Kabale University Interdisciplinary Research Journal (KURJ)

OPEN ACCESS

Application of artificial intelligence in e-governance: a comparative study of supervised machine learning and ensemble learning algorithms on crime prediction

***Ivan Niyonzima**[1], **Hussein Muhaise**[1] and **Aureri Akankwasa**[2]

[1]Department of Information Technology, Kabale University, Uganda

[2]Department of Computer Science, Kabale University, Uganda

**Corresponding Author*
iniyonzima@kab.ac.ug

KURJ
ISSN 2790-1394

pp. 4 - 13
Vol 2. Issue 4.
Sept 2024

ABSTRACT

In the developing world, the daily activities of humans' social, political and economic life make it vital and easy to encounter the phenomenon of crime. Crime is an unnecessary evil in society and for any economic, social and political activities to run smoothly, criminal offenses must be completely eliminated from society. Advancement in information and communications technology enables law enforcement agencies to collect a huge amount of crime data, and the data collected by these organizaions have been doubling every two years. It has been found out that only 17% of the collected crime data is used in their operations today and several studies have noted that Law Enforcement Agencies are data rich but information poor. Machine learning, a subfield of artificial intelligence, has been used by government agencies in developed countries in different operations like face recognition, computer forensics, image and video analysis to identify criminals and crime predictions. It is therefore time for developing countries to leverage such technologies in order to reduce crimes. Therefore, this study proposes the application of supervised machine learning techniques in the prediction of crimes basing on the past crime data. During this study, we used open-source crime data from the UCI Machine learning repository to train and validate our algorithms. The performance of supervised machine learning and ensemble learning algorithms was done using crime data. The supervised machine learning algorithms used include K-Nearest Neighbour (KNN), decision tree classifier (CART), Naïve Bayes (NB) and Support vector machine (SVM). The ensemble learning algorithms used include AdaBoost (AD), Gradient Boosting Classifier (GBM), Random Forest (RF) and Extra Trees (ET). We used an accuracy metric to measure the performance of the algorithms. Python 3 was used in all the experiments using windows 10 laptop with 8GB RAM and 2.0GHZ processor.

The performance of the supervised machine learning algorithms using the original datasets includes 60.33%, 56.24%, 57.01% and 59.06% for KNN, CART, NB, and SVM respectively. The performance of ensemble learning algorithms using the original datasets includes 58.58%, 59.81%, 55.23% and 55.74% for AD, GBM, RF and ET respectively. Experimental results revealed that KNN generally performed better when compared to the rest of the algorithms. we then developed a crime prediction model based on KNN and its prediction accuracy was 66% on our test dataset.

The use of Artificial Intelligence has the potential to ameliorate several existing structural inefficiencies in the discharge of governmental functions. Machine learning, a subfield of artificial intelligence, has been used by government agencies in developed countries in crime analysis and predictions. It is therefore time for developing countries to leverage such technologies in order to reduce crimes.

Keywords: Artificial Intelligence, Crime Prediction, E-governance, Machine Learning, Ensemble learning

Background

Governance, broadly understood as the "*action or manner of governing a state*" thrives on the ability of the government to ensure efficient, effective, transparent and responsive administration. In the developing world, the daily activities of humans' social, political and economic life makes it vital and easy to encounter the phenomenon of crime. It has been estimated that over 90% of the data that exists in the world today has been created over the last two years alone and crime data is not exceptional as it also comes in many formats (e.g., videos, images, audios, satellite data, and sensor data). Intelligently analyzed data can assist decision makers to make actionable data driven decisions and therefore it is a valuable resource in the era of big data. It can lead to new insights and, in commercial settings, to competitive advantages.

Law enforcement has always relied on intelligence information enhanced by analysis to combat all crime and identify threats; however, the information is often narrowly focused and inconsistently updated or shared. The rise of digital technologies has made possible more powerful methods for collecting, analyzing and sharing information and has fostered the development of Intelligence-Led Policing. Global business data is becoming an essential component for Law Enforcement and Intelligence Agencies which increasingly rely on Intelligence-Led Policing (ILP) strategies. Commercial business data can help law enforcement uncover money laundering schemes, financial fraud, illegal business fronts, and a variety of other criminal activities. Global business data assists law enforcement in proactively assessing and monitoring threats and exposing businesses and executives that are involved in nefarious activity (Dun and Bradstreet, 2012). According to (OSAC Report, 2017), crimes in Uganda can occur anywhere at any time. The report shows that there was an increase of crime activity in the central region of Kampala and the northern region specifically Gulu and lira. These included both serious and moderate crimes and there was moderate cybercrime.

