The decision tree construction starts at the root node where you take actions from the root node you split the remaining nodes constantly irrespective of the decision tree learning algorithm. The output is a decision tree with each branch having a decision made at each step [13]. Classification task is a rule of the decision being made [15]. It expresses if-then' explicitly compared to neural networks due to its strength and popularity. Decision tree works with computer-generated rules which can provide an explanation for its action unlike neural networks.

### *2.4. Application of Data Mining in Securities*

Data mining has been applied in security issues which are the main focus of this project; it has also been used in business operations. The following sections have used data mining:

#### *2.4.1. Terrorist Modus Operandi Detection System (TMODS) [12]*

TMODS controls the task of hunting down and separating instances of pattern activity that appears threatening. The analyst can characterize a relational graph with an attribute to represent the pattern of threatening activity he or she is searching for and hunt down the threat pattern through an input graph that speaks to the observed data with a large volume. TMODS recognizes the subset of data that will coordinate the threat pattern characterized by the analyst, which will turn a normal search into an efficient automated graph matching tool [6]. Pattern graph can be produced with possible network ontology which will be matched against observed activity graph. The analyst goes through the matches that are listed against the input graph, powerful and mature distributed java software that has been under development since October 2001 is TMODS [6]. An analyst and a pattern graph are needed to run the system. It tailors the result alongside pre-defined threatening activity like a supervised learning algorithm. A downside in TMODS is the graph utilized; they appear to have multi-node which can render further analysis pointless.

#### *2.4.2. The Over Project [9]*

Decision support system was used to assist the policing of the volume crime of Burglary from Dwelling Houses (BDH). The OVER project started in 2000 in UK with West Midlands Police. The techniques used for the OVER Project are:

I. Creating profiles of offenders (using Kohonen neural networks); and the use of offender profile to expose attribute of a certain offender with pending crimes and provide an ordered list of possible offenders based on their profiles.
II. Bayesian belief network ([10], [5], [3]).

### *2.5. Other Related Works*

The law enforcement and investigating agency in Ethiopia now extract hidden knowledge rooted in their data warehouse which is of importance in fighting crime. The police department will like to expose regular crime patterns from past records in order to investigate new cases and avoid future cases by using preventive measures based on previous patterns [8] this has helped to reduce training time of their officers that will be assigned to a new location.

In America, for the investigation of crime data mining technique was adopted to search through large volume of data gotten from various sources to track down criminals [2]. Also, the treasury department made use of data mining to extract suspicious fraud patterns. The aim is to detect crime in money laundry.

Associative classification uses association rules for new tuples. It consists of association rule mining and classification. Strong association between frequent patterns and class labels was searched for. The main aim was to improve the accuracy of a classifier, which can be achieved by producing all types of negative class association rules [11].

According to McClendon and Meghanathan [7], data mining and machine learning have become a vital part of crime detection and prevention. In [7], WEKA was used to carry out a comparative study between the violent crime patterns from the communities and crime unnormalized dataset, which was provided by the University of California-Irvine repository and actual crime statistical data for the state of Mississippi that has been provided by neighborhoodscout.com. Linear Regression, Additive Regression, and Decision Stump algorithms using the same finite set of features were the algorithms deployed on the communities and crime dataset. Their results showed that the linear regression algorithm had the best performance among the three selected algorithms.

Another recent study by Barnadas [1], provided an understandable explanation of what machine learning is and also solved a real data classification problem (namely San Francisco crimes classification) using three different and popular algorithms, which are: K-Nearest Neighbours, Parzen windows and Neural Networks.

As at date, artificial intelligence and machine learning techniques are already being used to assist humans in fighting cyber crimes, as they provide flexibility and learning capabilities. It has become a widely acceptable method used in decision making processes. Intelligent decision support in cyber defence can also be successfully achieved using these methods. Available academic resources show that artificial intelligence and machine learning techniques already have numerous applications in combating cyber crimes. Advances made so far in these areas of using artificial intelligence to combat crimes, their current limitations and desired characteristics were presented in [4].

