A Performance Analysis of Business Intelligence Techniques on Crime Prediction

Emmanuel Ahishakiye
School of Computing and Informatics,
University of Nairobi, P.O. Box 30197 – 00100,
GPO Nairobi-Kenya
Email: ahishema [AT] gmail.com

Elisha Opiyo Omulo
School of Computing and Informatics,
University of Nairobi, P.O. Box 30197 – 00100,
GPO Nairobi-Kenya

Ruth Wario
Department of Computer Science and Informatics,
Faculty of Natural and Agricultural Sciences,
University of Free State, Private Bag X13, Kestell 9866,
Republic of South Africa

Ivan Niyonzima
Department of Information and Communication
Technology, Faculty of science,
Kabale University, P.O. Box 317
Kabale - Uganda

Abstract - **Law Enforcement agencies are faced with a problem of effectively predicting the likelihood of crime happening given the past crime data which would otherwise help them to do so. There is a need to identify the most efficient algorithm that can be used in crime prediction given the past crime data. In this research, Business intelligence techniques considered was based on supervised learning (Classification) techniques given that labeled training data was available. Four different classification algorithms that is; decision tree (J48), Naïve Bayes, Multilayer Perceptron and Support Vector Machine were compared to find the most effective algorithm for crime prediction. The study used classification models generated using Waikato Environment for Knowledge Analysis (WEKA). Manual method of attribute selection was used; this is because it works well when there is large number of attributes. The dataset was acquired from UCI machine learning repository website with a title 'Crime and Communities'. The data set had 128 attributes of which 13 were selected for the study. The study revealed that the accuracy of J48, Naïve bayes, Multilayer perceptron and Support Vector Machine (SMO) is approximately 100%, 89.7989%, 100% and 92.6724%, respectively for both training and test data. Also the execution time in seconds of J48, Naïve bayes, Multilayer perceptron and SVO is 0.06, 0.14, 9.26 and 0.66 respectively using windows7 32 bit. Hence, Decision Tree (J48) out performed Naïve bayes, Multilayer perceptron and Support Vector Machine (SMO) algorithms, and manifested higher performance both in execution time and in accuracy. The scope of this project was to identify the most effective and accurate Business intelligence technique that can be used during crime data mining to provide accurate results.**

Keywords - Law Enforcement Agencies; crime prediction; Business Intelligence; WEKA; Performance Analysis.

I. INTRODUCTION

Business Intelligence is a broad category of applications and technologies for gathering, storing, analyzing and providing access to data to help enterprise users make better business decisions. BI improves decisions by supplying timely, accurate, valuable, and actionable insights. With the rapid advancement and development of Information and Communication Technologies (ICT), organizations are now able to generate, collect and distribute huge amount data from internal and external sources, and use this data in decision making [1]. Business Intelligence Systems (BI) aims to gather in depth information from the company data and to analyze that data using different types of computer techniques and plotting which results in different types of graphs that facilitate the company in the future decisions [2].

Intelligence, security, and public safety agencies gathers large amounts of data from multiple sources - from criminal records of terrorism incidents, and from cyber security threats to multilingual open-source intelligence. Also they face the daunting task of defending against cyber security threats and protecting their intellectual assets and infrastructure. Defiantly, the large volume of criminal data from different sources (social media, historical crime data in files and crime records systems) creates many problems in different domain for instance data storage, data warehousing and data analysis [3]. Many law enforcement agencies are facing the problem of being "data rich but information poor" [4]. Data generated from different systems was usually unreliable, Data quality is poor and not validated and therefore not trusted fully. Concerns over data accuracy eroded confidence to make important decisions. During decision making, there is usually lack of timeliness of key information. Because of this huge data, there is a problem by law enforcement agencies to

predict the likelihood of crime happening given the available information and therefore a need to analyze crime data to find the most effective algorithm for crime prediction.

