HYBRID DRAGONFLY ALGORITHM WITH NEIGHBOURHOOD COMPONENT ANALYSIS AND GRADIENT TREE BOOSTING FOR CRIME RATES MODELLING

ALIF RIDZUAN KHAIRUDDIN

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

SEPTEMBER 2021

DEDICATION

This thesis is dedicated to my supervisors, families, friends and relatives for their dedication and support in my work.

ACKNOWLEDGEMENT

In this opportunity, I, Alif Ridzuan Khairuddin would like to express my appreciation to my main supervisor that is Dr. Razana Alwee and my co-supervisor that is Prof. Dr. Habibollah Haron in providing a very helpful guidance, encouragement and invaluable knowledge during my postgraduate study in Doctor of Philosophy.

Also, I would like to express my appreciation to all UTM Faculty of Engineering staffs, lecturers, postgraduate committee members, examiners and also to all that involved during my study. Your helps and guidance are invaluable to me.

Finally, I would like to thank my parents and relatives for all the encouragement and advice that helped me during my postgraduate study. Not to forget, I also would like to express my appreciation to all my friends in helping me during the study. Any guidance provided by them really helped me a lot.

Thank you.

ABSTRACT

In crime studies, crime rates time series prediction helps in strategic crime prevention formulation and decision making. Statistical models are commonly applied in predicting time series crime rates. However, the time series crime rates data are limited and mostly nonlinear. One limitation in the statistical models is that they are mainly linear and are only able to model linear relationships. Thus, this study proposed a time series crime prediction model that can handle nonlinear components as well as limited historical crime rates data. Recently, Artificial Intelligence (AI) models have been favoured as they are able to handle nonlinear and robust to small sample data components in crime rates. Hence, the proposed crime model implemented an artificial intelligence model namely Gradient Tree Boosting (GTB) in modelling the crime rates. The crime rates are modelled using the United States (US) annual crime rates of eight crime types with nine factors that influence the crime rates. Since GTB has no feature selection, this study proposed hybridisation of Neighbourhood Component Analysis (NCA) and GTB (NCA-GTB) in identifying significant factors that influence the crime rates. Also, it was found that both NCA and GTB are sensitive to input parameter. Thus, DA^2-NCA-eGTB model was proposed to improve the NCA-GTB model. The DA^2-NCA-eGTB model hybridised a metaheuristic optimisation algorithm namely Dragonfly Algorithm (DA) with NCA-GTB model to optimise NCA and GTB parameters. In addition, DA^2-NCA-eGTB model also improved the accuracy of the NCA-GTB model by using Least Absolute Deviation (LAD) as the GTB loss function. The experimental result showed that DA^2-NCA-eGTB model outperformed existing AI models in all eight modelled crime types. This was proven by the smaller values of Mean Absolute Percentage Error (MAPE), which was between 2.9195 and 18.7471. As a conclusion, the study showed that DA^2-NCA-eGTB model is statistically significant in representing all crime types and it is able to handle the nonlinear component in limited crime rate data well.

ABSTRAK

Dalam kajian jenayah, ramalan siri masa kadar jenayah membantu dalam membuat keputusan bagi pencegahan jenayah yang strategik. Model statistik biasanya digunakan dalam meramal siri masa kadar jenayah. Walau bagaimanapun, data siri masa kadar jenayah adalah terhad dan kebanyakannya tidak linear. Satu kelemahan dalam model statistik adalah model ini kebanyakannya hanya dapat memodelkan hubungan yang linear sahaja. Oleh itu, kajian ini mencadangkan model peramalan siri masa kadar jenayah yang dapat menangani masalah komponen tidak linear serta data kadar jenayah yang terhad. Baru-baru ini, model Kecerdasan Buatan (AI) semakin dikenali kerana ia dapat menangani komponen data sampel yang tidak linear dan fleksibel terhadap data kadar jenayah yang sedikit. Oleh itu, model jenayah yang dicadangkan menerapkan model kecerdasan buatan iaitu Penambahbaikan Pokok Kecerunan (GTB) dalam memodelkan kadar jenayah. Kadar jenayah dimodelkan menggunakan kadar jenayah tahunan Amerika Syarikat (AS) sebanyak lapan jenis jenayah dengan sembilan faktor yang mempengaruhi kadar jenayah. Oleh kerana GTB tiada pemilihan fitur, kajian ini mencadangkan hibridisasi Analisis Komponen Kejiranan (NCA) dan GTB (NCA-GTB) bagi mengenal pasti faktor-faktor penting yang mempengaruhi kadar jenayah. Juga didapati bahawa NCA dan GTB sensitif terhadap parameter input. Oleh itu, model DA^2-NCA-eGTB dicadangkan untuk memperbaiki model NCA-GTB. Model DA^2-NCA-eGTB menghibridisasi algoritma pengoptimuman metaheuristik iaitu Algoritma Pepatung (DA) dengan model NCA-GTB bagi mengoptimumkan parameter NCA dan GTB. Selain itu, model DA^2-NCA-eGTB juga meningkatkan ketepatan model NCA-GTB dengan menggunakan Sisihan Mutlak Paling Sedikit (LAD) sebagai fungsi kehilangan dalam GTB. Hasil eksperimen menunjukkan bahawa model DA^2-NCA-eGTB adalah lebih baik berbanding model AI yang sedia ada dalam semua jenis lapan jenayah yang dimodelkan. Ini dibuktikan oleh nilai Ralat Peratusan Mutlak Min (MAPE) yang lebih kecil iaitu antara 2.9195 dan 18.7471. Sebagai kesimpulan, kajian menunjukkan bahawa model DA^2-NCA-eGTB secara statistik adalah signifikan untuk mewakili semua jenis jenayah dan ia mampu menangani komponen tidak linear dalam data kadar jenayah yang terhad dengan baik.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	-	Artificial Neural Network
SVR	-	Support Vector Regression
RF	-	Random Forest
GTB	-	Gradient Tree Boosting
eGTB	-	Improved Gradient Tree Boosting
DA	-	Dragonfly Algorithm
NCA	-	Neighbourhood Component Analysis

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Crime is an act or action committed by an individual or a group of people intending to inflict damage to a targeted victim. Crime is mostly influenced by certain objectives or motives of the suspects towards their victims. In the real world, crime is a part of society which cannot be predicted by the police (Ghazvini et al., 2015). The crime rate itself represents the degree of public safety of a country. The analysis of crime rate data helps in understanding the behaviour of the crime trend and future values may be forecast from past observations (Shrivastav and Ekata, 2012). Hence, crime forecasting is an essential analysis which affects the relative profits of people's life and properties (Yao-Lin et al., 2015).

In literature, several types of crime forecasting models have been introduced such as statistical models (Gorr et al., 2003; Omar et al., 2007; Shoesmith, 2012; Huddleston and Brown, 2013; Cesario et al., 2016) and artificial intelligence models (Olligschlaeger, 1997; Kianmehr and Alhajj, 2006; Yu et al., 2011; Vineeth et al., 2016; Yang et al., 2018). The crime models introduced by various researchers analyse past or present crime data trends to estimate future crime occurrence. Examples of the statistical models are linear regression, exponential smoothing, moving average (MA), and autoregressive integrated moving average (ARIMA). Among the examples of artificial intelligence models are artificial neural network (ANN), support vector regression (SVR), gradient tree boosting (GTB), and random forest (RF).

There are several factors that influence crime rate such as social instability, demographic, and economic disadvantages (Mittal et al., 2020). Previous studies provide evidence that crime occurrence is influenced by various factors (Nolan and

James, 2004; Rosenfeld and Fornango, 2007; Habibullah and Bhahrom, 2009; Goulas and Zervoyianni, 2013; Stansfield et al., 2017; Northrup and Klaer, 2014; Rosenfeld et al., 2019). By including influence factors in forecasting crime rates, new crime patterns that never occurred in the past could be discovered. Hence, the crime model accuracy can be improved.

In crime forecasting, the time series data have been used by various researchers in building the crime models (Greenberg, 2001; Saridakis, 2006; Huang et al., 2015; Mahmud et al., 2016). The time series data of crime rate is mostly limited, has complex relationships and exist in a nonlinear representation with small portions of linear patterns. Such characteristics pose difficulties in modelling an accurate crime rate model. One limitation of statistical models is that they are only able to capture linear patterns. Hence, it is difficult to model time series data using linear statistical methods (Du et al., 2020). In contrast, artificial intelligence models are robust in presenting various representations of time series data (Bontempi et al., 2013). This makes artificial intelligence models more suitable for modelling crime rates.

Therefore, this study proposes a suitable model that is able to handle crime rate data, identify significant factors that influence crime rates, and accurately forecast crime rates using the available data sets. The aim of this study is to propose an accurate crime rate forecasting model that is able to forecast the annual crime rates. The proposed crime rate forecasting model was developed to model the crime rate based on the sample data sets from the United States (US) annual crime rates from year the 1960 to 2015 (56 data samples). There were eight types of crime rate data to forecasts namely murder and non-negligent manslaughter, forcible rape, aggravated assaults, robbery, burglary, larceny-theft, motor vehicle theft and total crime rates for all types of crimes.

The social and economic stability of a country was often influenced by the trends of the annual crime rate. Hence, this study used annual data for forecasting the crime rates. It helps in increasing crime awareness among the public community. In addition, the changes in annual crime rate trends usually serve as an indicator for the government to incorporate macroeconomic models in formulating efficient economic strategies. Figure 1.1 illustrates the problem formulation in criminology related to this study.