## 3. Methodology

### *3.1. Data Collection and Description*

The data used in this research was collected from the Directorate of Students and Services Development (DSSD) of Redeemer's University, Nigeria. The data collected consists of crime records for about ten years, that is, from 2005-2015, which consists of the offence, number of male students, number of female students, number of 100 level students, number of 200 level students, number of 300 level students, number of 400 level students and number of 400+ (i.e. 500) level students. The data used in this work consists of records, which were described by the attributes: programme, sex, offence, expulsion period and level. A fraudulent record forms a transaction. Theoretically, we cannot have a fraudulent record. The behaviour of a person appears to be the only way crime can be investigated and this can be done by checking out the relating attributes which was built on forensic data then developed on record attributes. The major crimes are immoral act, theft and drugs (drug abuse). The aim is to use data mining technology to detect complex patterns hidden in the data. Decision tree algorithm was used to analyze the patterns better than a data analyst would do. Table 1 shows a description of data used in the work. It indicated the offence, sex, and level of expelled students because of crime.

***Table 1.*** *Description of Data.*

| OFFENCE | MALES | FEMA-LES | 100 L | 200 L | 300 L | 400 L | 500 L |
|---|---|---|---|---|---|---|---|
| DRUGS | 105 | 2 | 26 | 31 | 22 | 25 | 3 |
| THEFT | 41 | 4 | 5 | 8 | 14 | 14 | 4 |
| IMMORAL ACT | 29 | 22 | 8 | 15 | 7 | 18 | 3 |
| TOTAL NUMBER | 175 | 28 | 39 | 54 | 43 | 57 | 10 |

### *3.2. The ID3 Decision Tree Algorithm*

The ID3 decision tree model was used for training our model for this work. This tree speaks to a number of choices that exist between their alternatives utilizing every branch node and the decision being made is done by the leaf node. The principal purpose of decision making is for increasing important information. The development of the decision tree starts at the root node where users will take the right decision. Users will isolate each node from the root node using this algorithm recursively. Where every branch will speak to its result and conceivable situation of decision will produce a decision tree. Ross Quinlan added to the ID3 decision tree algorithm and made it straightforward. The ID3 algorithm is to gather a decision tree with the utilization of a top-down and avaricious hunt utilizing the given set for the test of every attribute that exists at every node. To get the attribute that is more valuable for the characterization of the given set, the presentation of metric-information gain is performed. This algorithm is an essential one in classification of decision tree and it is broadly connected, it seeks through the training instances with attributes and brings out the relevant attributes that isolates the examples in the most ideal way. ID3 stops if the attributes of the training set is splendidly classified, else it proceeds with operation on x (i.e., where x= number of conceivable values of an attribute) isolated subsets to achieve the best attribute. The algorithm chooses the best attribute with the utilization of avaricious hunt and it doesn't return to prior decisions being made. The standard of ID3 algorithm is Information theory. ID3 was used because it is easy to use and also very effective. A Pseudocode of the ID3 algorithm is presented here.

Step 1: Start

Step 2: Create a root node

Step 3: Develop classification attribute

Step 4: Calculate entropy of classification

Step 5: For every attribute in R, Use the attribute of classification to compute Information Gain

Step 6: Starting from the root node select attribute with the maximum gain that is the next node in the tree.

Step 7: Develop reduced table R's by removing node attribute

Step 8: Repeat steps 5- pending all attributes have been used

Step 9: Stop

Step 10: Return root

### *3.3. Data Training*

The flowchart for training of the data is shown in Figure 1 and the overall flowchart of the system is also shown in Figure 2. In Figure 1, the data selected is converted into Attribute Relation File Format (ARFF) using the ARFF converter and classification was done using WEKA and in order to produce decision trees. The "IF-THEN" rules gotten from the decision tree (Figure 2) is placed into the knowledge base of the prediction system and the user will use the user interface for interaction and related questions will be asked which will now be used to put together rules in the knowledge base of the system.

