Using AI Techniques To Predict Property Crime Rates - A Comparison And Analysis

Rebaz Mala[a] Nabi, Soran Ab. M. Saeed[b], Habibollah Haron[c]

[a]Computer Network, Sulaimani Polytechnic University, Sulaimani, KRG, Iraq
[b]Sulaimani Polytechnic University, Sulaimani, KRG, Iraq
[c]Qaiwan International University, Sulaiman, KRG, Iraq

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Property crime is an increasing concern worldwide and one which causes trauma to victims and puts pressure on law enforcement agencies. One essential weapon in the fight against property crime is that of effective forecasting which helps law enforcement task forces to put in place effective preventative measures. Researchers and law enforcers now choose to harness the power of modern technology, namely artificial intelligence, to help them to predict property crime rates and to therefore create proactive preventative solutions. The objective of this study is to perform a comparative analysis on three different artificial intelligence techniques which are; Random Forest Classifier (RFC), Gradient Tree Boosting (GTB) and Support Vector Regression (SVR). These techniques are applied to four separate crime types in the USA in order to compare and contrast in terms of quantitative measurement of error. The results of the study show that GTB is the most effective method as it produced the lowest error measurements and the highest level of forecast accuracy in comparison to RFC and SVR

1. Introduction

Worldwide, property crime is on the rise and, although crimes such as burglary decreased during the pandemic, rates remain high enough to be a pressing concern on a global scale. In addition to the personal trauma of these crimes, property crime also poses a significant threat to the financial stability of communities [1]. After the initial lockdown period of the pandemic, the ensuing uncertainty within economies and politics has acted as an accelerator for the occurrences of property crimes. In an effort to tackle this issue, criminologists and criminal research teams worldwide have conducted a number of studies and analyses in a bid to apply their observations of property crime patterns to crime forecasting.

The prediction of criminal patterns is important as this applied approach combines the analysis of crime patterns with an educated estimate of future occurrences of property crime. This, in turn, arms state and federal law enforcement services with the tools needed to create preventative solutions. By using these tools, agencies such as police and security services, can design and manage effective crime prevention measures and communicate these measures to the public [2].

Historically, studies show that data concerning real world (as opposed to intellectual property) crimes has been made up of complex, non-linear data structures with differing characteristics. This presents a significant challenge for researchers when attempting to identify the best model or technique with which to manage large amounts of complex data.

In this respect, artificial intelligence presents an exciting opportunity for researchers due to its ability to manage and analyse huge amounts of data at a rapid rate. In recent years, AI techniques have advanced rapidly, enabling researchers to complete studies of the application of AI in the use of non-linear functions for successful identification of non-linear patterns in big data; thereby increasing the accuracy of forecasting [2], [3].

Based on this significant advantage, this study will apply the forecasting performance of the aforementioned AI techniques (RFC, SVR and GTB) in order to compare and analyze the results in terms of accurately forecasting USA property crime rate data.

This study will be made up of the following sections:

Section2: An overview of applied AI techniques within the realm of crime forecasting.

Section3: A description of the experimental conduct of the study.

Section4: An analysis of the performance results of the different AI techniques applied within the conducted experiment.

Section5: A conclusion of the study

2. An Overview Of Applied AI Techniques Within The Realm Of Crime Forecasting

When using AI techniques to predict crime, researchers generally focus on machine learning applications for the forecasting and estimation of the target data (the crime rate) values. Artificial intelligence and machine learning are particularly popular with researchers and data analysts as it features superior adaptability and robustness in analyzing and forecasting data sets. Also, artificial intelligence techniques are able to learn and to model some incredibly complex data relationships, including non-linear types [4].

Within real world crime, data structures can appear in a number of different forms, representations and methods of distribution. Because of this, artificial intelligence techniques are an ideal solution as they transcend the data structure limitations which can occur in traditional forecasting and analysis of crime rate data [5], [8].