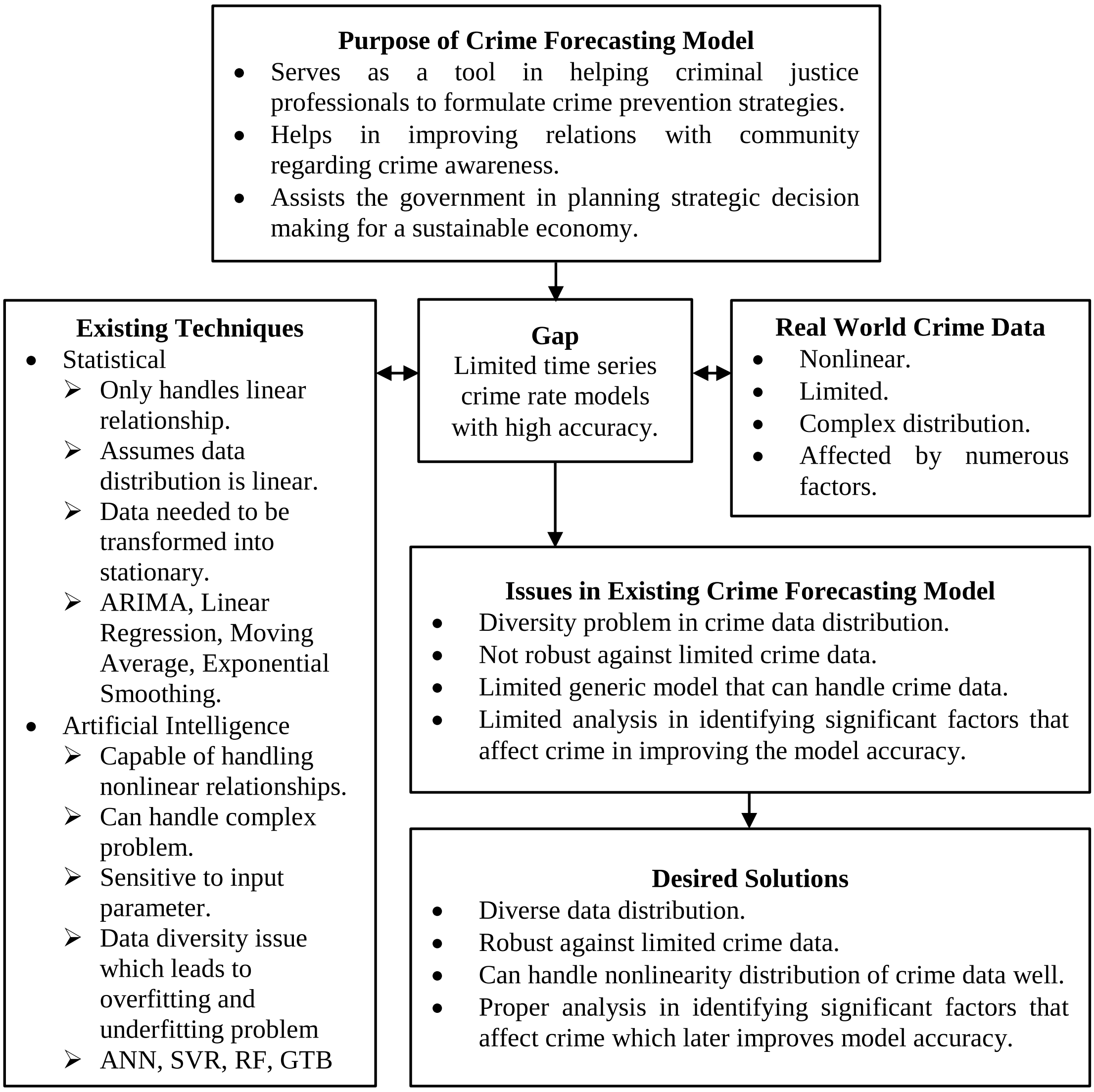

Figure 1.1 Problem Formulation in Criminology

1.2 Problem Background

In criminology, the application of time series models in forecasting crime is limited and rarely applied in most countries (Suzilah and Nurulhuda, 2013; Alwee, 2014). This is because in the real world, crime data are limited and difficult to obtain

(Zhao and Tang, 2018; Wang et al., 2019b). In addition, there is no universal crime model that is able to handle all types of crime data representations. In literature, various crime forecasting models have been introduced in handling time series data such as statistical models (Chen et al., 2008; Huddleston et al., 2015; Cesario et al., 2016) and artificial intelligence models (Huang et al., 2015; Vineeth et al., 2016; Wang et al., 2019b).

The application of a statistical model is conducted with an assumption that the obtained time series data are stationary and linear (Alwee, 2014). Such a limitation causes the linear model to be unable to capture the nonlinearity of the data (Rather et al., 2017). In crime rate data, the structure is complex and exists in a nonlinear pattern. Hence, it is difficult to model the crime rate data using a statistical model. In recent years, artificial intelligence has been favoured by most researchers in forecasting crimes due to its high generalisation capabilities (Vaquero, 2016). The reason is that an artificial intelligence model contains several nonlinear functions that are able to identify nonlinear patterns in the data and also possesses high generalisation capabilities that a statistical model lacks. Therefore, an artificial intelligence model that is able to model limited crime rate data with nonlinear structures is needed.

Artificial neural network (ANN) and support vector regression (SVR) are among the popularly applied artificial intelligence models in crime forecasting. Although ANN and SVR models are favoured by researchers, there are several drawbacks. ANN suffers from parameter control, possibility of overfitting, and network weight uncertainty (Alwee, 2014). As for SVR, it is sensitive to parameter, lacks transparency in result accuracy, and is computationally demanding (Awad and Khanna, 2015). In this study, an artificial intelligence technique called gradient tree boosting (GTB) is used to model the crime rate. GTB has been applied in various research domains by different researchers (Kim et al., 2015; Mayrink and Hippert, 2016; Persson et al., 2017; Cai et al., 2019; Tan et al., 2020). However, the application of GTB in time series crime forecasting is limited with minimal improvements made (Kumar and Bhalaji, 2016; Nguyen et al., 2017). GTB adopts numerical optimisation methods to minimise the loss function of the predictive model which later improves GTB's overall capabilities. GTB is diverse to data structure, produces output with low variance (error), and able to generate interpretable solutions for regression problems (Nguyen et al., 2017; Ke et al., 2015; Chandrasekar et al., 2015).

It is known that the crime rate is influenced by various factors (Hanslmaier et al., 2015). Studies on the influence of several factors such as economic (Habibullah and Baharom, 2009; Alwee, 2014), social (Hanslmaier et al., 2015; Hipp et al., 2011) and demographic (Ranson, 2013; Brown and Males, 2011) towards crime have been conducted by previous researchers. The study will analyse the significant impact of various factors towards crime occurrence. This is to ensure that the irrelevant factors that negatively affect crime model accuracy can be eliminated. When considering various factors in modelling crime rate, a multivariate analysis is required. Multivariate analysis uses more than one time series data in model development. The analysis is done to find the cross-correlation between multiple time series data (Preez and Witt, 2003). It is very useful when discovering a new pattern of data that never occurred in the past (Alwee, 2014).

In identifying relevant factors, feature selection is the popular approach by researchers recently. There are various feature selection approaches proposed by different research to identify and select significant factors that influence crime such as metaheuristic algorithm (Anuar et al., 2014; Liu et al., 2019) and statistical approach (Shalabi, 2017; Ingilevich and Ivanov, 2018). In this study, neighbourhood component analysis (NCA) is used as the feature selection method in identifying and selecting relevant factors that significantly affect crime rate. Previous researchers have introduced the application of NCA as the feature selection method in various research domains (Yang et al., 2012b; Wu et al., 2018; Jin and Deng, 2018; Tuncer and Ertam, 2020). This is because the capability of NCA in identifying the significant features is better than the other feature selection methods such as Principal Component Analysis (PCA), Sequential Feature Selection (SFS) and ReliefF in improving the model accuracy (Jin and Deng, 2018; Tuncer and Ertam, 2020).

From the studies conducted, both GTB and NCA share one drawback. The drawback is both GTB and NCA's accuracy is sensitive to input parameters. Optimising the parameters in GTB is challenging because an inappropriate parameter configuration leads to overfitting or underfitting problems. Thus, rather than GTB attempting to predict the functional dependence between input and response variables, instead GTB will predict the training data itself (Natekin and Knoll, 2013). There are three parameters that impact GTB' accuracy; number of trees, size of individual trees, and learning rate (Saha et al., 2015; Jalabert et al., 2010; Guelman, 2012; Elith, 2008; Zhang and Haghani, 2015). As for NCA, the performance is controlled by one parameter which is regularisation parameter λ. This parameter alleviates the overfitting problem in feature selection when applying NCA and is able to improve the selection of relevant factors. An optimal regularisation parameter value is able to minimise the generalisation error in NCA (Yang et al., 2012b).

Previous researchers have proposed various solutions to assess such drawbacks in optimising the parameters of GTB (Qi et al., 2018; Zhang et al., 2019; Yu et al., 2020) and NCA (Raghu and Sriraam, 2018; Malan and Sharma, 2019). In most work, researchers implemented a metaheuristic optimisation algorithm as a solution to optimise the input parameter values in various applications (Alwee, 2014; Ebrahimi et al., 2016; Hou et al., 2018). Examples of metaheuristic optimisation algorithms are genetic algorithm (Vlahogianni et al., 2005; Oliveira et al., 2010) and particle swarm optimisation (Ren et al., 2014; Chatterjee et al., 2016). The metaheuristic optimisation algorithm is a popular solution as it is able to produce robust output and converges to global optimum. In this study, an implementation of the dragonfly algorithm (DA) in optimising the input parameters in GTB and NCA is considered. DA is capable to improve the random population for a given problem, converges towards global optimum, and produces robust results (Mirjalili, 2016).

Another issue found is that the applied loss function in GTB plays a critical role that consecutively fits the new model in order to provide a more accurate forecast (Freeman et al., 2015). In GTB, the least square function is used as a loss function to consecutively minimise its 'pseudoresponses' value (error-fitting) over the response variable. It is known that the distribution of crime data varies and is not constant. Thus, the appropriate application of the loss function is beneficial as it is able to provide a flexibility in model design that fits different application needs (Guelman, 2012). Such an approach provides a robustness to GTB that fits the crime rate data. In this study, a mathematical function called least absolute deviation (LAD) was considered in replacing the GTB least square loss function. LAD is advantageous as it provides a robust regressive fitting with multiple solutions that the least square function does not possess (Natekin and Knoll, 2013; Kržić and Seršić, 2018). Based on the identified problems, Figure 1.2 defines the issues and improvement measures taken in developing the proposed crime forecasting model.