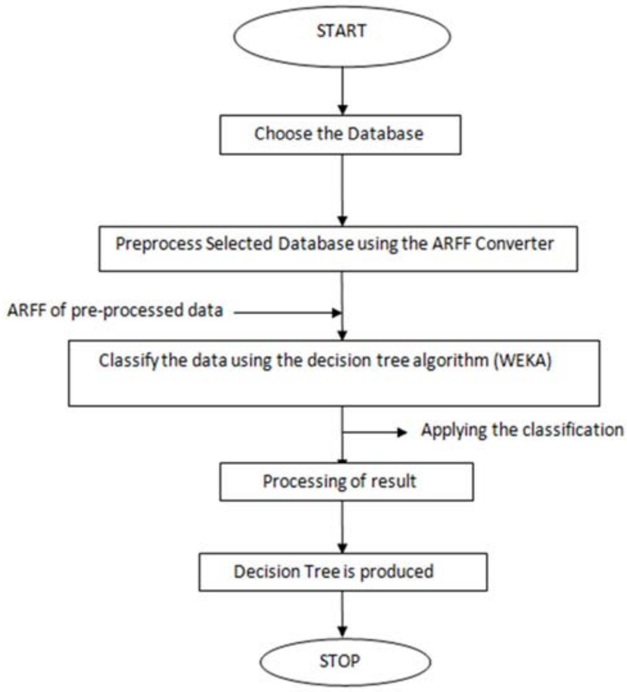


***Figure 1.*** *Flowchart for training of the data.*

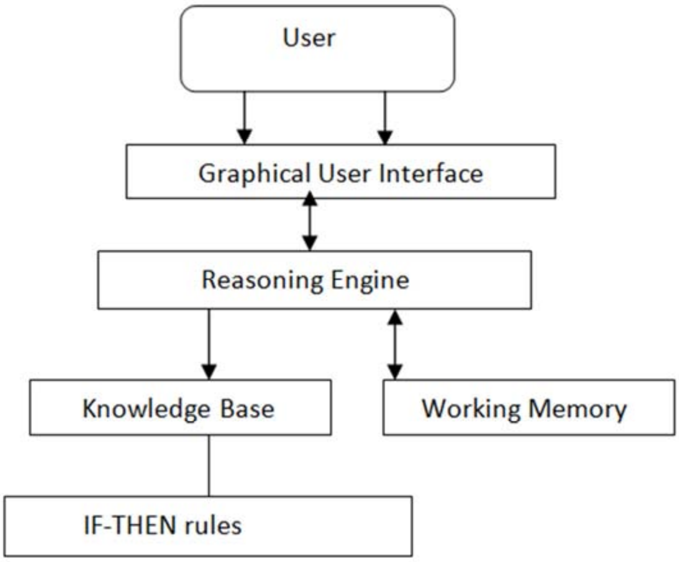


***Figure 2.*** *System Architecture.*

### *3.4. Description of the Attributes*

The paper uses a single major table called rules. The rules in the table will serve as the knowledge base. It stores information about all users of the application. Level is the primary key for this table. The other major fields and their attributes are described as follows.

Programme: This field contains the various programmes in the university.

Sex: This is the gender of the student. It occurred in two forms, i.e., as “MALE” or “FEMALE”.

Offence (OFF): In this field, we record the types of crime could have been committed by various students. For this research work, this field is very vital because other fields depend on it. This means that before any record can be entered into the database, a crime must have been committed.

Level: This was used to represent the current level of the student as at the time the offence was committed. It will be used to reference the records within the database.

Letter of warning (LOW): This is a record of the number of warning letters a student has been given by the DSSD. This will determine if the student will be eventually expelled.

### *3.5. The Design of the Prediction System*

The developed system has a knowledge base that assembles knowledge and an arrangement of rules for the application of the knowledge base to every situation being depicted to the program in the Figure below. This is a knowledge-base that incorporates past records of bad behavior with the demerit point and its relationship with new records after facing panel. It gives advice about the expulsion of students from the university. Figure 3 shows the architecture of the prediction system.

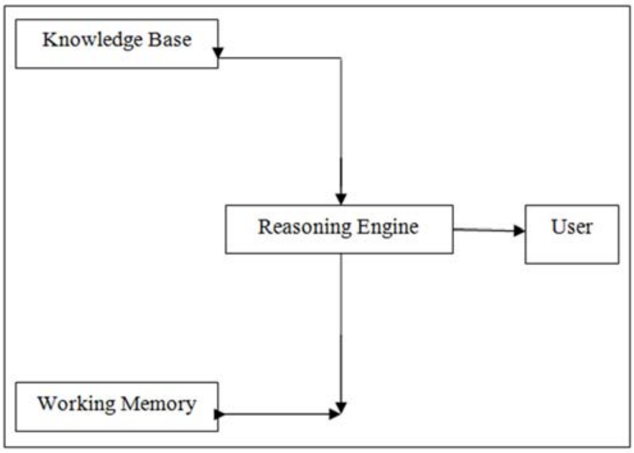


***Figure 3.*** *Architecture of the prediction system.*

Knowledge Base: It contains the domain knowledge (Domain knowledge means having specialized knowledge about a certain problem). Here, we have the rules and it will be implemented using WEKA “IF-THEN” rules

Working Memory: The working memory accepts information from users about the current problem through the inference engine and it will match the information stored in the working memory alongside the domain knowledge in the knowledge base to arrive at a conclusion.