A well-planned crime prevention strategy not only prevents crime and victimization, but also promotes community safety and contributes to the sustainable development of countries [5]. Effective, responsible crime prevention enhances the quality of life of all citizens. The key benefit of applying business intelligent in crime Prediction is that often there are multiple complex factors which influence crimes to be committed. Business intelligence tools enables us to analyze historical crime data sets, identify the combination of factors which are most closely correlated with crimes and build a model which allows us to predict the likelihood of a crime being committed.

II. LITERATURE SURVEY

A. *Crime Prediction and Management*

Criminology is an area that focuses on the scientific study of crime and criminal behavior and law enforcement and is a process that aims to identify crime characteristics [7]. It is one of the most important fields where the application of data mining techniques can produce important results. Crime analysis, a part of criminology, is a task that includes exploring and detecting crimes and their relationships with criminals. The high volume of crime datasets and also the complexity of relationships between these kinds of data have made criminology an appropriate field for applying data mining techniques. Identifying crime characteristics is the first step for developing further analysis. The knowledge that is gained from data mining approaches is a very useful tool which can help and support police forces [8]. Solving crimes is a complex task that requires human intelligence and experience and data mining is a technique that can assist them with crime detection problems. The idea here is to try to capture years of human experience into computer models via data mining [9].

B. *Why Crime Is Predictable*

There is a strong body of evidence to support the theory that crime is predictable (in the statistical sense) mainly because criminals tend to operate in their comfort zone [10]. That is, they tend to commit the type of crimes that they have committed successfully in the past, generally close to the same time and location. Although this is not universally true, it occurs with sufficient frequency to make these methods work reasonably well. There are major theories of criminal behavior, such as routine activity theory, rational choice theory, and crime pattern theory. These theories are consolidated into what is referred to as a blended theory.

C. *Business Intelligence (BI)*

Business Intelligence is described by [11] to have evolved from Decision Support Systems. They outlined that the concept has its roots in Management Information Systems of the 1970s. Executive Information Systems then emerged in the 1980s and, eventually, was transformed into Business Intelligence. The term "Business Intelligence" was coined in the mid-1990s as a broad category of software tools and solutions for gathering, consolidating, analyzing and providing access to data in a way that lets enterprise users make better business decisions. By 2005, Business Intelligence systems had artificial intelligence and analytical support functions [6].

D. *Applications of Business Intelligence Techniques in crime Prediction and Management*

1) Using Business Intelligence Tools to Pursue Identity Thieves

Data analytics, using business intelligence (BI) tools, is perfectly situated to comb through the mountains of data to identify repeating patterns in bank frauds where theft of personally identifiable information (PII) was an associative factor to the crime. The development of BI data analytical tools specifically for law enforcement has led to the emergence of predictive policing as law enforcement agencies throughout the country attempt to prevent crimes, such as bank fraud, from happening and to apprehend criminals, such as the identity thieves who commit the bank fraud [12].

2) Application of Business Intelligence in Security and Public Safety

Security issues are a major concern for most organizations. According to the research firm International Data Corporation, large companies were expected to spend $32.8 billion in computer security in 2012, and small- and medium-size companies were to spend more on security than on other IT purchases [13]. A significant challenge in security IT research is the information stovepipe and overload resulting from diverse data sources, multiple data formats, and large data volumes. Selected Business Intelligence and Analytics (BI&A) technologies such as criminal association rule mining and clustering, criminal network analysis, spatial-temporal analysis and visualization, multilingual text analytics, sentiment and affect analysis, and cyber attacks analysis and attribution should be considered for security informatics research [14] [15].

E. *Crime prediction using data mining*

[16], analyzed crime data using decision tree and Naïve Bayes algorithms, the accuracy was 83.9519% and 70.8124% respectively. Hence he concluded that decision tree performs better than Naïve Bayes. [17], compared Apriori algorithm, Naïve Bayes classifier and decision tree for accuracy on crime

prediction. The analysis found that Bayesian classifier yielded the best overall performance. [18], Analyzed crime data using linear regression model, additive regression model and decision stump model. The analysis was that linear regression model was very effective and accurate in predicting crime data based on training data set.