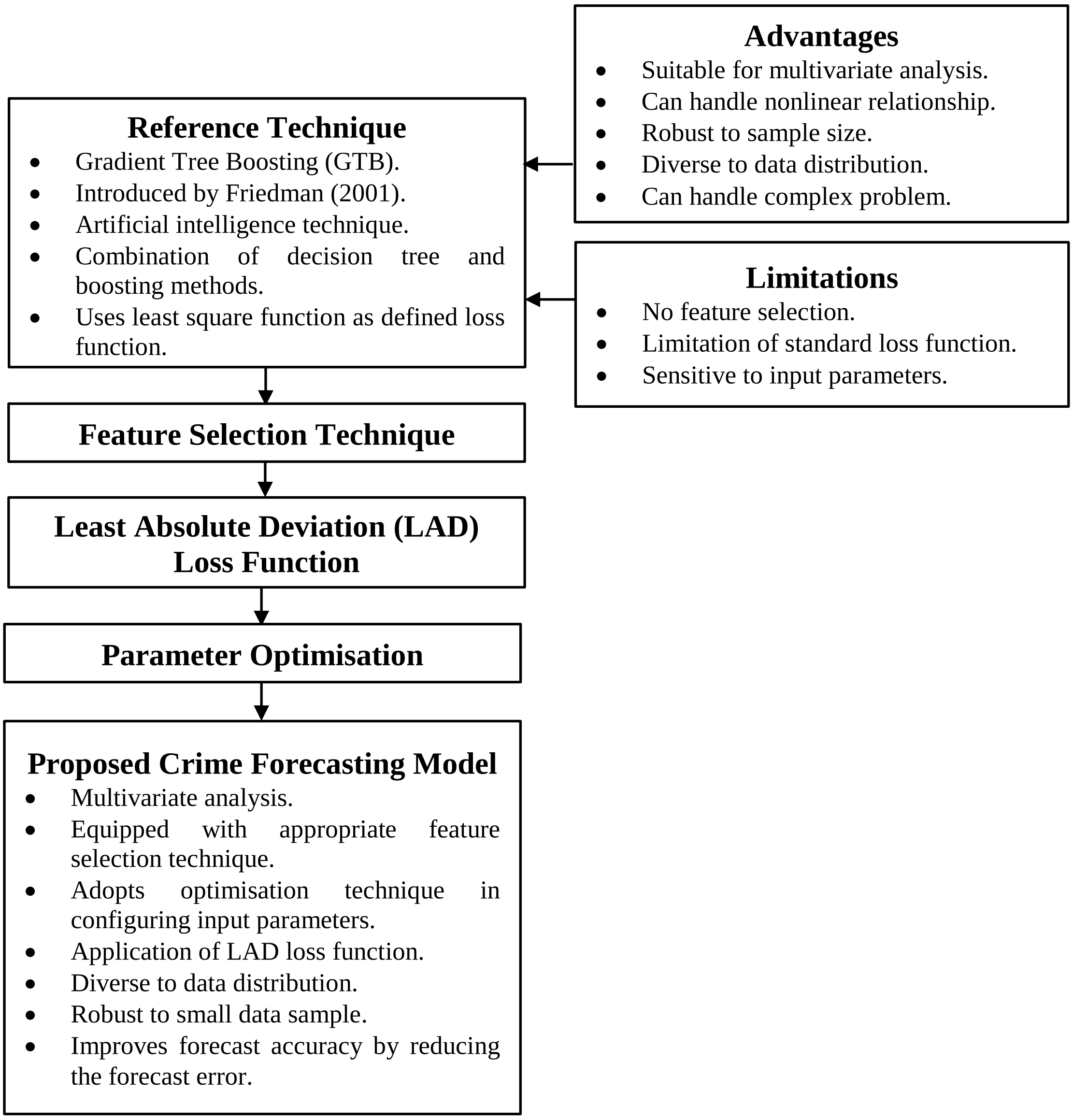

Figure 1.2 Defined Issues and Improvement Measures Taken in Developing the Proposed Crime Forecasting Model

1.3 Problem Statement

In most existing real world crime rate data, the distribution pattern is nonlinear. The crime rate pattern is mostly influenced by several factors such as economic and social conditions. Thus, a robust time series forecasting model is needed to handle such complex behaviour. Recently, the AI technique is favoured by researchers as it is robust to a data structure provided with proper configuration. Among the introduced AI techniques, GTB shows a promising result. Like other AI techniques, GTB is also sensitive to input parameters. A proper parameter configuration is required to ensure that GTB is able to produce a better and reliable forecast result. The DA algorithm is selected in assessing this problem.

Another issue found in GTB is that it uses the least square function as the standard loss function. This is because the least square function is only able to approximate one solution and never reaches global minimum (Kržić and Seršić, 2018). Hence, a study on the application of a suitable mathematical function to replace the GTB least square loss function is recommended. This is to ensure that the developed GTB crime model with a suitable loss function can fit this study's crime data. In determining an appropriate mathematical function as the loss function in GTB, the process is often influenced by the characteristics of data distribution (Natekin and Knoll, 2013). Thus, a mathematical function called least absolute deviation (LAD) was selected as a potential solution to replace the least square function in GTB for this study.

As mentioned before, the crime rate is influenced by several factors such as economic disadvantages and social mistreatment. By considering these factors, it helps in discovering new patterns in the crime rate and later increases forecast accuracy. However, not all factors influence the crime rate as some of them might negatively affect forecast accuracy. Hence, a proper analysis to observe the relationship between factors and crime data is needed to select the significant factors that influence the crime rate. In this study, NCA feature selection is equipped into GTB to identify and select the significant factors. In NCA, the regularisation parameter determines the overall NCA complexity. This issue potentially causes

NCA to overfit if the regularisation parameter value is too high. Thus, the DA algorithm is applied to tackle this issue. Based on the previous discussion, the study research hypothesis is defined as follows:-

The accuracy of crime rates model could be improved by implementing multivariate crime forecasting model using gradient tree boosting (GTB), neighbourhood component analysis (NCA) as feature selection to identify the significant factors that influence the crime rate, applying dragonfly algorithm (DA) for parameters optimisation and further improved the model accuracy by implementing least absolute deviation (LAD) loss function in GTB.

The following research question for this study is defined as follows:-

(a) How to design a new multivariate crime forecasting model that is able to accurately forecast the crime rate with limited time series data?
(b) How to select factors that significantly influence the crime rate in order to improve the model accuracy?
(c) What is the suitable standard mathematical function to replace the gradient tree boosting’s least square loss function for better accuracy?
(d) How to apply the metaheuristic optimisation algorithm in parameter estimation for better accuracy?

1.4 Research Goal and Objective

The research goal is to propose a new multivariate crime forecasting model with feature selection method by integrating neighbourhood component analysis (NCA) into gradient tree boosting (GTB) and further improving the performance of NCA and GTB with parameter optimisation through the hybridisation of dragonfly algorithm (DA) and implementation of least absolute deviation (LAD) as the GTB loss function for better accuracy. The research objectives are defined as follows:-

(a) To develop a multivariate time series model for modelling the crime rate using gradient tree boosting (GTB).

(b) To integrate neighbourhood component analysis into gradient tree boosting (NCA-GTB) as a feature selection model to identify the significant factors in modelling the crime rate.

(c) To propose DA^2-NCA-eGTB model through hybridisation of dragonfly algorithm with NCA-GTB for parameters optimisation and application of least absolute deviation loss function in improving the accuracy of the proposed model.

1.5 Research Scope

In this study's research scope, multivariate crime analysis is the focused domain. A hybridisation approach becomes the main focus in this study since the proposed model is a combination of various techniques in producing one complete hybrid crime model. First, an AI technique namely GTB is chosen as the base model in modelling the crime rate. Next, NCA is integrated into GTB to identify and select the significant factors that affect crime. After that, DA is hybridised with NCA and GTB. The hybridisation purpose is to optimise the parameter values of both NCA and GTB. Lastly, based on the studied loss function, three mathematical functions i.e. least absolute deviation (LAD), Huber, and quantile are selected.

For data definition, the data set used in this study is divided into two types; crime data and factors data. Both types of data are annual data collected from 1960 to 2015 which is equivalent to 56 data samples for each year. A detailed explanation about data definitions is discussed in Chapter 3. In this study, sample data sets from the United States' (US) annual crime rates from 1960 to 2015 are collected. The collected data sets were obtained from the Uniform Crime Reporting Statistics website (https://www.ucrdatatool.gov) provided by the Federal Bureau of Investigation (FBI) of the United States. There are eight types of crime rates: murder and non-negligent manslaughter, forcible rape, aggravated assault, robbery, burglary, larceny theft, motor vehicle theft, and total crime rate for all types of crime.

There are nine factors data selected and obtained in this study. These are unemployment rate (UR), immigration (IR), population rate (PR), consumer price index (CPI), gross domestic product (GDP), consumer sentiment index (CSI), poverty rate (PoR), inflation rate (InR), and tax revenue (TR). Data for the selected factors were obtained from the US Bureau of Labour Statistics (UR and CPI), US Bureau of Economic Analysis (GDP), US Census Bureau (PR), University of Michigan consumers' survey (CSI), US Department of Homeland Security (IR), US Inflation Calculator (InR) website, World Bank website (PN), and US Internal Revenue Service (TR).

For evaluation and validation analysis, three types of quantitative error measurement analyses are applied to evaluate and compare the performance of the proposed crime model with others. The quantitative error measurement analyses used are root mean square error (RMSE), mean absolute deviation (MAD), and mean absolute percentage error (MAPE). In addition, a statistical test analysis (paired sample t-test) is also applied to validate the proposed crime forecasting model.

In terms of software and tools, the experiment is primarily conducted on the Python and Matlab platforms. In Python, Scikit-learn tools are used in modelling GTB. Scikit-learn was developed by Pedregosa et al. (2011) and is a Python module package that implements varieties of state-of-the-art machine learning algorithms for various problem-solving solutions. It offers good flexibility in configuring the parameters and produces a consistent result. Matlab is used in implementing the NCA for feature selection and DA module for parameter optimisation purposes. In addition, Matlab is also used for calculating the quantitative error measurement result produced from the developed crime model. Other than that, the statistical analysis (paired sample t-test) is conducted on the SPSS platform to validate the developed crime model. Also, OriginPro software is used for the result's data visualisation and representation such as graph and scatter diagram. In addition, Microsoft Office software is used for documentation purpose.

1.6 Research Significance

The crime rate discusses the nature of emerging and continuing crime problems in different areas of the jurisdiction. The crime rate is often linked with the social and economic stability of a country. Governments mostly incorporate macroeconomic models in formulating efficient economic policies or strategies. The change in crime rates is used as an indicator for the macroeconomic development. The purpose of the crime rate is for strategic decision making in formulating crime prevention strategies. The crime rate data also help in improving relations in a community regarding crime awareness. Thus, a crime model to accurately forecast the crime rate is very beneficial and needed.