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 (Ahishakiye, Omulo, Taremwa, *et al.*, 2017). The use of Artificial Intelligence has the potential to ameliorate several existing structural inefficiencies in the discharge of governmental functions (Basu *et al.*, 2018). In fact, scientists are spending time studying crime and criminal behaviors in order to understand the characteristics of crime and to discover crime patterns. In particular, issues arise as to how to choose accurate techniques for analyzing data due to the inconsistency and inadequacy of these kinds of data. These issues motivate scientists to conduct research on these kinds of data to enhance crime data analysis. Dealing with crime data is very challenging as the size of crime data grows very fast, so it can cause storage and analysis problems. 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. 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 (Ahishakiye, Omulo, Taremwa, *et al.*, 2017). Therefore, the major objective of this study was to perform a comparative study of classification algorithms and ensemble methods in crime prediction. Experimental results helped us to determine which algorithm that works better, and we developed, trained and validated a crime prediction model based on KNN.

Related Work

The study by (Ahishakiye, Omulo, Taremwa, *et al.*, 2017) did a study on crime prediction using decision tree (J48) algorithm. They used dataset from UCI machine learning repository website. The title of the dataset is '*Crime and Communities*'. The experimental results revealed that J48 algorithm predicted the unknown category of crime data to the accuracy of 94.25287%.

The study by (Kim *et al.*, 2019) did a study on crime analysis using machine learning techniques. The study used Vancouver crime datasets. Experimental results revealed that K-nearestneighbour and boosted decision tree algorithms had an accuracy between 39% to 44%.

The study by (Ahishakiye, Omulo, Wario, *et al.*, 2017) did a performance study of machine learning algorithms on crime prediction. The dataset used was secondary data from UCI machine learning repository website. Experimental results 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 on test data.

The study by (Wu *et al.*, 2020) did a study on crime prediction using data mining and machine learning techniques. The study revealed that the classification effect of Random Trees is better than that of Neural Networks and Bayesian Networks. However, the performance accuracy of individual algorithms was not reported.

The study by (Toppireddy *et al.*, 2018) did a study on crime prediction and monitoring based on spatial analysis. In their work, various visualizing techniques and machine learning algorithms are adopted for predicting the crime distribution over an area. machine learning algorithms were used to extract the knowledge out of these large datasets and discover the hidden relationships among the data. However, the performance accuracy of individual algorithms was not reported.

The study by (Rumi *et al.*, 2018) did a study on crime prediction using dynamic features. The study revealed that dynamic information was very sparse compared to the relatively static information. To address this issue, the study developed a matrix factorization based approach to estimate the missing dynamic features across the city. Experimental results revealed that the crime prediction performance can be significantly improved with the inclusion of dynamic features across different types of crime events.

The study by (Grover *et al.*, 2007) did a review of crime prediction techniques. The study revealed three techniques namely: statistical methods, these mainly relate to the journey to crime, age of offending and offending behaviour; techniques using geographical information systems that identify crime hot spots, repeat victimisation, crime attractors and crime generators; a miscellaneous group which includes machine learning techniques to identify patterns in criminal behaviour and studies involving reoffending. The majority of current techniques involve the prediction of either a single offender's criminality or a single crime type's next offence. For further details, refer to (Ahishakiye, Omulo, Taremwa, *et al.*, 2017) (Kim *et al.*, 2019)- (Grover *et al.*, 2007).

Methods

Data sources
The dataset used in this study was obtained from the UCI Machine Learning repository and the description of the dataset was well explained on the platform. Python 3.7 notebook was used for all the implementations on 64-bit windows 10, 2GHZ processor computer.

Data Preparation
During this study, we used 492 samples with 8 variables. some variables that were deemed irrelevant in crime prediction were removed. The dataset had some missing values which were imputed using the most frequent value (mode) of that particular column. After this step, the correlation was also done on the dataset to determine which variables were related to our outcome or target variable.