F. Review of BI Predictive Algorithms

1) Decision Tree classifier (DT)

Decision tree learning uses a decision tree as a predictive model which maps observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). It is one of the predictive modeling approaches used in statistics, data mining and machine learning. In data mining, a decision tree describes data (but the resulting classification tree can be an input for decision making). There are several popular decision tree algorithms such as ID3, C4.5, and CART (classification and regression trees). DT is in the form of a tree structure, where each node is either a leaf node (indicating the value of the target class of examples) or a decision node (specifying a test to be carried out on a single attribute value, with one branch and sub-tree for each possible outcome of the test) [19]. DTs have many advantages such as very fast classification of unknown records, easy interpretation of small-sized trees, robust structure to the outliers' effects, and a clear indication of most important fields for prediction but DTs are very sensitive to over-fitting particularly in small data-sets.

2) Multilayered perceptrons

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate outputs. MLP utilizes a supervised learning technique called back propagation for training the network. The multilayer perceptron consists of three or more layers (an input and an output layer with one or more *hidden layers*) of nonlinearly-activating nodes and is thus considered a deep neural network [20]. Each layer is made up of units. The inputs to the network correspond to the attributes measured for each training tuple. The inputs are fed simultaneously into the units making up the input layer. These inputs pass through the input layer and are then weighted and fed simultaneously to a second layer of "neuron like" units, known as a hidden layer. The outputs of the hidden layer units can be input to another hidden layer, and so on. The number of hidden layers is arbitrary, although in practice, usually only one is used. To improve the classification accuracy we should reduce the training time of neural network and reduce the number of input units of the network [21].

3) Naive Bayes classifiers

Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. Naive Bayes has been studied extensively since the 1950s. Naive Bayes models are known under a variety of names, including simple Bayes and independence Bayes. All these names reference the use of Bayes' theorem in the classifier's decision rule, but naive Bayes is not (necessarily) a Bayesian method [22].

4) Support Vector Machines

Support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall [23]. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high- dimensional feature spaces.

G. Gaps to be filled

Despite the fact that BI can play an important role in crime data analysis for decision making and strategic planning and address the issues of crime prediction, most of the information systems used in law enforcements agencies are just a collection of crime data. BI technologies have not been widely used in law enforcement agencies. This study determines the most effective BI classification technique for crime prediction. This will help law enforcement researchers in developing effective crime prediction systems.

H. Predictive Model

From the previous work done in this field and literature studied and cited above, this study identifies the most appropriate approach. These approaches form the building blocks of a conceptual model used for this research.

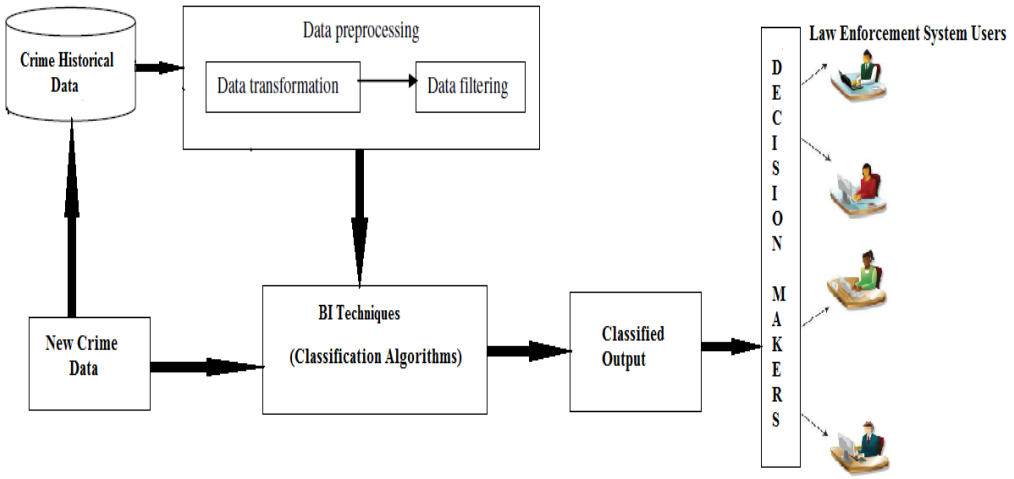

Figure 1: Block diagram of the proposed crime predictive System.