As mentioned previously, within this study, three separate artificial intelligence techniques were applied to predict occurrences of violent crime. These were; Random Forest Classifier (RFC), Support Vector Regression (SVR), and Gradient Tree Boosting (GTB). The study then compared the forecast accuracy performance of each of the separate techniques. These three techniques are described as follows:

2.1. Random Forecast Classifier (RFC)

This AI technique applies aggregation concept for both feature selection and issues relating to classification and regression [9]. This study introduced two factors; recursive partitioning (the tree method) and bootstrap aggregation (bagging). This method involves combining a number of created decision trees along with a number of bootstrap samples extracted from a learning sample which is randomly selected for each node of the predictor features' subset or section [10], [11].

2.2. Support Vector Regression (SVR)

This AI technique is based on the implementation of a symmetrical loss function in order to gain a real value function.[12]. This popular method has been inspired by the linear regression function computation within high dimensional feature space, in which inputted data is mapped by way of a non-linear function [13]. The benefit of this method is in its capabilities of generalizing unseen data, along with the ability to manage a number of different data representations in a case where appropriate kernel functions are supplied. [13]. Support Vector Regression is widely used by researchers and experts due to the fact that it combines outstanding generalization capabilities and superior prediction accuracy. [2], [14], [15].

2.3. Gradient Tree Boosting (GTB)

This widely used AI technique was introduced by [16] and focuses on a combination of decision trees and boosting techniques. Within this technique, the decision tree plays the role of a base. The boosting technique plays a vital role in the reduction of errors during the decision tree learning curve process and is conducted iteratively in that every boost iteration increases the accuracy of the tree from the previous iteration. [17]. The major benefit to using the Gradient Tree Boosting method is that is able to override the problem of overfitting in an instance whereby new independent data is introduced. [17]. In the arena of crime forecasting, this method is still considered to be relatively new and, therefore, available studies on its application within crime analysis are fairly minimal due to its fledgling status [18].

3. Experimental Conduct

Within the study, the three artificial intelligence models have been primed using statistics and machine learning. Each crime model (RFC, SVR and GTB) has been analyzed through the application of what we call multivariate analysis, which means that a number of factors which might reasonably influence a crime rate has been taken into consideration. Additionally, the model has been targeted to solve regression issues which can often occur within the prediction of crime values within each data set for the crime type.

Before we can forecast the crime rate value, we first need to train the model and, for this, we implement the previously established training data set used for data fitting. After training and fitting, the model can then be used to accurately predict crime rate values, using the aforementioned testing data set. The result of this forecast can then be used to calculate the performance of each of the three artificial intelligence crime models. The study will feature three separate methods of measuring and analyzing quantitative error to calculate the forecasting performance of our three AI models and, these are: Root Mean Square Error (RMSE), Mean Absolute Deviation (MAD), and, finally, Mean Absolute Percentage Error (MAPE).

3.1. Data Collection

In this research, crime rate and factors data set were used. Three types of US property crime rates in the crime rate data set, namely burglary, larceny, and motor vehicle theft, 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. Preparing and Processing the Data

Throughout this experiment, the data sets for the crime rate and factors were split into separate training sections which were: (in sample) and testing (out sample) data sets. The prepared training data was then used to effectively train each crime model. The testing data meanwhile was implemented for the testing and forecasting of crime rate values. For the purposes of the study, the data was divided to a precise ratio of 9:1 whereby the training data was made up of fifty data samples ranging from 1960 through to 2009 and, the testing data was made up of six samples ranging from 2010 to 2015 inclusive.

A possible issue with this was the possibility of unexpected errors caused by the different units of measurement. To counteract this, each of the data sets was correctly processed and then converted to a dimensionless format through a method of normalization which was done through a feature scaling methodology where the scale ranges between 0 and 1 and where the normalization can be defined by the equation (1). We can explain this as follows:

$$x' = (x_i - \min_x) / (\max_x - \min_x) \tag{1}$$

In the equation marked (1), xi denotes the raw value of the selected sample within our respective data series x (max) denotes the highest value of the raw data within the respective data series x (min) denotes the lowest value of the raw data within the respective data series x, and, finally, x ' denotes the new normalized value for the corresponding sample xi.