In existing crime rate models, several problems arise such as non-robustness to small data samples and diversity issues in crime data distribution. The application of AI techniques serves as a viable solution in handling such problems. This is because AI techniques are able to perform well even when the data sample is small and also diverse to complex distribution. As crime rates are mostly influenced by several factors, the feature selection method proposed in this study is able to identify and select the significant factors. Hence, the proposed model includes the impact of various factors in the crime rates. In addition, by incorporating the metaheuristic optimisation algorithm into both AI technique and feature selection method, the crime model accuracy can be further increased. The assessment of problems that arise in modelling the time series crime model makes this study significant to the field of criminology and multivariate time series forecasting.

1.7 Research Methodology

This study is divided into seven main phases. They are literature review and problem definition, data definition and preparation, GTB crime model development, NCA-GTB factor selection, development of hybrid DA-NCA-GTB model, parameter optimisation with LAD loss function and lastly, evaluation and validation analysis. In the first phase, a thorough investigation and study in crime forecasting is conducted to observe recent work, identify issues or problems that arise and formulate potential solutions to the problems. In phase two, the required data set is defined, collected, and prepared. For the third phase, a base crime model using GTB is modelled using the prepared data set.

Next, in phase four, the GTB crime model is equipped with NCA feature selection in analysing and identifying significant factors that influence crime rates. In phase five, the development of a hybrid DA-NCA-GTB model was conducted. After that, in phase six, the proposed DA^2-NCA-eGTB crime model is modelled by optimising the input parameters for NCA and GTB using DA, and implementing the LAD loss function in GTB. Lastly, in the final phase, the proposed model output is evaluated based on quantitative measurement error analysis. Also, the statistical test analysis is performed to validate the model. Figure 1.3 shows an overview of the research methodology in this study.

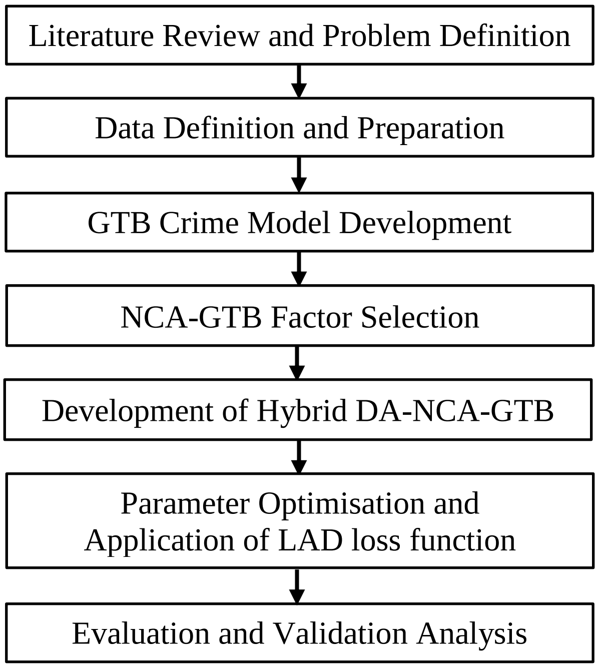

Figure 1.3 Overview of Research Methodology

1.8 Research Contribution

This study's main contribution is the new nonlinear crime forecasting model to accurately forecast crime rate data. The proposed model is based on a multivariate time series analysis that is designed to handle limited crime rate data. The model is equipped with the feature selection method to identify and select the significant factors that influence the crime rate. The selection of factors is based on types of crime rates. The purpose of feature selection is to analyse and identify the relationship between factors and the crime rate. By identifying the significant factors, the crime model's forecast accuracy can be improved for each type of crime rate.

The proposed hybrid crime model is modelled based on GTB. NCA is equipped into GTB for feature selection. The proposed model is able to accurately model the limited crime rate data with nonlinear structure. Further, the proposed model is improved by optimising both the NCA and GTB input parameters. In addition, GTB is further improved by implementing LAD as a loss function that is more suitable for the crime rate data. The improvement made is to overcome the limited crime rate data constraint.

1.9 Thesis Organisation

Chapter 1 is the introduction that briefly summarises and provides an overview of the study. Chapter 2 provides discussions of literature reviews concerning recent research findings, issues, and solutions. Chapter 3 presents the research methodology that explains the framework and procedures in conducting the research. Chapter 4 discusses the development of GTB as the base model in modelling crime rate. Chapter 5 proposes an NCA-GTB feature selection model in identifying significant factors that influence crime which later improves the overall model prediction accuracy. Chapter 6 proposes an improvement to the proposed NCA-GTB crime model by hybridising DA to optimise the NCA and GTB parameters, and replacing the GTB least square loss function with LAD function. Lastly, Chapter 7 provides the conclusion of the study.

REFERENCES

Aadil, F., Ahsan, W., Rehman, Z. U., Shah, P. A., Rho, S. and Mehmood, I. (2018). Clustering algorithm for internet of vehicles (IoV) based on dragonfly optimizer (CAVDO). *The Journal of Supercomputing*, 1-26.

Åberg, M. B., Löken, L. and Wessberg, J. (2008). An Evolutionary Approach to Multivariate Feature Selection for FMRI Pattern Analysis. *BIOSIGNALS*, 2, 302-307.

Abhishek, K., Singh, M., Ghosh, S. and Anand, A. (2012). Weather forecasting model using artificial neural network. *Procedia Technology,* 4, 311-318.

Aghababaei, S. and Makrehchi, M. Mining Social Media Content for Crime Prediction. (2016). IEEE/WIC/ACM International Conference on Web Intelligence (WI), 526-531.

Ahn, H. and Kim, K.-J. (2009). Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. *Applied Soft Computing,* 9, 599-607.

Al Boni, M. and Gerber, M. S. (2016). Area-Specific Crime Prediction Models. 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 671-676.

Al Shalabi, L. (2017). Perceptions of crime behavior and relationships: rough set based approach. *International Journal of Computer Science and Information Security (IJCSIS)*, 15.

Alam, M. A., Zehra, B. and Agrawal, N. A. (2014). Comprehensive Study of Artificial Neural Networks. National Conference on Recent Advances in Electronics and Communication Engineering, 28-29 March.

Aldehim, G. and Wang, W. (2017). Determining appropriate approaches for using data in feature selection. *International Journal of Machine Learning and Cybernetics,* 8, 915-928.

Ali, J., Khan, R., Ahmad, N. and Maqsood, I. (2012). Random forests and decision trees. *International Journal of Computer Science Issues (IJCSI),* 9, 272.

Almazroey, A. A. and Jarraya, S. K. (2020). Abnormal Events and Behavior Detection in Crowd Scenes Based on Deep Learning and Neighborhood Component Analysis Feature Selection. Joint European-US Workshop on Applications of Invariance in Computer Vision. *Springer*, 258-267.

Alsharif, M. H., Younes, M. K. and Kim, J. (2019). Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea. *Symmetry*, 11, 240.

Altindag, D. T. (2014). Crime and International Tourism. *Journal of Labor Research*, 35, 1-14.

Alves, L. G., Ribeiro, H. V. and Rodrigues, F. A. (2018). Crime prediction through urban metrics and statistical learning. *Physica A: Statistical Mechanics and its Applications*, 505, 435-443.

Alwee, R. (2014). *Swarm Optimized Support Vector Regression with Autoregressive Integrated Moving Average for Modeling of Crime Rate*. PhD Thesis, Universiti Teknologi Malaysia.

Alwee, R., Shamsuddin, H., Mariyam, S. and Sallehuddin, R. (2013). Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators. *The Scientific World Journal*, 2013.

Amiri, F., Yousefi, M. R., Lucas, C., Shakery, A. and Yazdani, N. (2011). Mutual information-based feature selection for intrusion detection systems. *Journal of Network and Computer Applications*, 34, 1184-1199.

Amroune, M., Bouktir, T. and Musirin, I. (2018). Power System Voltage Stability Assessment Using a Hybrid Approach Combining Dragonfly Optimization Algorithm and Support Vector Regression. *Arabian Journal for Science and Engineering*, 1-14.

Anuar, M. S., Selamat, A. and Sallehuddin, R. (2014). *Particle Swarm Optimization Feature Selection for Violent Crime Classification.* Advanced Approaches to Intelligent Information and Database Systems. Springer International Publishing.

Awad, M. and Khanna, R. (2015). *Efficient learning machines: theories, concepts, and applications for engineers and system designers*. Apress.

Babakura, A., Sulaiman, M. N. and Yusuf, M. A. (2014). Improved method of classification algorithms for crime prediction. IEEE International Symposium on Biometrics and Security Technologies (ISBAST), 250-255.

Baglivio, M. T., Wolff, K. T., Epps, N. and Nelson, R. (2017). Predicting adverse childhood experiences: The importance of neighborhood context in youth trauma among delinquent youth. *Crime and Delinquency,* 63, 166-188.

Bai, Y., Jin, X., Wang, X., Su, T., Kong, J. and Lu, Y. (2019). Compound autoregressive network for prediction of multivariate time series. *Complexity*, 2019.

Baliyan, A., Gaurav, K. and Mishra, S. K. (2015). A Review of Short Term Load Forecasting using Artificial Neural Network Models. *Procedia Computer Science,* 48, 121-125.

Bandara, K., Bergmeir, C. and Smyl, S. (2020). Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert Systems with Applications*, 140, 112896.

Basak, D., Pal, S. and Patranabis, D. C. (2007). Support vector regression. *Neural Information Processing-Letters and Reviews,* 11, 203-224.

Batabyal, S. (2011). Temporal causality and the dynamics of crime and delinquency. *Atlantic Economic Journal,* 39, 421-441.

Bergstra, J., Pinto, N. and Cox, D. (2012). Machine learning for predictive auto-tuning with boosted regression trees. IEEE Innovative Parallel Computing (InPar), 1-9.

Berk, R. (2011). Asymmetric loss functions for forecasting in criminal justice settings. *Journal of Quantitative Criminology,* 27, 107-123.

Berk, R. A., Sorenson, S. B. and Barnes, G. (2016). Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies*, 13, 94-115.

Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F. and Pentland, A. (2014). Once upon a crime: towards crime prediction from demographics and mobile data. Proceedings of 16th international conference on multimodal interaction, 427-434.

Bolón-Canedo, V., Sánchez-Maroño, N. and Alonso-Betanzos, A. (2013). A review of feature selection methods on synthetic data. *Knowledge and information systems*, 34, 483-519.