Performance Metric Used
Classification Accuracy was used during this study and it is the most common evaluation metric for classification problems. To calculate Classification accuracy, we divide the number of correct predictions by the total number of instances. It is therefore reported as a ratio that can be converted into a percentage by multiplying the value by 100. In order for accuracy to hold any substantial value, the dataset must contain an equal number of instances belonging to each class. If the dataset is unbalanced, accuracy will be affected (Brownlee, 2016).

$$Accuracy = \frac{Correct\ Predictions}{Total\ Instances}$$

Algorithms Selection
Basing on our dataset, we selected classification algorithms and ensemble methods. Classification algorithms were selected because most of the variables in our dataset were categorical in nature and therefore the selected algorithms would work well, while ensemble methods were selected since they consolidate a few machine learning techniques into one model so as to diminish variance, bias, or improve performance. Also, ensemble methods are not sensitive to data distributions and therefore work well even when data is not standardized. The following section discusses Ensemble methods and classification algorithms that were used during this study.

Ensemble Methods
Ensemble methods use a combination of techniques that allows multiple machine learning models, called base learners, sometimes called weak learners, to merge their predictions and output a single, optimal prediction, given their respective inputs and outputs (see algorithm 1). Ensembles attempt to solve two issues, specifically, bias and variance, as well as the relationship between them. Ensemble methods are divided into two major classes or taxonomies: generative and non-generative methods. Non-generative methods are focused on combining the predictions of a set of pre-trained models. Generative methods, on the other hand, are able to generate and affect the base learners that they use. They can either tune their learning algorithm or the dataset used to train them, in order to ensure diversity and high model performance. During this study, we focused on generative methods namely Bagging and Boosting algorithms, and we discussed them in detail.

Algorithm1: General Pseudo-code for Ensemble Methods (Adapted from (*Ensemble Methods*, n.d.)**)**

Given M as the original training data, p as the number of base classifiers, and N as the test data.
for i=1 to p ***do***
    Create a training set M_i from M.
    Build a base classifier K_i from M.
end for
for each test record n ∈ N ***do***
    $K^*(n) = Vote(K_1(n), K_2(n),..., K_p(n))$
end for

Bagging Methods

Bagging (or Bootstrap Aggregation) uses various examples from the training dataset (with replacement) and training a model for each example. The final output prediction is found by taking the mean across the predictions of all of the sub-models (Breiman, 2001). Bagging uses bootstrap sampling to reduce variance just as the progress of the accuracy (Zhang & Ma, 2012). The usage of the bagging methods improves the classification results at whatever point the base classifiers are unstable, this being the vital inspiration why the bagging approach works splendidly for classification (Zhang & Ma, 2012). For more details, refer to (Zhang & Ma, 2012). Two Bagging algorithms were used during this study, namely; Random Forests (Gislason et al., 2006)(Scornet, 2010) (we used the python library sklearn's implementation of a random forest) and Extra Trees (Geurts et al., 2006) (we used the python library sklearn's implementation of Extra Trees).

Boosting Algorithms

The first boosting technique was proposed by Schapire in (Schapire, 1990), where the key result is that the weak and strong learnability are equivalent, in the sense that strong learning can be performed by combining weak learners. "Boosting" is a general technique for improving the performance of any learning algorithm. In principle, boosting can be utilized to on a very basic level lessen the error of any "weak" learning algorithm that reliably produces classifiers that need simply be fairly better than random guessing (Freund & Schapire, 1996). For more details, refer to (Freund & Schapire, 1996). Two Boosting algorithms were used during this study, namely; AdaBoost (Schapire, n.d.) (We used the python library sklearn's implementation of AdaBoost) and Stochastic Gradient Boosting (Friedman, 1999) (We used the python library sklearn's implementation of Gradient Boosting classifier).

Classification Algorithms

Classification techniques have been seen as a key bit of machine learning, with a tremendous proportion of applications published over the most recent couple of years (Pérez-Ortiz et al., 2016) (Kotsiantis et al., 2007). The target of classification algorithms is to isolate the classes of the problem by using the training data. In the event that the output variable has two possible values, the problem is referred to as binary classification. Then again, if there are multiple classes, the problem is named multiclass or multinomial classification.

K-Nearest Neighbor (KNN)

The study by Alkhatib et al. of 2013 as cited by (Vainionpää & Davidsson, 2014) revealed that the KNN algorithm is viewed as a lazy learning algorithm, with a low computational cost and it is easy to execute. KNN is a case of nonparametric models on the grounds that the viable number of parameters is unbounded, for example it develops with the number of examples. This methodology is known as instance-based learning or memory-based learning. Prediction with KNN is computed as follows (Alkhatib *et al.*, 2013).