Reasoning Engine: This is the processor of the prediction system. It combines facts and specialized knowledge to get new information in order to draw conclusions then adds the conclusion to the working memory.

User: The individual will use the system to generate predictions from a set of data.

## 4. Implementation and Results

This section deals with the implementation of the Crime Prediction System using the ID3 decision tree algorithm. The tools that were used for the system implementation are described here. JDK (Java Developer's Kit) with NetBeans IDE 8.0.2 (Integrated Development Environment) were used to develop the prediction application interface. WEKA Explorer was used to train the prediction model. It contains a set of visualization tools and algorithms for the analysis of data and modelling prediction, alongside the GUI for easy access to its functionality. MySQL relational database management system was used to store and manage data directly within the database. WEKA stores data in flat files i.e. ARFF format. The ARFF converter was used to convert the data from excel file into ARFF file format acceptable in WEKA.

### *4.1. Model Construction using ID3 Decision Tree*

ARFF converter was used to select and pre-process the raw data into the format acceptable by WEKA for model construction. The ARFF pre-processed data was then trained by the WEKA implementation tool, specifically with the use of ID3 algorithm under the classify panel in WEKA. Set of rules were generated by the system and the resulting model from the classifier are shown in figures 4. Figure 4 shows the ID3 rules generated by WEKA, the error margins and prediction accuracies of the classified model.