Once the forecasting process is complete, the dimensionless forms of the normalized forecast crime rate values for each of our crime models can then be reverted to the accurate raw crime rate values through the process of this denormalization. Our denormalization calculation has been performed on the basis of the mathematical transformation from equation (1) which is then described within the following equation (2).

$$x_i = (\max_x - \min_x) + \min_x \tag{2}$$

3.3. AI Technique Parameter Setups

A number of configurations relating to a number of input parameters were employed for each of the three AI techniques to create the crime model and this was in place prior to the creation of the crime model and, therefore, prior to the forecasting. For the RFC model, the set parameters were two hundred and fifty trees and, the chosen method was bagger. For the SVR method, Gaussian, 0.1 and sequential, minimal optimization parameters were chosen for the kernel function, the value of epsilon and, also, the optimization solver method as this was considered to be the most efficient and effective. Finally, for the GTB method, the parameters involved setting the learning rate to 100, 3 and 0.1 for the number of and size of individual trees.

4. The Results and Discussion Points

The study calculated the collated predicted property crime rate values of each artificial intelligence crime model and for each type of property crime using a form of quantitative error measurement analysis. The results of this can be observed in Table 1 below.

Table 1. Result of The AI model Quantitative Error Measurement per crime type.

Property Crime Type	AI Techniques	Quantitative Error Measurement

		RMSE	MAD	MAPE
Burglary	RFC	0.7621	0.7381	15.6378
	SVR	0.6690	0.6428	13.8879
	GTB	0.4517	0.4228	9.0082
Larceny Theft	RFC	6.5897	5.3570	19.7650
	SVR	4.5072	4.4144	16.2266
	GTB	4.3757	4.4043	16.3718
Motor vehicle Theft	RFC	36.4164	35.8245	32.6903
	SVR	27.1325	26.2597	23.1433
	GTB	24.5597	23.3599	21.8367

The results shown above in Table 1 demonstrate that the GTB scores highly in its ability to forecast performance when compared directly to both RFC and SVR within this experiment. We can further prove this by the fact that GTB also displays lower error measurement values within RMSE, MAD and MAPE when compared to the other two artificial intelligence techniques employed across all modelled crime types.

These results clearly show that, of all three models, GTB, while working with limited data samples within the study, is able to most efficiently forecast and predict rates for each of the crime types. This tells us that, despite a relatively small amount of data, GTB is robust and is able to provide an impressively accurate forecast for the different crime types.

In comparison, we see that the technique with the lowest level of performance was RFC which shows the highest levels for RMSE, MAPE and MAD values across all of the different types of crime. RFC significantly underperforms here while using the same limited crime data samples as GTB and SVR. Our overall result is that GTB is the superior technique in predicting violent crime rates in the USA, SVR appears in the middle and, RFC displays poor performance.

5. Conclusion

Within this study, I have examined the performance of three separate AI techniques - RFC, SVR and GTB to calculate their effectiveness at forecasting crime figures for the USA. I conducted this study as the use of artificial intelligence is now becoming more and more mainstream and can be applied to a number of sectors and industries, including law enforcement. By using three separate techniques on crimes in the USA which include burglary, larceny theft and motor vehicle theft, I was able to prove conclusively that the GTB technique produced considerably superior results and can therefore be considered the most appropriate tool for conducting this kind of forecasting, particularly when using limited crime rate data.

References

O. Oluduro, “Crime Forecasting and Planning in Developing Countries : Emerging Issues Crime Forecasting and Planning in Developing Countries : Emerging Issues,” ResearchGate, vol. 8, no. January 2012, pp. 2–7, 2017.