There is also a need to cleanse crime data and remove noisy data and identify the missing data before applying BI techniques. [24], state that data pre-processing techniques are mainly used for producing high-quality mining results. Since raw data can come in different format, collected from various sources and stored in the data bases and data warehouses. Some steps included to cleanse data involve extract, transform and load. This includes data cleaning, data integration, data transformation and data reduction. The preprocessed data is then passed through business intelligence techniques (classification algorithms) and then the classified output is got which is used by law enforcement stakeholders for planning and making decisions.

III. FRAMEWORK AND METHODOLOGY

The first step is to upload the dataset and then pre-process and clean the data so that it is ready for analysis. Noisy data can adversely affect the results of any data mining analysis therefore it is of essence to clean up the data [25]. The data flow diagram in Figure 3-1 shows how data is flowing from each process to the other.

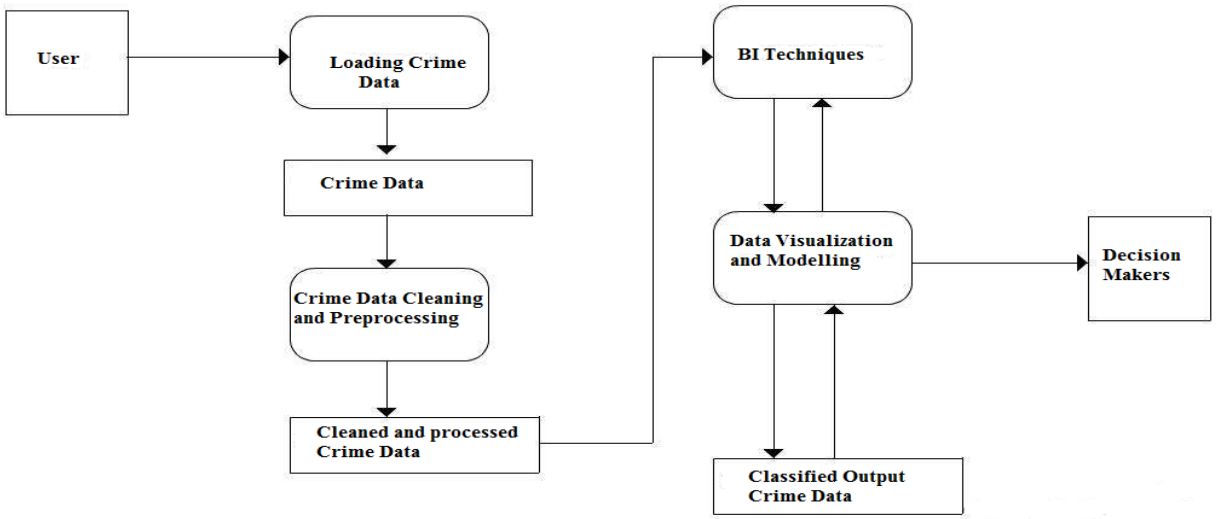

Figure 2: DFD showing how data is flowing from each process to the other.

A. The BI Techniques Used

The business intelligence model considered in our study was based on supervised learning (classification) techniques given that labeled training data was available. Classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. Also different classifiers were applied in the classification such as decision tree, naïve bayes, support vector machine and multilayer perceptrons. Our methodology consists of data collection, data-preprocessing, building classification model using training data and evaluation of the generated models using test data. Trained and tested model was then used to score incoming data.

B. Sources of Data

The dataset used for this study is real and authentic. The dataset was acquired from UCI machine learning repository website. The title of the dataset is 'Crime and Communities'. This dataset contains a total number of 128 attributes and 1994 instances. All data provided in this dataset is numeric and normalized. The complete details of all 128 attributes can be acquired from the UCI machine learning repository website [25].

1) Data Partition

The input data was randomly divided into three datasets: a training data set, test data set and validation set. The training data set was used to build the model. Model was then tested using test data to compute a realistic estimate of the performance of the model on unobserved data. We used a ratio of 70% of the data used for training, and 30% for testing.