Bontempi, G., Ben Taieb, S. and Le Borgne, Y.-A. (2013). *Machine Learning Strategies for Time Series Forecasting*. Tutorial Lectures of Business Intelligence: Second European Summer School, Brussels. Springer Berlin Heidelberg.

Botsch, M. and Nossek, J. A. (2007). Feature selection for change detection in multivariate time-series. IEEE Symposium on Computational Intelligence and Data Mining, 590-597.

Breiman, L. (1997). Arcing the edge. Technical Report 486, Statistics Department, University of California, Berkeley.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32.

Brown, E. and Males, M. (2011). Does age or poverty level best predict criminal arrest and homicide rates? A preliminary investigation. *Justice Policy Journal*, 8, 1-30.

Budur, E., Lee, S. and Kong, V. S. (2015). Structural analysis of criminal network and predicting hidden links using machine learning. Social and Information Networks Report, Stanford University.

Buonanno, P. and Montolio, D. (2008). Identifying the socio-economic and demographic determinants of crime across Spanish provinces. *International Review of law and Economics*, 28, 89-97.

Bye, E. K. (2007). Alcohol and violence: use of possible confounders in a time-series analysis. *Addiction*, 102, 369-376.

Cai, L., Gu, J., Ma, J. and Jin, Z. (2019). Probabilistic wind power forecasting approach via instance-based transfer learning embedded gradient boosting decision trees. *Energies*, 12, 159.

Camacho-Collados, M. and Liberatore, F. (2015). A decision support system for predictive police patrolling. *Decision Support Systems*, 75, 25-37.

Cao, D.-S., Xu, Q.-S., Liang, Y.-Z., Zhang, L.-X. and Li, H.-D. (2010). The boosting: A new idea of building models. *Chemometrics and Intelligent Laboratory Systems*, 100, 1-11.

Cao, L. and Tay, F. E. (2001). Financial forecasting using support vector machines. *Neural Computing and Applications*, 10, 184-192.

Caplan, J. M., Kennedy, L. W. and Piza, E. L. (2013). Joint utility of event-dependent and environmental crime analysis techniques for violent crime forecasting. *Crime and Delinquency*, 59, 243-270.

Castelli, M., Sormani, R., Trujillo, L. and Popovič, A. (2017). Predicting per capita violent crimes in urban areas: an artificial intelligence approach. *Journal of Ambient Intelligence and Humanized Computing,* 8, 29-36.

Cesario, E., Catlett, C. and Talia, D. (2016). Forecasting Crimes Using Autoregressive Models. 14th IEEE Intl Conf on Dependable, Autonomic and Secure Computing (DASC), 14th IEEE Intl Conf on Pervasive Intelligence and Computing (PiCom), 2nd IEEE Intl Conf on Big Data Intelligence (DataCom) and 2nd IEEE Intl Conf on Computing and Cyber Science and Technology Congress (CyberSciTech), 795-802.

Chakraborty, B. (2007). Feature selection and classification techniques for multivariate time series. IEEE Second International Conference on Innovative Computing, Information and Control, 42-42.

Chandra, A., Khatri, S. K. and Simon, R. (2019). Filter-based Attribute Selection Approach for Intrusion Detection using k-Means Clustering and Sequential Minimal Optimization Technique. IEEE Amity International Conference on Artificial Intelligence (AICAI), 740-745.

Chandra, B., Gupta, M. and Gupta, M. A. (2008). Multivariate time series clustering approach for crime trends prediction. IEEE International Conference on Systems, Man and Cybernetics, (SMC), 892-896.

Chandrasekar, A., Raj, A. S. and Kumar, P. (2015). Crime Prediction and Classification in San Francisco City. *CS229 Technical Report : Machine Learning.* Stanford Computer Science Department: Stanford University.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, 40, 16-28.

Chang, D. (2011). Social crime or spatial crime? Exploring the effects of social, economical, and spatial factors on burglary rates. *Environment and Behavior*, 43, 26-52.

Chapelle, O., Shivaswamy, P., Vadrevu, S., Weinberger, K., Zhang, Y. and Tseng, B. (2010). Multi-task learning for boosting with application to web search ranking. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 1189-1198.

Chatterjee, S., Sarkar, S., Hore, S., Dey, N., Ashour, A. S. and Balas, V. E. (2017). Particle swarm optimization trained neural network for structural failure prediction of multistoried RC buildings. *Neural Computing and Applications,* 28, 2005-2016.

Chen, P., Yuan, H. and Shu, X. (2008). Forecasting crime using the arima model. Fifth IEEE International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 627-630.

Chen, S.-M. and Chen, C.-D. (2011). TAIEX forecasting based on fuzzy time series and fuzzy variation groups. *IEEE Transactions on Fuzzy Systems*, 19, 1-12.

Chen, S.-M. and Tanuwijaya, K. (2011). Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. *Expert Systems with Applications,* 38, 10594-10605.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM international conference on knowledge discovery and data mining (SIGKDD), 785-794.

Chen, T. and He, T. Higgs boson discovery with boosted trees. (2015). NIPS Workshop on High-energy Physics and Machine Learning, 69-80.

Chen, Y., Jia, Z., Mercola, D. and Xie, X. (2013). A gradient boosting algorithm for survival analysis via direct optimization of concordance index. *Computational and mathematical methods in medicine,* 2013.

Chen, Y., Yang, B., Dong, J. and Abraham, A. (2005). Time-series forecasting using flexible neural tree model. *Information sciences,* 174, 219-235.

Chen, Y. H., Hong, W.-C., Shen, W. and Huang, N. N. (2016). Electric load forecasting based on a least squares support vector machine with fuzzy time series and global harmony search algorithm. *Energies*, 9, 70.

Chen, Y. M. and Lin, C.-T. (2007). Dynamic parameter optimization of evolutionary computation for on-line prediction of time series with changing dynamics. *Applied Soft Computing,* 7, 1170-1176.

Cheng, C.-H., Cheng, G.-W. and Wang, J.-W. (2008). Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Systems with Applications,* 34, 1235-1242.

Chniti, G., Bakir, H. and Zaher, H. (2017). E-commerce time series forecasting using LSTM neural network and support vector regression. Proceedings of International Conference on Big Data and Internet of Thing, 80-84.

Chun, Y. (2014). Analyzing space-time crime incidents using eigenvector spatial filtering: an application to vehicle burglary. *Geographical Analysis*, 46, 165-184.

Cook, P. J. and Durrance, C. P. (2013). The virtuous tax: lifesaving and crime-prevention effects of the 1991 federal alcohol-tax increase. *Journal of health economics*, 32, 261-267.

Corcoran, J. J., Wilson, I. D. and Ware, J. A. (2003). Predicting the geo-temporal variations of crime and disorder. *International Journal of Forecasting*, 19, 623-634.

Corder, G. W. and Foreman, D. I. (2014). *Nonparametric statistics: A step-by-step approach*, John Wiley and Sons.

Corman, H. and Mocan, H. N. (2000). A time-series analysis of crime, deterrence, and drug abuse in New York City. *The American economic review*, 90, 584-604.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273-297.

Coulibaly, P. and Baldwin, C. K. (2005). Nonstationary hydrological time series forecasting using nonlinear dynamic methods. *Journal of Hydrology*, 307, 164-174.

Das, P., Das, A. K. and Nayak, J. (2020). Feature selection generating directed rough-spanning tree for crime pattern analysis. *Neural Computing and Applications*, 32, 7623-7639.

Das, S. P. and Padhy, S. (2018). A novel hybrid model using teaching learning-based optimization and a support vector machine for commodity futures index forecasting. *International Journal of Machine Learning and Cybernetics*, 9, 97-111.

De Angelis, J., Benz, T. A. and Gillham, P. (2017). Collective Security, Fear of Crime, and Support for Concealed Firearms on a University Campus in the Western United States. *Criminal Justice Review*, 42, 77-94.

De Livera, A. M., Hyndman, R. J. and Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American statistical association*, 106, 1513-1527.

Deadman, D. (2003). Forecasting residential burglary. *International Journal of Forecasting*, 19, 567-578.

Ding, F., Ge, Q., Jiang, D., Fu, J. and Hao, M. (2017). Understanding the dynamics of terrorism events with multiple-discipline datasets and machine learning approach. *PloS one,* 12, Article e0179057.

Dingeman, K. and Rumbaut, R. G. (2010). The immigration-crime nexus and post-deportation experiences: en/countering stereotypes in Southern California and El Salvador. *University of La Verne Review Report*, 31(2), 363-402.

Dittman, D. J., Khoshgoftaar, T. M., Wald, R. and Van Hulse, J. (2010). Comparative analysis of DNA microarray data through the use of feature selection techniques. 9th IEEE International Conference on Machine Learning and Applications, 147-152.

Drew, P. J. and Monson, J. R. (2000). Artificial neural networks. *Surgery,* 127, 3-11.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A. and Vapnik, V. (1997). Support vector regression machines. *Advances in neural information processing systems*, 9, 155-161.

Du Preez, J. and Witt, S. F. (2003). Univariate versus multivariate time series forecasting: an application to international tourism demand. *International Journal of Forecasting*, 19, 435-451.

Du, S., Li, T., Yang, Y. and Horng, S.-J. (2020). Multivariate Time Series Forecasting via Attention-based Encoder-Decoder Framework. *Neurocomputing*, 388, 269-279.

Dubey, N. and Chaturvedi, S. K. (2014). A Survey Paper on Crime Prediction Technique Using Data Mining. *Int. Journal of Engineering Research and Applications*, 4(3), 396-400.

Ebrahimi, E., Monjezi, M., Khalesi, M. R. and Armaghani, D. J. (2016). Prediction and optimization of back-break and rock fragmentation using an artificial neural network and a bee colony algorithm. *Bulletin of Engineering Geology and the Environment,* 75, 27-36.

Egrioglu, E., Aladag, C. H., Yolcu, U., Uslu, V. R. and Basaran, M. A. (2009). A new approach based on artificial neural networks for high order multivariate fuzzy time series. *Expert Systems with Applications,* 36, 10589-10594.

Elith, J., Leathwick, J. R. and Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology,* 77, 802-813.

Entorf, H. and Winker, P. (2008). Investigating the drugs–crime channel in economics of crime models: Empirical evidence from panel data of the German States. *International Review of Law and Economics,* 28, 8-22.

Fajnzylber, P., Lederman, D. and Loayza, N. (2002). What causes violent crime? *European Economic Review,* 46, 1323-1357.

