A Comparative Analysis of Artificial Intelligence Techniques in Forecasting Violent Crime Rate

Alif Ridzuan Khairuddin¹, Razana Alwee¹ and Habibollah Haron¹

¹Applied Industrial Analytics Research Group (ALIAS), School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310, Johor Bahru, Johor, Malaysia

E-mail: alifjam1991@gmail.com

Abstract. The increase in the occurrence of violent crimes is a major concern in all countries around the globe. Various approaches of crime analyses have been implemented in reducing the number of violent crimes and among them is crime forecasting. Crime forecasting is an effective solution as it assists law enforcement agencies in planning efficient crime prevention measures. It has been observed recently that the application of artificial intelligence (AI) techniques in crime forecasting and analysis is favoured by researchers. Motivated by this development, this study aims to conduct a comparative analysis on the forecasting performance of three artificial intelligence (AI) techniques, namely artificial neural network (ANN), support vector regression (SVR), and gradient tree boosting (GTB) in forecasting the rates of four types of crimes in the United States (US). The forecasting performance of each AI technique was compared in terms of quantitative error measurement. From the results obtained, GTB showed the highest forecast accuracy compared to ANN and SVR as the observed error measurements were the smallest.

1. Introduction

The recent increase in crime rates especially violent crimes has been a worrying factor and a major concern for all countries around the globe. Crimes cause immense financial losses and inflict injury upon various parties and individuals [1]. Crimes are also a huge threat on the stability of communities and societies. Recent unpredictable economic and political conditions also serve as catalysts towards the occurrence of violent crimes. In the effort to reduce violent crimes, various approaches of crime analyses have been implemented and proposed by researchers and criminologists to analyze and observe violent crime patterns.

Among the applied approaches, crime forecasting or prediction has become a popular approach that not only analyses the violent crime pattern but is also able to extrapolate the potential violent crime occurrence into the future. The advantage of crime forecasting is that it can assist many federal and state law enforcement agencies such as police forces and security services by providing important information in planning and managing efficient crime prevention measures. Prior studies have shown that the real world crime data consists of complex nonlinear data structures with different representations. These characteristics provide a challenge to researchers in identifying an appropriate model or technique to handle such complex data.

However, the rapid development of artificial intelligence (AI) techniques in the past few years has resulted in them being favored and studied by researchers recently. One advantage of AI is that it
contains nonlinear functions which are able to identify nonlinear patterns in data and thus improving the overall forecasting performance [2]. Motivated by this fact, the aim of this study is to apply, compare and analyze the forecasting performance of three selected AI techniques, namely artificial neural network (ANN), support vector regression (SVR), and gradient tree boosting (GTB) in forecasting violent crime rates data in the US.

The remainder of this paper is organized as follows. Section 2 provides an overview of applied artificial intelligence techniques in crime forecasting. Section 3 explains the experimental conduct of this study. Section 4 discusses the performance result of each AI technique based on the conducted experiment. Finally, Section 5 provides the conclusion of the study.

2. Overview of Applied Artificial Intelligence (AI) Technique in Crime Forecasting

In forecasting crime using the AI technique, researchers mostly implement machine learning techniques to predict and estimate the value of target data (crime rate). The AI technique has been a popular option for researchers, experts and data analysts to analyze and forecast various data sets due to its robustness and adaptability. Another reason is that the AI technique is able to learn and model a complex and non-linear relationship of data [3].

As mentioned before, real world crime data structures exist in many forms, distribution and representations. Thus, the AI technique is a highly recommended solution in analyzing and forecasting crime rate data without worrying about the data structure limitations. In this study, three AI techniques were selected to forecast violent crime rates i.e. artificial neural network (ANN), support vector regression (SVR), and gradient tree boosting (GTB). The forecast accuracy performance of each selected AI technique was then compared and observed.

2.1. Artificial Neural Network (ANN)

Artificial neural network (ANN) is an AI technique which was inspired by the human nervous system. The structure of ANN comprises many interconnected layers of neurons, each with its own weight matrix, bias vector and output vector that later form a massive parallel network [4]. Neurons in ANN act as processing units or a local memory that receive information from other neurons, then process and transmit it to other neurons. The application of ANN in crime forecasting has been widely implemented by researchers for many years due to its ability to produce results which are not restricted to the provided input data due to its self-learning capabilities [5-7].

2.2. Support Vector Regression (SVR)

Support vector regression (SVR) is an AI technique that implements a symmetrical loss function to estimate a real-value function [4]. It was inspired from the linear regression function computation in high dimensional feature space where the input data are mapped via a nonlinear function [8]. The advantage of SVR is that it is able to generalize the unseen data and offers flexibility in handling various data representations provided with appropriate kernel functions [9]. SVR is one of the most popularly applied AI techniques in crime forecasting due to its excellent generalization capability with high prediction accuracy [10-12].