2) Attribute selection

The objective of my analysis (crime prediction) did not require all the variables recorded hence there was need for data preparation, reduction and pre-processing. Data reduction is performed by selecting the most informative attributes in a dataset, while attempting to lose no critical information for classification [15]. There was need for removal of the variables which I did not need. From the 128 attributes only 12 were of use for the analysis. There are different methods available for attribute or feature selection but manual method is usually chosen or attribute selection based on human understanding of data set. When dealing with a large number of attributes it is practical to use human knowledge to make decisions on the attributes and also taken in account that only those attributes are chosen which do not contain any missing values.

3) Variables used in this study

State, population, MedIncome **(**Median household income**),** MedFamInc (Median family income (differs from household income for non-family households)), PerCapInc (Per capita income), NumUnderPov (Number of people under the poverty level), PctLess9thGrade (Percentage of people 25 and over with less than a 9th grade education), PctNotHSGrad (Percentage of people 25 and over that are not high school graduates), PctBSorMore (Percentage of people 25 and over with a bachelor's degree or higher education), PctUnemployed (Percentage of people 16 and over, in the labor force, and unemployed), PctEmploy (Percentage of people 16 and over who are employed), ViolentCrimesPerPop (Total number of violent crimes per 100K population), Crime Category (Crime categorization in to three categories, namely). The new added nominal attribute have three values, which are 'Low', 'Medium', and 'High'. If the value in 'Violent Crimes Per Pop' is less than 25 percent than the value of 'Crime Category' is 'Low', If the value in 'Violent Crimes Per Pop' is equal to or greater than 25 percent and less than 40 percent, than the value of 'Crime Category' is 'Medium', If the value in 'Violent Crimes Per Pop' is equal to or greater than 40 percent than the value of 'Crime Category' is 'High'.

C. Predictive Techniques and Tools Used

The BI model considered in our study was based on supervised learning (classification) technique. The software tool used was WEKA 3.8.0, an open-source and free software used for knowledge analysis and downloadable from the internet and used under the GNU license. Weka supports several standard data mining tasks, more specifically, data preprocessing, clustering, classification, regression, visualization, and feature selection. All techniques of Weka's software are predicated on the assumption that the data is available as a single flat file or relation, where each data point is described by a fixed number of attributes (normally, numeric or nominal attributes, but some other attribute types are also supported).

IV. RESULTS AND DISCUSSION

A. Predictive model/ Basic Classification Results using WEKA

In the classification we used J48, Naïve bayes, Multilayer perceptron and SVM. These classification algorithms were selected because they are considered as "white box" classification model, that is, they provide explanation for the classification and can be used directly for decision making. Each classifier belongs to a different family of classifiers implemented in WEKA. J48 relate to Decision trees, the multilayer perceptron belong to neural networks, Naïve bayes belongs to Bayesian network and SMO belong to support vector machine. Since they are from different classifiers family, they yielded different models that classify differently on some inputs. Attribute importance analysis was carried out using manual method because attribute selection is based on human understanding of data set. When dealing with a large number of attributes, it is practical to use human knowledge to make decisions on the attributes and also taken in account that only those attributes are chosen which do not contain any missing values (Witten and Frank, 2011).

B. Performance Evaluation

In the experiment, the performance analysis of J48, Naïve bayes, Multilayer perceptron and SMO algorithms was carried out, over the crime dataset. During experiment, the pre-processed dataset was converted to .ARFF file, which is the

standard file type for WEKA input. The Accuracy of J48, Naïve bayes, Multilayer perceptron and SVO is 100%, 89.7989%, 100% and 92.6724%, respectively. The figures below show the detailed results on both the training and test data.

C. Comparison of Predictive algorithms

No single learning algorithm can uniformly outperform other algorithms over all datasets. Features of learning techniques are compared in Tables below from the models built to find the most accurate and effective algorithm.