Fakhraei, S., Soltanian-Zadeh, H. and Fotouhi, F. (2014). Bias and stability of single variable classifiers for feature ranking and selection. *Expert systems with applications*, 41, 6945-6958.

Fang, L., Zhao, H., Wang, P., Yu, M., Yan, J., Cheng, W. and Chen, P. (2015). Feature selection method based on mutual information and class separability for dimension reduction in multidimensional time series for clinical data. *Biomedical Signal Processing and Control,* 21, 82-89.

Felson, M. and Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting*, 19, 595-601.

Felson, R. B. and Staff, J. (2017). Committing economic crime for drug money. *Crime and Delinquency*, 63, 375-390.

Feng, S., Zhou, H. and Dong, H. (2019). Using deep neural network with small dataset to predict material defects. *Materials and Design*, 162, 300-310.

Fergusson, D., Swain-Campbell, N. and Horwood, J. (2004). How does childhood economic disadvantage lead to crime? *Journal of Child Psychology and Psychiatry*, 45, 956-966.

Ferreira, J., João, P. and Martins, J. (2012). GIS for Crime Analysis-Geography for Predictive Models. *The Electronic Journal Information Systems Evaluation,* 15.

Fitzgerald, J., Curtis, K. A. and Corliss, C. L. (2012). Anxious publics: Worries about crime and immigration. *Comparative Political Studies*, 45, 477-506.

Flaxman, S. R. (2014). A General Approach to Prediction and Forecasting Crime Rates with Gaussian Processes. *Heinz College Second Paper Report*, 1-32.

Fonoberova, M., Fonoberov, V. A., Mezic, I., Mezic, J. and Brantingham, P. J. (2012). Nonlinear dynamics of crime and violence in urban settings. *Journal of Artificial Societies and Social Simulation*, 15, 2.

Foon, T, C. (2011). An exploration of dynamic relationship between tourist arrivals, inflation, unemployment and crime rates in Malaysia. *International Journal of Social Economics*, 38, 50-69.

Foucquier, A., Robert, S., Suard, F., Stéphan, L. and Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23, 272-288.

Freeman, E. A., Moisen, G. G., Coulston, J. W. and Wilson, B. T. (2015). Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. *Canadian Journal of Forest Research*, 46, 323-339.

Freilich, J. D., Adamczyk, A., Chermak, S. M., Boyd, K. A. and Parkin, W. S. (2015). Investigating the applicability of macro-level criminology theory to terrorism: A county-level analysis. *Journal of Quantitative Criminology*, 31, 383-411.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Froelich, W. and Salmeron, J. L. (2014). Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series. *International Journal of Approximate Reasoning*, 55, 1319-1335.

Gandomi, A. H., Yang, X.-S., Talatahari, S. and Alavi, A. H. (2013). Metaheuristic algorithms in modeling and optimization. *Metaheuristic applications in structures and infrastructures*, 1-24.

Ganjisaffar, Y., Caruana, R. and Lopes, C. V. (2011a). Bagging gradient-boosted trees for high precision, low variance ranking models. Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, 85-94.

Ganjisaffar, Y., Debeauvais, T., Javanmardi, S., Caruana, R. and Lopes, C. V. (2011b) Distributed tuning of machine learning algorithms using MapReduce clusters. Proceedings of the Third ACM Workshop on Large Scale Data Mining: Theory and Applications, 2.

Gao, Y., Pan, J., Ji, G. and Gao, F. (2011). A time-series modeling method based on the boosting gradient-descent theory. *Science China Technological Sciences*, 54, 1325.

Garcia, R. C., Contreras, J., Van Akkeren, M. and Garcia, J. B. C. (2005). A GARCH forecasting model to predict day-ahead electricity prices. *IEEE transactions on power systems*, 20, 867-874.

Genuer, R., Poggi, J. M. and Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31, 2225-2236.

Ghazvini, A., Nazri, M. Z. B. A., Abdullah, S. N. H. S., Junoh, M. N. and Bin Kasim, Z. A. (2015). Biography commercial serial crime analysis using enhanced dynamic neural network. IEEE 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR), 334-339.

Ghosh, B., Basu, B. and O'mahony, M. (2009). Multivariate short-term traffic flow forecasting using time-series analysis. *IEEE transactions on intelligent transportation systems,* 10, 246-254.

Gill, M. K., Kaheil, Y. H., Khalil, A., Mckee, M. and Bastidas, L. (2006). Multiobjective particle swarm optimization for parameter estimation in hydrology. *Water Resources Research,* 42.

Gordini, N. and Veglio, V. (2017). Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. *Industrial Marketing Management,* 62, 100-107.

Gorr, W. and Harries, R. (2003). Introduction to crime forecasting. *International Journal of Forecasting,* 19, 551-555.

Gorr, W., Olligschlaeger, A. and Thompson, Y. (2003). Short-term forecasting of crime. *International Journal of Forecasting*, 19, 579-594.

Goulas, E. and Zervoyianni, A. (2013). Economic growth and crime: does uncertainty matter? *Applied Economics Letters*, 20, 420-427.

Greenberg, D. F. (2001). Time series analysis of crime rates. *Journal of quantitative criminology*, 17, 291-327.

Guelman, L. (2012). Gradient boosting trees for auto insurance loss cost modeling and prediction. *Expert Systems with Applications*, 39, 3659-3667.

Gutierrez Rufrancos, H., Power, M., Pickett, K. E. and Wilkinson, R. (2013). Income inequality and crime: a review and explanation of the time-series evidence. *Sociology and criminiology,* 1(1), 1-9.

Halicioglu, F., Andrés, A. R. and Yamamura, E. (2012). Modeling crime in Japan. *Economic Modelling*, 29, 1640-1645.

Han, M. and Wang, Y. (2009). Analysis and modeling of multivariate chaotic time series based on neural network. *Expert Systems with Applications,* 36, 1280-1290.

Hanslmaier, M., Kemme, S., Stoll, K. and Baier, D. (2015). Forecasting crime in Germany in times of demographic change. *European Journal on Criminal Policy and Research,* 21, 591-610.

Hapfelmeier, A. and Ulm, K. (2013). A new variable selection approach using Random Forests. *Comput. Stat. Data Anal.,* 60, 50-69.

Hart, T. and Zandbergen, P. (2014). Kernel density estimation and hotspot mapping: examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting. *Policing: An International Journal of Police Strategies and Management,* 37, 305-323.

He, X. and Xu, S. (2010). *Process neural networks: Theory and applications*. Springer Science & Business Media, First Edition.

Hipp, J. R. and Yates, D. K. (2011). Ghettos, thresholds, and crime: Does concentrated poverty really have an accelerating increasing effect on crime? *Criminology,* 49, 955-990.

Hira, Z. M. and Gillies, D. F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in bioinformatics,* 2015, 1-14.

Hong, W.-C. (2011). Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. *Energy,* 36, 5568-5578.

Hou, S. and Li, Y. (2009). Short-term fault prediction based on support vector machines with parameter optimization by evolution strategy. *Expert Systems with Applications,* 36, 12383-12391.

Hou, Y., Zhao, L. and Lu, H. (2018). Fuzzy neural network optimization and network traffic forecasting based on improved differential evolution. *Future Generation Computer Systems,* 81, 425-432.

Hua, J., Tembe, W. D. and Dougherty, E. R. (2009). Performance of feature-selection methods in the classification of high-dimension data. *Pattern Recognition*, 42, 409-424.

Huang, Y.-L., Lin, C.-T., Yu, Y.-S., Hsieh, W.-H. and Pai, S.-M. (2015). Crime Forecasting Based on BP Neural Network. National Conference on Information Technology Practice and Application (NCITPA), China, 55-63.

Huarng, K.-H., Yu, T. H.-K. and Hsu, Y. W. (2007). A multivariate heuristic model for fuzzy time-series forecasting. IEEE Transactions on Systems, Man, and Cybernetics, (Cybernetics), 37, 836-846.

Huddleston, S. H., Porter, J. H. and Brown, D. E. (2015). Improving forecasts for noisy geographic time series. *Journal of Business Research*, 68, 1810-1818.

Ingilevich, V. and Ivanov, S. (2018). Crime rate prediction in the urban environment using social factors. *Procedia Computer Science*, 136, 472-478.

Iqbal, R., Murad, M. A. A., Mustapha, A., Panahy, P. H. S. and Khanahmadliravi, N. (2013). An experimental study of classification algorithms for crime prediction. *Indian Journal of Science and Technology*, 6, 4219-4225.

Islam, M. (2017). Income Inequality and Economic Growth Nexus in Japan: A Multivariate Analysis. *The Ritsumeikan Economic Review*, 65(4), 439-456.

Ismail, S. and Ramli, N. (2013). Short-term crime forecasting in Kedah. *Procedia-Social and Behavioral Sciences*, 91, 654-660.

Jahrer, M., Töscher, A. and Legenstein, R. (2010). Combining predictions for accurate recommender systems. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 693-702.

Jain, A. and Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, 7, 585-592.

Jalabert, S., Martin, M., Renaud, J. P., Boulonne, L., Jolivet, C., Montanarella, L. and Arrouays, D. (2010). Estimating forest soil bulk density using boosted regression modelling. *Soil use and management,* 26, 516-528.

James, A. and Smith, B. (2017). There will be blood: crime rates in shale-rich US counties. *Journal of Environmental Economics and Management*, 84, 125-152.

Jiang, H., Zou, Y., Zhang, S., Tang, J. and Wang, Y. (2016). Short-term speed prediction using remote microwave sensor data: machine learning versus statistical model. *Mathematical Problems in Engineering*, 2016.

Jilani, T. A., Burney, S. A. And Ardil, C. (2007). Multivariate high order fuzzy time series forecasting for car road accidents. *International Journal of Computational Intelligence*, 4, 15-20.

Jilani, T. A. and Burney, S. M. A. (2007). M-factor high order fuzzy time series forecasting for road accident data. *Analysis and design of intelligent systems using soft computing techniques.* Springer.

Jiménez, F., Palma, J., Sánchez, G., Marín, D., Palacios, F. and López, L. (2020). Feature selection based multivariate time series forecasting: An application to antibiotic resistance outbreaks prediction. *Artificial Intelligence in Medicine*, 101818.

Jin, M. and Deng, W. (2018). Predication of different stages of Alzheimer's disease using neighborhood component analysis and ensemble decision tree. *Journal of neuroscience methods*, 302, 35-41.