I. Determine the quantity of closest neighbors, k.
II. Compute the separation between the training samples and the query record.
III. Sort all training records as per the distance values.
IV. Use a dominant part vote in favor of the class labels of k nearest neighbors, and allot it as a prediction value of the query record.

We used the python library sklearn's implementation of KNN (KNeighborsClassifier).

Support Vector Machine (SVM)

SVM is one of the best binary classifiers (Pérez-Ortiz *et al.*, 2016) (Madge, 2015) (Min & Lee, 2005) that limits their choice such that most points in one class fall on one side of the boundary while most points in the other category fall on the opposite side of the boundary. During this study, RBF kernel was used. The main advantage of RBF Kernel is that it can deal with different input sets. Furthermore, it classifies test examples dependent on the example's Euclidean distance to the training points, and weights closer training points all the more intensely. This implies classification depends intensely on the most similar training examples and exploits patterns in the data (Madge, 2015). We used the python library sklearn's implementation of SVM (SVC).

Naïve Bayes Classifiers

Naive-Bayes classifier assumes class conditional independence. Given the test data Bayesian classifier predicts the likelihood of data having a place with a specific class, to foresee likelihood it utilizes the idea of Bayes' theorem. Bayesian classifiers additionally fill in as a theoretical legitimization for different classifiers that don't unequivocally utilize Bayes' theorem. For instance, under explicit suppositions, it very well may be shown that numerous neural networks and curve-fitting algorithms output the maximum posterior hypothesis, as does the naive Bayesian classifier (Predicting Direction of Movement of Stock Price and Stock Market Index, 2012) (Huang & Liu, 2019) (Ren et al., 2009). We used the python library sklearn's implementation of Naïve Bayes (GaussianNB).

Classification and Regression Trees (CART)

CART algorithm can be utilized to build both Classification and Regression Decision Trees (Zacharis, 2018). For the most part, for any classification or regression problem, CART algorithm has three significant undertakings: (i) how to divide the data at each step, (ii) when to stop dividing the data, and (iii) how to forecast the value of y for each x in a partition. During this study, we dealt with a classification problem (binary classification-whether one has cancer or not) and therefore, we used a decision tree classifier for CART. We used the python library sklearn's implementation of CART for classification (DecisionTreeClassifier).

Experimental Results

During this study, our dataset was divided into two parts: the training set and the test set. The proportion of the training set was 80% while 20% was for testing. Also, the 10-fold cross-validation was used. The algorithms were evaluated using the accuracy metric. Accuracy is defined as the number of correct predictions made as a ratio of all predictions made (Brownlee, 2016). This metric was used because it gives a quick idea of how good a given model is, and it also works well on binary classification problems.

Performance of the Classification Algorithms

In this section, we represent the performance of our classification algorithms in crime prediction. The percentage accuracy is obtained by multiplying the mean and standard deviation scores by 100. Table 1 below shows the individual performance of each of the classification algorithms. from the table1, we note that K-Nearest Neighbor (KNN) outperformed other classification algorithms in crime prediction with accuracy of 60.33%.

Table 1: Performance of the Classification Algorithms.

	Classification Accuracy	
Algorithm	**Mean**	**Standard Deviation**
K-Nearest Neighbor (KNN)	60.33%	0.112616
Classification and Regression Trees (CART)	56.24%	0.077752
Naïve Bayes Classifier (NB)	57.01%	0.076089
Support Vector Machine (SVM)	59.06%	0.092983

Comparison of performance of the Classification Algorithms used

In this section, we compared the performance of K-Nearest Neighbor (KNN), Classification and Regression Trees (CART), Naïve Bayes Classifier (NB) and Support Vector Machine (SVM) using box and whisker plot. Box and whisker plots present and effective way to compare the accuracy of more than one machine learning algorithms. The yellowish line shows the mean accuracy score.