Table 1: Comparison of Predictive algorithms on training data

Algorithm	**Execution time in seconds (windows7 32 bit)**	**Accuracy (correctly classified instances)**	**Incorrectly classified instances**	**Precision**	**Recall**	**F measure**
J48	0.06sec	100%	0%	100%	100%	100%
Naïve Bayes	0.14sec	89.9425%	10.0575%	90.4%	89.9%	90.1%
Multilayer perceptron	9.26sec	100%	0%	100%	100%	100%
SVM	0.66sec	93.6782%	6.3218%	93.5%	93.7%	93.4%

Table 2: Comparison of Predictive algorithms on test data

Algorithm	**Execution time in seconds (windows7 32 bit)**	**Accuracy (correctly classified instances)**	**Incorrectly classified instances**	**Precision**	**Recall**	**F measure**
J48	0.1sec	100%	0%	100%	100%	100%
Naïve Bayes	0.103sec	89.6104%	10.3896%	89.4%	89.6%	89.3%
Multilayer perceptron	9.21sec	100%	0%	100%	100%	100%
SVM	0.28sec	92.2078%	7.7922%	91.9%	92.2%	92.8%

D. Explanation of the comparisons in tables 1 and 2

Based on the above comparisons, J48 and Multilayer Perceptron have the best accuracy, precision, recall and F measure of 100% but the execution time of J48 and Multilayer Perceptron are 0.1 and 9.21 seconds respectively. This means that J48 can predict accurately crime data within a shortest time period than Multilayer Perceptron. Also Support Vector Machine is the third in this comparison. It has the accuracy of approximately 92% which can produce relatively very good results when dealing with crime prediction. Naïve Bayes has relatively good accuracy when compared to Support Vector Machine but it executes in relatively little time when compared with Support Vector Machine. Therefore when one requires very good accurate predictions, decision tree (J48) and Multilayer Perceptron are the best options but when the execution time is a factor to consider also, then decision tree (J48) can work best. In conclusion, Based on all the benchmarks used to measure the algorithms employed in this study, it is discovered that J48 performance is better than all other algorithms.

V. CONCLUSIONS AND RECOMMENDATIONS

A. Conclusions

This study presented a comparison between four classification algorithms namely, Decision trees, the multilayer perceptron, Naïve bayes and support vector machine for predicting the ‘Crime Category’ attribute, having labels, namely ‘Low’, ‘Medium’, and ‘High’. The results of this data mining could potentially be used to lessen and even prevent crime for the forth coming years. From the encouraging results, we believe that crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis. It is evident that law enforcing agencies can take great advantage, using BI techniques like Decision Tree to effectively fight crime and war against terrorism.

B. Recommendations

This project would further be developed by incorporating real-time Business Intelligence on crime data. This implies that, a complete business intelligence system should be developed and incorporated with the best performing BI classification Algorithms like decision tree algorithm (J48) which works well on the prediction of crime data. Also Developers can take advantage of social media and email unstructured content and aggregate to structured data of information systems so as to improve analytics in the Law Enforcement Agencies on crime data.

ACKNOWLEDGMENT

The authors would like to take this opportunity to show our appreciation to all our colleagues for their support throughout the period of this research. Without such help, this work would not be completed.