Jones, S. S., Evans, R. S., Allen, T. L., Thomas, A., Haug, P. J., Welch, S. J. and Snow, G. L. (2009). A multivariate time series approach to modeling and forecasting demand in the emergency department. *Journal of biomedical informatics,* 42, 123-139.

Ju, X., Cheng, M., Xia, Y., Quo, F. and Tian, Y. (2014). Support Vector Regression and Time Series Analysis for the Forecasting of Bayannur's Total Water Requirement. *Procedia Computer Science,* 31, 523-531.

Junbo, K., Xinyue, L. and Jiajia, C. (2015). San Francisco Crime Classification. *CSE 255 Report.* San Diego: University of California.

Kadar, C., Maculan, R. & Feuerriegel, S. (2019). Public decision support for low population density areas: An imbalance-aware hyper-ensemble for spatio-temporal crime prediction. *Decision Support Systems*, 119, 107-117.

Kalekar, P. S. (2004). Time series forecasting using holt-winters exponential smoothing. *Kanwal Rekhi School of Information Technology,* 4329008, 1-13.

Kapuscinski, C. A., Braithwaite, J. and Chapman, B. (1998). Unemployment and crime: Toward resolving the paradox. *Journal of Quantitative Criminology,* 14, 215-243.

Karamizadeh, S., Abdullah, S. M., Halimi, M., Shayan, J. and Javad Rajabi, M. (2014). Advantage and drawback of support vector machine functionality. IEEE International Conference on Computer, Communications and Control Technology (I4CT), 63-65.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 2017, 3146-3154.

Kecman, V., Huang, T.-M. and Vogt, M. (2005). Iterative single data algorithm for training kernel machines from huge data sets: Theory and performance. *Support vector machines: Theory and Applications*, 255-274.

Khaidem, L., Saha, S. and Dey, S. R. (2016). Predicting the direction of stock market prices using random forest. *Applied Mathematical Finance*, 1-20.

Khashei, M. and Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*, 11, 2664-2675.

Kianmehr, K. and Alhajj, R. (2006). Crime hot-spots prediction using support vector machine. IEEE International Conference on Computer Systems and Applications, 952-959.

Kim, K.-J. and Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, 19, 125-132.

Kim, T., Lee, D., Choi, J., Spurlock, A., Sim, A., Todd, A. and Wu, K. (2015). Extracting baseline electricity usage using gradient tree boosting. IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), 734-741.

Klaer, J. and Northrup, B. (2014). Effects of GDP on Violent Crime. *Econometric Analysis Undergraduate Research Papers*. School of Economics, Ivan Allen College (IAC).

Koop, G. (2008). *An introduction to econometrics*, John Wiley and Sons.

Koper, C. S., Guterbock, T. M., Woods, D. J., Taylor, B. and Carter, T. J. (2013). The effects of local immigration enforcement on crime and disorder. *Criminology and Public Policy*, 12, 239-276.

Kouziokas, G. N. (2017). The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment. *Transportation Research Procedia*, 24, 467-473.

Kriesel, D. (2007). *A Brief Introduction to Neural Networks*. Zeta Edition. Available Online : http://www.dkriesel.com/en/science/neural_networks. Accessed : March 2018.

Kržić, A. S. and Seršić, D. (2018). L1 minimization using recursive reduction of dimensionality. *Signal Processing*, 151, 119-129.

Kumar, K. S. and Bhalaji, N. (2016). A study on classification algorithms for crime records. International Conference on Smart Trends for Information Technology and Computer Communications, 873-880.

Lafree, G. and Tseloni, A. (2006). Democracy and crime: A multilevel analysis of homicide trends in forty-four countries, 1950-2000. *The Annals of the American Academy of Political and Social Science*, 605, 25-49.

Lauritsen, J. L. and Cork, D. L. (2016a). Modernizing Crime Statistics: Report 1: Defining and Classifying Crime, Washington, DC, The National Academies Press.

Lauritsen, J. L., Rezey, M. L. and Heimer, K. (2016b). When Choice of Data Matters: Analyses of U.S. Crime Trends, 1973–2012. *Journal of Quantitative Criminology*, 32, 335-355.

Li, C., Tao, Y., Ao, W., Yang, S. and Bai, Y. (2018). Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. *Energy*, 165, 1220-1227.

Li, S.-T., Kuo, S.-C. and Tsai, F.-C. (2010). An intelligent decision-support model using FSOM and rule extraction for crime prevention. *Expert Systems with Applications*, 37, 7108-7119.

Linden, A. and Yarnold, P. R. (2016). Using machine learning to identify structural breaks in single-group interrupted time series designs. *Journal of Evaluation in Clinical Practice*, 22, 855-859.

Lindley, DV. (1990). *Regression and correlation analysis*, in Palgrave, M. (ed.) *Time Series and Statistics*, London:Springer, 237-243.

Liu, G., Zhou, D., Xu, H. and Mei, C. (2010). Model optimization of SVM for a fermentation soft sensor. *Expert Systems with Applications*, 37, 2708-2713.

Liu, H., Yang, C., Zhang, M., Mcloone, S. and Sun, Y. (2017). A computational intelligent approach to multi-factor analysis of violent crime information system. *Enterprise Information Systems*, 11, 161-184.

Liu, R., Xu, S., Fang, C., Liu, Y.-W., Murphey, Y. L. and Kochhar, D. S. (2012). Statistical modeling and signal selection in multivariate time series pattern classification. Proceedings of IEEE 21st International Conference on Pattern Recognition (ICPR), 2853-2856.

Liu, X., Shen, C., Wang, W. and Guan, X. (2019). CoEvil: A Coevolutionary Model for Crime Inference Based on Fuzzy Rough Feature Selection. IEEE Transactions on Fuzzy Systems, 28, 806-817.

Lu, C.-J., Lee, T.-S. and Chiu, C.-C. (2009). Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems,* 47, 115-125.

Luthra, I., Chaturvedi, S. K., Upadhyay, D. and Gupta, R. (2017). Comparative study on nature inspired algorithms for optimization problem. IEEE International conference of Electronics, Communication and Aerospace Technology (ICECA), 143-147.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*, Springer Science and Business Media, First Edition.

Macdonald, J. M., Hipp, J. R. and Gill, C. (2013). The effects of immigrant concentration on changes in neighborhood crime rates. *Journal of Quantitative Criminology*, 29, 191-215.

Mahmud, N., Zinnah, K. I., Rahman, Y. A. and Ahmed, N. (2016). Crimecast: A crime prediction and strategy direction service. IEEE 19th International Conference on Computer and Information Technology (ICCIT), 414-418.

Malan, N. S. and Sharma, S. (2019). Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals. *Computers in biology and medicine*, 107, 118-126.

Martinez Jr, R. and Stowell, J. I. (2012). Extending immigration and crime studies: national implications and local settings. *The Annals of the American Academy of Political and Social Science,* 641, 174-191.

Martinez, R., Stowell, J. I. and Lee, M. T. (2010). Immigration and crime in an era of transformation: A longitudinal analysis of homicides in San Diego neighborhoods, 1980–2000. *Criminology,* 48, 797-829.

Matheny, M. E., Resnic, F. S., Arora, N. and Ohno-Machado, L. (2007). Effects of SVM parameter optimization on discrimination and calibration for post-procedural PCI mortality. *Journal of biomedical informatics*, 40, 688-697.

Mayrink, V. and Hippert, H. S. (2016). A hybrid method using Exponential Smoothing and Gradient Boosting for electrical short-term load forecasting. IEEE Latin American Conference on Computational Intelligence, 1-6.

McClendon, L. and Meghanathan, N. (2015). Using Machine Learning Algorithms to Analyze Crime Data. *Machine Learning and Applications: An International Journal (MLAIJ),* 2(1), 1-12.

McCulloch, W.S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.

McDowall, D. and Loftin, C. (2005). Are US crime rate trends historically contingent? *Journal of Research in Crime and Delinquency*, 42, 359-383.

Mirjalili, S. (2016). Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Computing and Applications*, 27, 1053-1073.

Mitchell, M. B., Brown, D. E. and Conklin, J. H. (2007). A crime forecasting tool for the web-based crime analysis toolkit. IEEE Symposium of Systems and Information Engineering Design Symposium (SIEDS), 1-5.

Mittal, M., Goyal, L. M., Sethi, J. K. and Hemanth, D. J. (2019). Monitoring the Impact of Economic Crisis on Crime in India Using Machine Learning. *Computational Economics*, 53, 1467-1485.

Mohd, F. and Noor, N. M. M. (2017). A comparative study to evaluate filtering methods for crime data feature selection. *Procedia computer science*, 116, 113-120.

Mohler, G. (2014). Marked point process hotspot maps for homicide and gun crime prediction in Chicago. *International Journal of Forecasting*, 30, 491-497.

Mori, Y., Iizuka, M., Tarumi, T. and Tanaka, Y. (2007). Variable selection in principal component analysis. *Statistical methods for biostatistics and related fields*, 265-283.

Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7, 21.

Nayak, P. C., Sudheer, K., Rangan, D. and Ramasastri, K. (2004). A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology*, 291, 52-66.

Nawi, N. M., Atomi, W. H. and Rehman, M. Z. (2013). The effect of data pre-processing on optimized training of artificial neural networks. *Procedia Technology*, 11, 32-39.

Necaise, C. M. (2013). *Effects of Illegal Immigration Upon Crime In the United States*. Bachelors of Arts, The University of Southern Mississippi.

NG, S. (2013). Variable selection in predictive regressions. *Handbook of economic forecasting*, 2, 752-789.

Nguyen, T. T., Hatua, A. and Sung, A. H. (2017). Building a Learning Machine Classifier with Inadequate Data for Crime Prediction. *Journal of Advances in Information Technology Vol*, 8.

Nie, J. (1997). Nonlinear time-series forecasting: A fuzzy-neural approach. *Neurocomputing*, 16, 63-76.

Nolan, J. J. (2004). Establishing the statistical relationship between population size and UCR crime rate: Its impact and implications. *Journal of Criminal Justice*, 32, 547-555.

Nordfjaern, T. (2017). Violence involvement among nightlife patrons: the relative role of demographics and substance use. *Aggressive behavior journal*, 43(4), 398-407.

Ogutu, J. O., Piepho, H.-P. and Schulz-Streeck, T. (2011). A comparison of random forests, boosting and support vector machines for genomic selection. BMC Proceedings of the 14th European workshop on QTL mapping and marker assisted selection (QTL-MAS), 1-5.