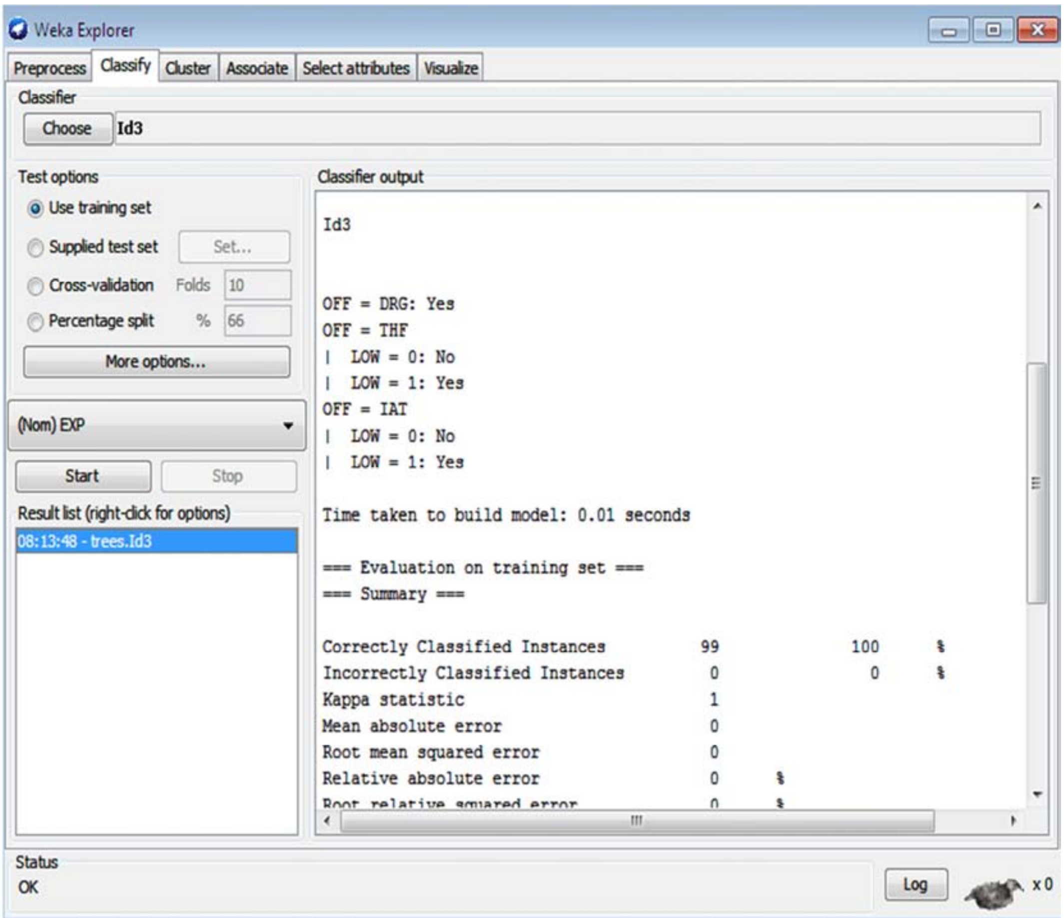


***Figure 4.*** *Some rules generated using WEKA.*

Figure 4 showed that the time taken to build the model was 0.01 seconds with 99% of the data reported as correctly classified instances while 1% was incorrectly classified. Koppa statistic was 1, which, which showed that there was agreement of prediction with true class. Also, Mean Absolute error, Root Mean Square Error, Relative Absolute Error, Root Relative Squared Error were all zero. This showed that the resulting model was very reliable. The decision tree generated from the rules as produced by WEKA is shown in figure 5.

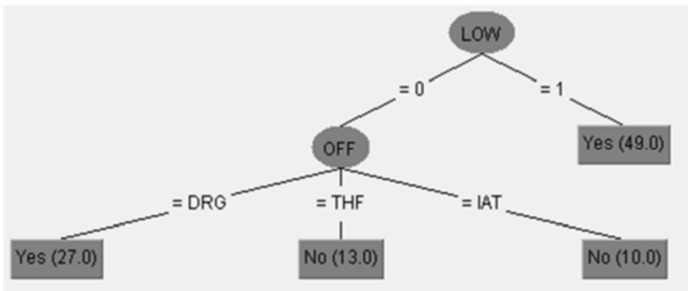


***Figure 5.*** *Decision Tree produced using WEKA.*

### *4.2. Knowledge Representation*

The knowledge represented by the decision tree generated was extracted and represented in the form of IF-THEN rules and this is presented in Table 2. For instance, rule one means that if the concerned student is in 100 levels and student has never been given a letter of warning and the new offence committed is drugs (i.e. drug abuse), then the prediction is that the student is a suspect. Likewise for the last rule, if the student is an extra year student and has been issued one letter of warning and the offence committed is drugs, then the prediction is that the student is a strong suspect based on his/her past records. So, further investigation could be carried by security agencies in charge of students to further verify this claim.

***Table 2.*** *RULE Set generated by ID3.*

- IF LEVEL ='100' and LOW = '0' and OFF ='DRG', THEN PRD =' Based on past records, student is a suspect'
- IF LEVEL ='200' and LOW ='0' and OFF ='DRG', THEN PRD =' Based on past records, student is a suspect'
- IF LEVEL ='300' and LOW ='0' and OFF = 'DRG', THEN PRD =' Based on past records, student is a suspect'
- IF LEVEL ='400' and LOW ='0' and OFF ='DRG', THEN PRD =' Based on past records, student is a suspect'
- IF LEVEL ='500' and LOW='0' and OFF ='DRG', THEN PRD =' Based on past records, student is a suspect'
- IF LEVEL='100' and LOW='0' and OFF ='THF', THEN PRD = 'Based on past records, this is a new case'
- IF LEVEL ='200' and LOW ='0' and OFF ='THF', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='300' and LOW ='0' and OFF ='THF', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='400' and LOW ='0' and OFF ='THF', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='500' and LOW ='0' and OFF ='THF', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='100' and LOW ='0' and OFF ='IAT', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL='200' and LOW ='0' and OFF ='IAT', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='300' and LOW ='0' and OFF ='IAT', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='400' and LOW ='0' and OFF ='IAT', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL ='500' and LOW ='0' and OFF ='IAT', THEN PRD =' Based on past records, this is a new case'
- IF LEVEL= '100' and LOW ='1' and OFF ='DRG', THEN PRD =' Based on past records, student is a strong suspect '
- IF LEVEL= '200' and LOW ='1' and OFF ='DRG', THEN PRD =' Based on past records, student is a strong suspect'
- IF LEVEL= '300' and LOW ='1' and OFF ='DRG', THEN PRD =' Based on past records, this is a strong suspect'
- IF LEVEL= '400' and LOW ='1' and OFF ='DRG', THEN PRD =' Based on past records, this is a strong suspect'
- IF LEVEL= '500' and LOW ='1' and OFF ='DRG', THEN PRD =' Based on past records, this is a strong suspect'