2.3. Gradient Tree Boosting (GTB)

Gradient tree boosting (GTB) is an AI technique introduced by [13] which is based on a combination of decision trees and boosting techniques. In GTB, the decision tree acts as a base learner while the boosting technique is used to reduce the error during the decision tree learning process. The boosting process is performed iteratively where each boosting iteration improves the tree accuracy of the previous iteration [14]. The advantage of GTB is that it is able to avoid the overfitting problem when new independent data are added [11]. In crime forecasting, GTB can be considered new and studies on the application of GTB in crime analysis are minimal [15-16].
3. Experimental Conduct

In this study, each AI-based crime type model is modelled by using a Statistics and Machine Learning Toolbox on Matlab platform. The AI-based crime model (ANN, SVR and GTB) is analyzed using a multivariate analysis where several factors that potentially influence crime rates are considered. In addition, it is focused on solving the regression problem where the crime rate values were forecast for each crime type data.

Before the crime rate value is forecast, the crime model is trained using the prepared training data set for data fitting. Once the crime model has been trained or fitted, it is then used to forecast the crime rate values using the prepared testing data set. The forecast crime rate value result is then used to evaluate the model performance for each AI crime model. In this study, three types of quantitative error measurement analyses namely root mean square error (RMSE), mean absolute deviation (MAD), and mean absolute percentage error (MAPE) were used in calculating and measuring the forecasting performance of each AI crime model.

3.1. Data Collection

In this study, two types of data sets were used i.e. crime rate and factors data sets. In the crime rate data set, four types of US violent crime rates namely murder and non-negligent manslaughter, forcible rape, aggravated assault, and robbery were used. The crime rate data set was obtained from the Uniform Crime Reporting Statistics website provided by the Federal Bureau of Investigation of the United States (US). For the factors data set, nine data series namely unemployment rate, immigration rate, population rate, consumer price index, gross domestic product, tax revenue, poverty rate, inflation rate, and consumer sentiment index were selected and used. The factors data set was obtained from numerous United States government agencies and other related data repository websites. Each data series of both data sets consists of 56 data samples of annual time series data that were collected from 1960 to 2015.

3.2. Data Preparation and Processing

During the experiment, the crime rate and factors data sets were divided into two groups of training (in-sample) and testing (out-sample) data. Training data is used to train each crime model while testing data is used to test and forecast the crime rate values based on the trained crime model. In this study, the data was divided into a ratio of 9:1 where the training data contained 50 data samples from 1960 to 2009 while the testing data contained six data samples from 2010 to 2015.

In order to avoid any unexpected error due to different measurement units, each data set was processed and transformed into a dimensionless form using the normalization technique. The raw data sets (crime rate and factors) were normalized by using the feature scaling method in a scale range between 0 and 1. The data normalization is defined in the following equation (1).

\[
x' = \frac{x - \text{min}_x}{\text{max}_x - \text{min}_x}
\]

In equation (1), \(x\) is a raw value of the selected sample in the respective data series \(x\), \(\text{max}_x\) is the largest value of raw data in the respective data series \(x\), \(\text{min}_x\) is the smallest value of raw data in the respective data series \(x\), and \(x'\) is the normalized value for the corresponding sample \(x\). After the forecasting process has been conducted, the dimensionless form of the normalized forecast crime rate values for each crime model are transformed back into the actual raw crime rate values through denormalization. The denormalization calculation is based on a mathematical transformation from equation (1) and it is expressed in the following equation (2).

\[
x = (\text{max}_x - \text{min}_x) + \text{min}_x
\]
3.3. Parameter Setup for Each Artificial Intelligence Technique

The configuration of several input parameters for each selected AI technique in modelling the crime model was performed before crime model modeling and forecasting took place. In the ANN parameters, the type, training function, number of neurons and layers were set to feed-forward back-propagation, Levenberg-Marquadt, 10 and 2 respectively. For the SVR parameters, the kernel function, epsilon value, and optimization solver method were set to Gaussian, 0.1, and sequential minimal optimization, respectively. Lastly, in the GTB parameters, the number of trees, size of individual trees, and learning rate were set to 100, 3 and 0.1 respectively.