R. M. Nabi, S. A. M. Saeed, and H. Haron, “Artificial Intelligence Techniques and External Factors used in Crime Forecasting in Violence and Property : A Review,” 2020.

A. E. Yahya, R. Samsudin, and A. S. Ilman, A genetic algorithm-based grey model combined with fourier series for forecasting tourism arrivals in langkawi island malaysia, vol. 1073. 2020.

S. R. Bandekar and C. Vijayalakshmi, “Design and analysis of machine learning algorithms for the reduction of crime rates in India,” Procedia Comput. Sci., vol. 172, no. 2019, pp. 122–127, 2020.

A. R. Khairuddin, N. A. Ali, R. Alwee, H. Haron, and A. M. Zain, “Parameter optimization of gradient tree boosting using dragonfly algorithm in crime forecasting and analysis,” J. Comput. Sci., vol. 15, no. 8, pp. 1085–1096, 2019.

Y. S. Triana and A. Retnowardhani, “Enhance interval width of crime forecasting with ARIMA model-fuzzy alpha cut,” Telkomnika (Telecommunication Comput. Electron. Control., vol. 17, no. 3, pp. 1193–1201, 2019.

A. Ridzuan Khairuddin, R. Alwee, and H. Haron, “A review on applied statistical and artificial intelligence techniques in crime forecasting,” IOP Conf. Ser. Mater. Sci. Eng., vol. 551, no. 1, 2019.

S. Arora, H. Singh, M. Sharma, S. Sharma, and P. Anand, “A New Hybrid Algorithm Based on Grey Wolf Optimization and Crow Search Algorithm for Unconstrained Function Optimization and Feature Selection,” IEEE Access, vol. 7, no. c, pp. 26343–26361, 2019.

L. BREIMAN, “Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy,” Remote Sens., vol. 10, no. 6, pp. 5–32, 2001.

J. O. Ogutu, H. P. Piepho, and T. Schulz-Streeck, “A comparison of random forests, boosting and support vector machines for genomic selection,” BMC Proc., vol. 5, no. SUPPL. 3, 2011.

A. Hapfelmeier and K. Ulm, “A new variable selection approach using Random Forests,” Comput. Stat. Data Anal., vol. 60, no. 1, pp. 50–69, 2013.

R. Alwee, S. M. Hj Shamsuddin, and R. Sallehuddin, “Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators,” Sci. World J., vol. 2013, no. May, 2013.

J. Wu and Z. Lu, “A novel hybrid Genetic Algorithm and Simulated Annealing for feature selection and kernel optimization in support vector regression,” 2012 IEEE 5th Int. Conf. Adv. Comput. Intell. ICACI 2012, pp. 999–1003, 2012.

X. Ju, M. Cheng, Y. Xia, F. Quo, and Y. Tian, “Support vector regression and time series analysis for the forecasting of Bayannur’s total water requirement,” Procedia Comput. Sci., vol. 31, pp. 523–531, 2014.

C. H. Wu, G. H. Tzeng, and R. H. Lin, “A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression,” Expert Syst. Appl., vol. 36, no. 3 PART 1, pp. 4725–4735, 2009.

J. H. Friedman, “Greedy Function Approximation : A Gradient Boosting Machine Author (s): Jerome H . Friedman Source : The Annals of Statistics , Vol . 29 , No . 5 (Oct ., 2001), pp . 1189-1232 Published by : Institute of Mathematical Statistics Stable URL : https://ww,” vol. 29, no. 5, pp. 1189–1232, 2001.

H. Jiang, K. Huang, and R. Zhang, “Field support vector regression,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10634 LNCS, no. 10, pp. 699–708, 2007.

F. Ding, Q. Ge, D. Jiang, J. Fu, and M. Hao, “Understanding the dynamics of terrorism events with multiple-discipline datasets and machine learning approach,” PLoS One, vol. 12, no. 6, pp. 1–11, 2017.