### *4.3. Prediction Interface*

The prediction system interface was tested in order to check if it works properly with the if-then rules used to model it, which served as the knowledge base in the prediction system. It has a login page which serves as an introduction page to the user of the application. The security check of the system is the username and password and it grants access to the prediction system if the user can enter a correct username and password. The prediction interface was based on a combination of rules to determine whether a student is a suspect when a crime is committed. The 'Student Information' tab allows the system administrator to enter details about students suspected to be involved in a new crime. The number of suspects is entered under the 'Suspects' tab. The 'Student Crime Data' tab is used to populate the prediction system with students' crime data details as they occur. Usually it checks the database whether the student involved has committed any crime in the past or it is a new case. The 'Prediction' tab contains the main commands for carrying out the actual prediction of a crime. Suspected students details including the offence (s) presently committed are entered into the system. The system thus predicts whether the concerned students are suspects based on the existing past knowledge in the system. A sample prediction is shown in figure 6.

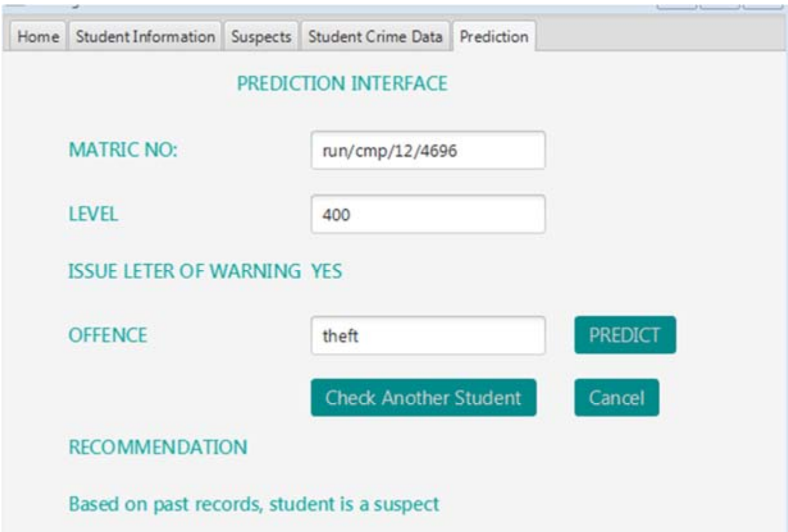


***Figure 6.*** *Prediction Interface showing "Based on past records, student is a suspect".*

### 4.4. Class-Wise Accuracy and Accuracy Model for Class Prediction

The system was tested with 30 percent the real dataset used for this work. Summary of predictions generated by the developed system are presented in Table 3. Table 3 shows the true positive and false positive classifications. In addition, the correct precision for the three-class outcome categories were presented.

***Table 3.*** *Class-wise accuracy for three class prediction.*

| CLASS | True Positive (TP) | False Positive (FP) | Correct Precision (%) |
|---|---|---|---|
| Drug | 0.940 | 0.347 | 0.734 |
| Theft | 0.435 | 0.013 | 0.909 |
| Immoral Act | 0.577 | 0.123 | 0.625 |
| Average | 0.651 | 0.161 | 0.756 |

From table 3, the average true positive rate was 0.765; this represented the ratio of correct predictions, which were the number of sample predictions that came out to be truly positive. The average false positive rate was 0.073; this represented the number of samples that were predicted positive but they were actually negative. Precision is the fraction of those predicted positives that were actually positive, which had an average of 0.756. These very positive values reported by the model implied that the model was predicting with high level of accuracy.

Table 4 shows the percentage accuracy of the developed crime prediction system. As shown in the table, correctly classified instances were 72.73% while the incorrectly classified instances were 27.27%. This showed that the model has a high level of accuracy in predicting crimes among students and it could be a very reliable tool to Students administrators and security personnel in any educational system.

***Table 4.*** *Percentage Accuracy of the System.*

| Algorithm | Correctly Classified Instances | Incorrectly Classified Instances |
|---|---|---|
| ID3 | 72.73% | 27.27% |

## 5. Conclusions and Future Works

This paper looked at the use of data mining for identifying crime patterns using classification rule mining techniques. Classification rule technique was adopted as the data mining method for this research work, which focused on developing a crime prediction system for the Directorate of Students and Services Development (DSSD) of Redeemer's University. Model generation for the prediction system was based on real-life students' crime related data obtained from DSSD, Redeemer's University. Crime prediction patterns as machine learning task were formulated in this work and to hereby utilize data mining to assist DSSD in predicting crimes and how to make use of classification rule mining for crime and suspects. By accurate analysis and representation, we were able to identify hidden relationships in the database. Predictions based on these relationships were made in order to detect crime and suspects. It could be concluded from the results of this study that the classification rule mining used in this work showed the hidden relationships between the crime data. The developed system also effectively predicted the list of possible suspects by simply analyzing data retrieved from the crime scene with already existing data in the database with 72.7% accuracy.