4. Results and Discussion

The collected forecast crime rate values of the AI crime models for each crime type were calculated using the quantitative error measurement analysis and the calculated result is presented in table 1.

| Crime Type               | AI Model | Quantitative Error Measurement |
|-------------------------|----------|-------------------------------|
|                         |          | RMSE | MAD | MAPE        |
| Murder and Non-negligent Manslaughter | ANN      | 0.8022 | 0.7770 | 16.4609    |
|                         | SVR      | 0.7043 | 0.6765 | 14.6189    |
|                         | GTB      | 0.4755 | 0.4451 | 9.4824     |
| Forcible Rape           | ANN      | 6.9366 | 5.6390 | 20.8053    |
|                         | SVR      | 4.7445 | 4.6468 | 17.0807    |
|                         | GTB      | 4.6061 | 4.4043 | 16.3718    |
| Robbery                 | ANN      | 38.3331 | 37.7101 | 34.4109    |
|                         | SVR      | 27.1325 | 26.2597 | 24.3614    |
|                         | GTB      | 25.8524 | 24.5894 | 22.8913    |
| Aggravated Assualts     | ANN      | 69.9341 | 64.7358 | 27.1261    |
|                         | SVR      | 40.2510 | 37.4862 | 15.7069    |
|                         | GTB      | 37.0387 | 35.6826 | 14.8152    |

The results in table 1 demonstrate that GTB possesses the best forecasting performance capabilities compared to ANN and SVR in this case study. This is proven by its smallest error measurement values in RMSE, MAD and MAPE compared to the other AI techniques in all modelled crime types. This shows that among the AI models, GTB is able to efficiently estimate and forecast crime rates for each crime type despite the limited data samples provided in the study. Hence, it is proven that the GTB is robust even for a small data size and is able to produce a more accurate forecast. In contrast, ANN performed the worst compared to the other AI techniques as the observed RMSE, MAD and MAPE values for all crime types are highest. This indicates that ANN performs poorly with small crime data samples. Overall, GTB shows the best result with the best forecast accuracy, followed by SVR in second place and finally ANN is the worst model in forecasting US violent crime rates for this case study.
5. Conclusion
The applications of AI techniques in crime forecasting and analyses have become popular recently due to their robustness in handling various types of real-world crime data structures and representations. Such an advantage has led to researchers focusing more on the application of AI in criminology studies recently. Various approaches have been introduced by researchers in modelling an efficient AI-based crime model to accurately forecast crime for their case studies. In this study, three AI techniques namely ANN, SVR and GTB were applied to forecast four types of US violent crime rates; murder and non-negligent manslaughter, forcible rape, aggravated assaults, and robbery. The results showed that that GTB outperformed the other AI techniques as shown by its smallest quantitative error measurement values. This shows that GTB is considered more appropriate in handling limited time series crime rate data as compared to ANN and SVR.

Acknowledgements
This work was funded and supported by Universiti Teknologi Malaysia (UTM) under the UTM Encouragement Research Grant with grant number Q.J130000.2651.18J46 for School of Computing, Universiti Teknologi Malaysia (UTM).

References
[1] Kadar C, Maculan R and Feuerriegel S 2019 Decis. Support Syst. 119 pp 107-17
[2] Rather A M, Sastry V and Agarwal A 2017 OPSEARCH 54 pp 558–79
[3] Baliyan A, Gaurav K and Mishra S K 2015 Procedia Comput. Sci. 48 pp 121-25
[4] Awad M and Khanna R 2015 Efficient learning machines: theories, concepts, and applications for engineers and system designers (Berkley: Apress)
[5] Huang Y L, Lin C T, Yu Y S, Hsieh W H and Pai S M 2015 Proc. of Ntl. Conf. on Information Technology Practice and Application (China)
[6] Corcoran J J, Wilson I D and Ware J A 2003 Int. J. of Forecasting 19 pp 623-34
[7] Wang Q, Jin G, Zhao X, Feng Y and Huang J 2019 Knowl-Based Syst. pp 105-20
[8] Wu J and Lu Z 2012 Proc. of 5th Int. Conf. on Advanced Computational Intelligence (Nanjing: IEEE) pp 999-1003
[9] Basak D, Pal S and Patranabis D C 2007 Neural Info. Process.-Lett. and Rev. 11 pp 203-24
[10] Kianmehr K and Alhajj R 2006 Proc. of Int. Conf. on Computer Systems and Applications (Dubai: IEEE) pp 952-59
[11] Alwee R, Shamsuddin H, Mariyam S and Salehuddin R 2013 The Scientific World J. 2013 951475
[12] Yang F, Wu C, Xiong N and Wu Y 2018 Int. J. Softw. Hardw. Res. Eng. 6 pp 1-10
[13] Friedman J H 2001 Ann. Statist. 29 pp 1189-232
[14] Budur E, Lee S and Kong V S 2015 Social and Information Networks arXiv:1507.05739
[15] Shi X, Paiement J F, Grangier D, and Yu P S 2012 Proc. of the 2012 SIAM Int. Conf. on Data Mining (California: SIAM) pp 224-35
[16] Sumner C, Byers A, Boochever R and Park G J 2012 Proc. of 11th Int. Conf. on Machine Learning and Applications 2 (Washington: IEEE Computer Society) pp 386–93