REFERENCES

[1] J. M. Ngemu, O. O. Elisha, O. O. William, and M. Bernard, M. Student Retention Prediction in Higher Learning Institutions: The Machakos University College Case. International Journal of Computer and Information Technology (ISSN: 2279 – 0764) Volume 04 – Issue 02, 2015.
[2] http://www.businessdictionary.com Retrieve 10 February, 2017.
[3] S. Umair, S. Muhammad, U. Amna, M. Aniqa, B. S. Abdul and K. R. Sheikh. Application of Machine learning Algorithms in Crime Classification and Classification Rule Mining. Research Journal of Recent Sciences; ISSN 2277-2502; Vol. 4(3), 106-114.2015
[4] H. Nicole. How BI is Becoming Crime's Biggest Enemy. 2012.
[5] S. O. Adeola, S. O. Falaki and O. Olabode. E-neighborhood Management Architecture for Crime Detection and Control in Nigeria; Science and Technology, 4(2): 17-21 DOI: 10.5923/j.scit.20140402.02. 2014.
[6] S. Negash. Business intelligence. Communications of the Association for Information System. 13: 177-195. 2004.
[7] A. Malathi, B. S. Santhosh. Algorithmic Crime Prediction Model Based on the Analysis of Crime Clusters; Global Journal of Computer Science and Technology; Volume 11 Issue 11 Version 1.0.2011.
[8] M. R. Keyvanpour, M. Javideh and M. R. Ebrahimi. Detecting and investigating crime by means of data mining: a general crime matching framework, Procedia Computer Science, World Conference on Information Technology, Elsvier B.V., Vol. 3, Pp. 872-830.2010.
[9] S. Nath. Crime data mining, Advances and innovations in systems, K. Elleithy (ed.), Computing Sciences and Software Engineering, Pp. 405-409. 2011.
[10] L. P. Walter, M. Brian, C. P. Carter, C. S. Susan and S. H. John. Predictive Policing; The Role Of Crime Forecasting In Law Enforcement Operations; ISBN: 978-0-8330-8148-3. 2013.
[11] E. R. Turban, R. Sharda and D. Delen. Decision Support Systems and Intelligent Systems. 9th Edn., Prentice Hall International, New York, pp: 2. 2011.
[12] M. B. Jon. Using Business Intelligence Tools to Pursue Identity Thieves; INFSY 556. 2016.
[13] N. Perlroth and E. M. Rusli. "Security Start-Ups Catch Fancy of Investors," New York Times, Technology Section. 2012.
[14] H. Chen, W. Chung, J. Xu, G. Wang, Y. Qin and M. Chau. Crime data mining: A general framework and some examples. Computer. v. 37 n. 4, p. 50-56, ©2004 IEEE. 2004.
[15] P. L. Brantingham. "Computational Criminology," Keynote Address to the European Intelligence and Security Informatics. 2011.
[16] R. Iqbal, A. A. M. Masrah, M. Aida, H. S. P. Payam and K. Nasim. An Experimental Study of Classification Algorithms for Crime Prediction. Indian Journal of Science and Technology. | Print ISSN: 0974-6846 | Online ISSN: 0974-5645. 2013.
[17] A. Tahani, M. Rsha and L. Elizabeth. Crime Prediction Based On Crime Types And Using Spatial And Temporal Criminal Hotspots. International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.5, No.4. 2015.
[18] M. Lawrence and M. Natarajan. Using Machine Learning Algorithms to Analyze Crime Data. Machine Learning and Applications: An International Journal (MLAIJ) Vol.2, No.1. 2015.
[19] G. T. Sav. Four-Stage DEA Efficiency Evaluations: Financial Reforms in Public University Funding. International Journal of Economics and Finance, 5(1). doi:10.5539/ijef.v5n1p24. 2012.
[20] G. Cybenko. Approximation by superpositions of a sigmoidal function Mathematics of Control, Signals, and Systems, 2(4), 303–314. 1989.
[21] A. S. Rohit. Comparative Analysis of Classification Algorithms on Different Datasets using WEKA. International Journal of Computer Applications (0975 – 8887) Volume 54– No.13. 2012.
*[22] D. J. Hand and K. Yu. "Idiot's Bayes — not so stupid after all?". International Statistical Review. **69** (3): 385–399.* doi:10.2307/1403452. ISSN 0306-7734. *2001.*
[23] V. Vapnik. The Nature of Statistical Learning Theory_ Springer_ NY. 1995.
[24] K. Revatthy and J. K. Satheesh. An Attribute Weighted Fuzzy Clustering Algorithm for Mixed Crime Data. Indian Journal of Science and Technology, Vol 9(8), DOI: 10.17485/ijst/2016/v9i8/87910. 2016.
[25] M. Rouse. 'What is noisy data? - Definition from WhatIs.com' [Online]. Techtarget Available from: http://searchbusinessanalytics.techtarget.com/definition/noisy-data [Accessed 22 January 2017].
[26] UCI MachineLearning Repository. Available from: http://archive.ics.uci.edu/ml/datasets.html. Retrieved on 20 January 2017.