The result of this work showed how to deal with crime prediction using decision tree techniques. Data mining had been applied in various areas of security, crime and criminal detection. In this study, encouraging results were obtained, a sample data was used for testing and training classifiers due to time constraint. It is appropriate to perform the experiments with very large training and testing datasets as well as making a number of trials to come out with more accurate classifiers. Finally, the Students Services Department of any University can react quickly and adequately to crime situations. This makes identification of criminals more efficient and more focused. Reduction in the number of pending cases and in the population of falsely accused individuals can be achieved.

## References


[1] Barnadas M. V. (2016), Machine Learning Applied to Crime Prediction, A Degree Thesis submitted to the Faculty of the Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona, Universitat Politècnica de Catalunya.

[2] Berry, M. and Linoff, G. (1997). Data mining techniques: For marketing, sales and support. New York. John Wiley and Sons. Inc.

[3] Cowell, R. G., Dawid, A. P., Lauritzen, S. L. and Spiegelhalter, D. J., (1999). Probabilistic Networks and Expert Systems, Springer-Verlag.

[4] Dilek S., Çakır H. and Aydın M. (2015), Applications of Artificial Intelligence Techniques to Combating Cyber Crimes: a Review, International Journal of Artificial Intelligence & Applications (IJAIA), Vol. 6, No. 1, pp 21-39.

[5] Jenson, F. V. (1996). An Introduction to Bayesian Networks. UCL Press.

[6] Marcus, S., Moy, M. & Coffman, T. (2007), Social Network Analysis, in Diane J. Cook and Lawrence B. Holder, 'Mining Graph Data', John Wiley & Sons.

[7] McClendon L. and Meghanathan N. (2015). Using Machine Learning Algorithms to Analyze Crime Data, Machine Learning and Applications: An International Journal (MLAIJ) Vol. 2, No. 1, Pp 1-12.

[8] Megaputer Intelligence, Inc. (2012). Crime Pattern Analysis: Megaputer Case Study. http://www.elon.edu/facstaff/mconklin/cis230/cases/crime_pattern_case.pdf.

[9] Oatley G. C., Ewart B. W and Zeleznikow J., (2004). Decision Support Systems for Police: Lessons from the Application of Data Mining Techniques to 'Soft' Forensic Evidence. Expert Systems with Applications.

[10] Pearl, J. (2000). Casualty: Models, Reasoning, and Inference. Cambridge University Press, UK.

[11] Ramasubbareddy, B., Govardhan, A., & Ramamohanreddy, A. (2011). Classification Based on Positive and Negative Association Rules. International Journal of Data Engineering (IJDE) Volume (2) Issue (2), Pp 84.

[12] Seifert, J. (2007). Data Mining: An overview. CRS Report for Congress.

[13] Surjeet Kumar Y., Saurabh P. (2012), “Data Mining Application in Enrollment Management: A Case Study “International Journal of Computer Applications (0975 – 8887), Volume 41–No. 5.

[14] Samrat S. & Vikesh K. (2012) “Classification of Student’s data Using Data Mining Techniques for Training & Placement Department in Technical Education” International Journal of Computer Science and Network (IJCSN) Volume 1, Issue 4.

[15] Thearling, K. (2003). Understanding Data Mining: It’s all in the Interaction. Available at www.thearling.com.

[16] Ye, F. (2011). Investigating the effects of sample size, model misspecification, and underreporting in crash data on three commonly used traffic crash severity models. Ph. D. Dissertation. Texas A&M University, UMI Number: 3471262.

[17] Tesfaye, H. (2002). Predictive Modeling Using Data Mining Techniques In Support to Insurance Risk Assessment. Masters Thesis Addis Ababa University, Addis Ababa.

[18] Witten I. H., & Frank E., (2000). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementation. Morgan Kaufmann, October 2000, ISBN 1-55860-552-5.