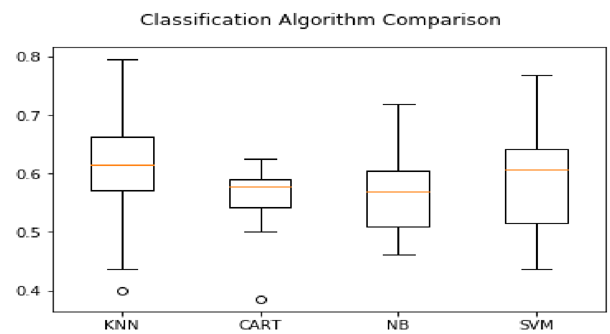

Figure 1: Box and Whisker Plots Comparing Classification Algorithm Performance.

Performance of the Ensemble Methods

In this section, we present the performance of ensemble methods using classification accuracy. Also, the percentage accuracy is gotten by multiplying each score by 100. From Table 2 below, boosting algorithms outperformed bagging algorithms, with Stochastic Gradient Boosting having a higher accuracy of 59.81%.

Table 2: Performance of the Ensemble Methods

	Classification Accuracy	
Algorithm	**Mean**	**Standard Deviation**
Random Forests (RF)	55.23%	0.090160
Extra Trees (ET)	55.74%	0.058384
AdaBoost (AD)	58.58%	0.092292
Stochastic Gradient Boosting (GBM)	59.81%	0.059553

Comparison of performance of the Ensemble Methods used

In this section, we compared the performance of Random Forests (RF), Extra Trees (ET), AdaBoost (AD), and Stochastic Gradient Boosting (GBM) using box and whisker plot. The yellowish line shows the mean accuracy score. It can be noted that Stochastic Gradient Boosting had higher accuracy when compared to the rest of the algorithms.

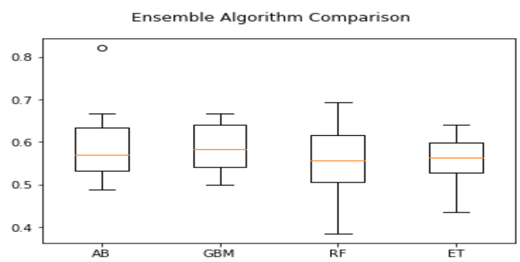

Figure 2: Box and Whisker Plots Comparing Ensemble Methods Performance.

Final Prediction Model

Based on the performance of our algorithms in Tables 1 and 2, we trained and tested our final model using KNN because it had a better performance when compared to the rest of the algorithms. The model was trained using the entire training dataset and the test dataset was used to confirm our findings. The performance of our final model had an accuracy of 66%.

Discussions and Conclusion

In this study, we used supervised machine learning algorithms and ensemble methods in crime prediction. Experimental results revealed that K-Nearest Neighbor (KNN), a classification algorithm outperformed Classification and Regression Trees (CART), Naïve Bayes Classifier (NB), Support Vector Machine (SVM), Random Forests (RF), Extra Trees (ET), AdaBoost (AD) and Support Vector Machine (SVM). We finally built and trained our final KNN prediction model and it had a performance accuracy of 66% on our test data. The use of Artificial Intelligence has the potential to ameliorate several existing structural inefficiencies in the discharge of governmental functions. Machine learning, a subfield of artificial intelligence, has been used by government agencies in developed countries in different operations like face recognition, computer forensics, image and video analysis to identify criminals and crime predictions. It is therefore time for developing countries to leverage such technologies in order to reduce crimes.

References

Ahishakiye, E., Taremwa, D., Omulo, E. O., & Niyonzima, I. (2017). Crime prediction using decision tree (J48) classification algorithm. *International Journal of Computer and Information Technology*, 6(3), 188-195.

Ahishakiye, E., Omulo, E. O., Wario, R., & Niyonzima, I. (2017). A Performance Analysis of Business Intelligence Techniques on Crime Prediction. *International Journal of Computer and Information Technology*, 6(2), 84–90. https://doi.org/10.1.1.404.1245

Alkhatib, K., Najadat, H., Hmeidi, I., & Shatnawi, M. K. A. (2013). Stock Price Prediction Using K-Nearest Neighbor (KNN) Algorithm. *International Journal of Business, Humanities and Technology*, 3(3), 32–44.

Basu, A., Hickok, E., Mohandas, S., Kundu, A., Sinha, A., Pranav, M. B., & Ramachandran, V. (2018). *Artificial Intelligence in the Governance Sector in India*. The Centre for Internet and Society, India.

Breiman, L. (2001). *Random Forests. Statistics Department*, University of California Berkeley, CA 94720, 1–33.

Brownlee, J. (2016). *Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models and Work Projects End-To-End*.

Ensemble Methods. (n.d.).

Freund, Y., & Schapire, R. E. (1996). *Experiments with a New Boosting Algorithm*.

Friedman, J. H. (1999). Stochastic Gradient Boosting. *CSIRO CMIS, Locked Bag* 17, North Ryde NSW 1670.

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Mach Learn*, June 2005. https://doi.org/10.1007/s10994-006-6226-1

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random Forests for land cover classification. *Pattern Recognition Letters*, 27, 294–300. https://doi.org/10.1016/j.patrec.2005.08.011

Grover, V., Adderley, R., & Bramer, M. (2007). Review of Current Crime Prediction Techniques. A*pplications and Innovations in Intelligent Systems* XIV, 233–237. https://doi.org/10.1007/978-1-84628-666-7_19

Huang, C., & Liu, Y. (2019). Machine Learning on Stock Price Movement Forecast : The Sample of the Taiwan Stock Exchange. *International Journal of Economics and Financial Issues*, 9(2), 189–201. https://doi.org/10.32479/ijefi.7560

Kim, S., Joshi, P., Kalsi, P. S., & Taheri, P. (2019). Crime Analysis Through Machine Learning. 2018 *IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON* 2018, 415–420. https://doi.org/10.1109/IEMCON.2018.8614828

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2007). Machine learning : a review of classification and combining techniques. *Artif Intell Rev,* 2006, 159–190. https://doi.org/10.1007/s10462-007-9052-3

Madge, S. (2015). *Predicting Stock Price Direction using Support Vector Machines*.

Min, J. H., & Lee, Y. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28, 603–614. https://doi.org/10.1016/j.eswa.2004.12.008

Pérez-Ortiz, M., Jiménez-Fernández, S., Gutiérrez, P. A., Alexandre, E., Hervás-Martínez, C., & Salcedo-Sanz, S. (2016). A Review of Classification Problems and Algorithms in Renewable Energy Applications. *Energies - MDPI*, 1–27. https://doi.org/10.3390/en9080607

Predicting Direction of Movement of Stock Price and Stock Market Index. (2012). 11–38.

Ren, J., Lee, S. D., Chen, X., Kao, B., Cheng, R., & Cheung, D. (2009). Naive Bayes Classification of Uncertain Data. 2009 *Ninth IEEE International Conference on Data Mining*. https://doi.org/10.1109/ICDM.2009.90

Rumi, S. K., Deng, K., & Salim, F. D. (2018). Crime event prediction with dynamic features. *EPJ Data Science*, 7(1). https://doi.org/10.1140/epjds/s13688-018-0171-7

Schapire, R. E. (n.d.). Explaining AdaBoost. In Princeton University, Dept. of Computer Science, 35 Olden Street, Princeton, NJ 08540 USA.

Schapire, R. E. (1990). The Strength of Weak Learnability. *In Machine Learning*, 227, 197–227.

Scornet, E. (2010). A Random Forest Guided Tour. *Mathematics Subject Classification*: 62G05, 62G20, 1–41.

Toppireddy, H. K. R., Saini, B., & Mahajan, G. (2018). Crime Prediction & Monitoring Framework Based on Spatial Analysis. *Procedia Computer Science*, 132(Iccids), 696–705. https://doi.org/10.1016/j.procs.2018.05.075

Vainionpää, I., & Davidsson, S. (2014). *Stock market prediction using the K Nearest Neighbours algorithm and a comparison with the moving average formula (Issue April)*. KTH Royal Institute of Technology.

Wu, S., Wang, C., Cao, H., & Jia, X. (2020). Crime prediction using data mining and machine learning. *Advances in Intelligent Systems and Computing*, 905(Ml), 360–375. https://doi.org/10.1007/978-3-030-14680-1_40

Zacharis, N. Z. (2018). Classification and Regression Trees (CART) for Predictive Modeling in Blended Learning. I.J. *Intelligent Systems and Applications*, 3(March), 1–9. https://doi.org/10.5815/ijisa.2018.03.01

Zhang, C., & Ma, Y. (2012). *Ensemble Machine Learning: Methods and Application*s. Springer New York Dordrecht Heidelberg London. https://doi.org/10.1007/978-1-4419-9326-7