Okawa, M. (2020). Online signature verification using single-template matching with time-series averaging and gradient boosting. *Pattern Recognition*, 102, 107227.

Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petković, D. and Sudheer, C. (2015). A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Solar Energy*, 115, 632-644.

Oliveira, A. L., Braga, P. L., Lima, R. M. and Cornélio, M. L. (2010). GA-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation. *information and Software Technology*, 52, 1155-1166.

Oliveira, J. F. L. and Ludermir, T. B. (2014). A distributed PSO-ARIMA-SVR hybrid system for time series forecasting. IEEE International Conference on Systems, Man and Cybernetics (SMC), 3867-3872.

Olligschlaeger, A. M. (1997). Artificial neural networks and crime mapping. *Crime mapping and crime prevention*, 313-348.

Omar, A., Kwanbunbumpen, A. and Alijani, D. (2007). Forecasting Computer Crime Complaints. *Information System Education Journal*, 15.

Osho, G. S. and Daramola, A (2010). An empirical infusion of time series crime forecasting modeling into crime prevention approach for use by law enforcement agencies: An econometrics perspective. *National Forum Of Multicultural Issues Journal*, 1-8.

Parejo, J. A., Ruiz-Cortés, A., Lozano, S. and Fernandez, P. (2012). Metaheuristic optimization frameworks: a survey and benchmarking. *Soft Computing,* 16, 527-561.

Parmezan, A. R. S., Souza, V. M. and Batista, G. E. (2019). Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. *Information Sciences*, 484, 302-337.

Patton, A. J. 2012. Copula methods for forecasting multivariate time series. *Handbook of economic forecasting,* 2, 899-960.

Pedregosa, F., Gaël, V., Alexandre, G., Vincent, M., Bertrand, T., Olivier, G., Mathieu, B., Peter, P., Ron, W., Vincent, D., Jake, V., Alexandre, P., David, C., Matthieu, B., Matthieu, P. and Édouard, D. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12, 2825-2830.

Peng, H., Ozaki, T., Haggan-Ozaki, V. and Toyoda, Y. (2003). A parameter optimization method for radial basis function type models. *IEEE Transactions on Neural Networks,* 14, 432-438.

Persson, C., Bacher, P., Shiga, T. and Madsen, H. (2017). Multi-site solar power forecasting using gradient boosted regression trees. *Solar Energy,* 150, 423-436.

Peterson, R. R. (2011). Employment, unemployment, and rates of intimate partner violence: Evidence from the National Crime Victim Surveys. *Economic Crisis and Crime.* Emerald Group Publishing Limited.

Pinotti, P. (2011). The economic consequences of organized crime: Evidence from Southern Italy. *The Economic Journal*, 125(586), 203-232.

Piza, E. L. (2012). *Using Poisson and negative binomial regression models to measure the influence of risk on crime incident counts*. Newark, NJ : Rutgers Center on Public Security.

Poutvaara, P. and Priks, M. (2011). Unemployment and gang crime: can prosperity backfire? *Economics of Governance,* 12, 259-273.

Prastyo, D. D., Nabila, F. S., Lee, M. H., Suhermi, N. and Fam, S.-F. (2018). VAR and GSTAR-based feature selection in support vector regression for multivariate spatio-temporal forecasting. International Conference on Soft Computing in Data Science, 46-57.

Pratt, T. C. and Cullen, F. T. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and justice*, 32, 373-450.

Qi, C., Fourie, A. and Zhao, X. (2018). Back-analysis method for stope displacements using gradient-boosted regression tree and firefly algorithm. *Journal of Computing in Civil Engineering*, 32, 04018031.

Qin, C., Song, S., Huang, G. and Zhu, L. (2015). Unsupervised neighborhood component analysis for clustering. *Neurocomputing*, 168, 609-617.

Qiu, J., Wu, Q., Ding, G., Xu, Y. and Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, 2016, 67.

Qu, J., Zheng, C., Zhao, W., Wang, X. and Zhang, S. (2010). Parallel Model of Forecasting Killer Residence and Place of Crime. IEEE Ninth International Symposium on Distributed Computing and Applications to Business Engineering and Science (DCABES), 228-230.

Raghu, S. and Sriraam, N. (2018). Classification of focal and non-focal EEG signals using neighborhood component analysis and machine learning algorithms. *Expert Systems with Applications*, 113, 18-32.

Ramadas, G. C., Fernandes, E. M., Ramadas, A., Rocha, A. M. A. and Costa, M. F. P. (2018). On Metaheuristics for Solving the Parameter Estimation Problem in Dynamic Systems: A Comparative Study. *Journal of Optimization*, 2018, 21.

Ranson, M. (2014). Crime, weather, and climate change. *Journal of environmental economics and management*, 67, 274-302.

Raphael, S. and Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. *The Journal of Law and Economics*, 44, 259-283.

Rather, A. M., Sastry, V. and Agarwal, A. (2017). Stock market prediction and Portfolio selection models: a survey. *OPSEARCH*, 1-22.

Redmond, M. and Baveja, A. (2002). A data-driven software tool for enabling cooperative information sharing among police departments. *European Journal of Operational Research*, 141, 660-678.

Ren, C., An, N., Wang, J., Li, L., Hu, B. and Shang, D. (2014). Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting. *Knowledge-Based Systems,* 56, 226-239.

Renushe, H. N., Joshi, M. J., Kumbhar, R. and Desai, A. (2011). Short term crime forecasting for prevention of crimes: A study of Satara district. *International Journal of Computer Technology and Applications,* 2.

Ridgeway, G. (2013). Generalized Boosted Models: A guide to the gbm package. *R Package Update.*

Rivas, V. M., Merelo, J., Castillo, P., Arenas, M. G. and Castellano, J. (2004). Evolving RBF neural networks for time-series forecasting with EvRBF. *Information Sciences,* 165, 207-220.

Robnik-Šikonja, M. and Kononenko, I. (2003). Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning,* 53, 23-69.

Rosenfeld, R. (2009). Crime is the problem: Homicide, acquisitive crime, and economic conditions. *Journal of Quantitative Criminology,* 25, 287-306.

Rosenfeld, R. and Fornango, R. (2007). The impact of economic conditions on robbery and property crime: the role of consumer sentiment. *Criminology,* 45, 735-769.

Rosenfeld, R. and Levin, A. (2016). Acquisitive crime and inflation in the United States: 1960–2012. *Journal of Quantitative Criminology*, 32, 427-447.

Rosenfeld, R., Vogel, M. and Mccuddy, T. (2019). Crime and Inflation in U. S. Cities. *Journal of Quantitative Criminology*, 35, 195-210.

Rotolo, T. and Tittle, C. R. (2006). Population size, change, and crime in US cities. *Journal of Quantitative Criminology*, 22, 341-367.

Ruta, D. and Gabrys, B. (2007). Neural network ensembles for time series prediction. IEEE International Joint Conference on Neural Networks, 1204-1209.

Saha, D., Alluri, P. and Gan, A. (2015). Prioritizing Highway Safety Manual's crash prediction variables using boosted regression trees. *Accident Analysis and Prevention,* 79, 133-144.

Said, G. A. E.-N. A., Mahmoud, A. M. and El-Horbaty, E.-S. M. (2014). A comparative study of meta-heuristic algorithms for solving quadratic assignment problem. *International Journal of Advanced Computer Science and Applications (IJACSA),* 5, 1-6.

Sallehuddin, R., Shamsuddin, S. M. and Hashim, S. Z. M. (2008). Hybridization model of linear and nonlinear time series data for forecasting. IEEE 2nd Asia International Conference on Modelling and Simulation (AMS), 597-602.

Santitissadeekorn, N., Lloyd, D. J. and Short, M. B. (2015). Exploring data assimilation and forecasting issues for an urban crime model. *European Journal of Applied Mathematics*, 27(3), 451-478.

Sapankevych, N. I. and Sankar, R. (2009). Time series prediction using support vector machines: a survey. *IEEE Computational Intelligence Magazine,* 4.

Sathyadevan, S., M. S, D. and Gangadharan, S. (2014). Crime analysis and prediction using data mining. IEEE First International Conference on Networks and Soft Computing (ICNSC), 406-412.

Saridakis, G. (2004). Violent crime in the United States of America: A time-series analysis between 1960-2000. *European Journal of Law and Economics*, 18, 203-221.

Sathyadevan, S., M. S, D. and Gangadharan, S. (2014). Crime analysis and prediction using data mining. IEEE First International Conference on Networks and Soft Computing (ICNSC), 406-412.

Savsani, P., Jhala, R. and Savsani, V. J. (2016). Comparative study of different metaheuristics for the trajectory planning of a robotic arm. *IEEE Systems Journal*, 10, 697-708.

Schliebs, S., Defoin-Platel, M., Worner, S. and Kasabov, N. (2009). Integrated feature and parameter optimization for an evolving spiking neural network: Exploring heterogeneous probabilistic models. *Neural Networks,* 22, 623-632.

Seals, A. and Nunley, J. (2007). The effects of inflation and demographic change on property crime: A structural time series approach. *Department of Economics and Finance Working Paper Series. Middle Tennessee State University, Murfreesboro, TN,* 1-39.

Seijo, P. B., Alonso, B. A., Bennett, K. P., Bolón, C. V., Josse, J., Saeed, M. and Guyon, I. (2019). Biases in feature selection with missing data. *Neurocomputing*, 342, 97-112.

Shah Habibullah, M. and Baharom, A. H. (2009). Crime and economic conditions in Malaysia. *International Journal of Social Economics,* 36, 1071-1081.

Shao, Y. E. and Dai, J.-T. (2018). Integrated feature selection of ARIMA with computational intelligence approaches for food crop price prediction. *Complexity*, 2018.

Sharadga, H., Hajimirza, S. and Balog, R. S. (2020). Time series forecasting of solar power generation for large-scale photovoltaic plants. *Renewable Energy*, 150, 797-807.

Shi, T. (2020). A Method of Predicting Crime of Theft Based on Bagging Ensemble Feature Selection. IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), 140-143.

Shi, X., Paiement, J.-F., Grangier, D. And Yu, P. S. (2012). Learning from heterogeneous sources via gradient boosting consensus. Proceedings of the SIAM International Conference on Data Mining, 224-235.

Shoesmith, G. L. (2013). Space–time autoregressive models and forecasting national, regional and state crime rates. *International Journal of Forecasting*, 29, 191-201.

Shri, T. P. and Sriraam, N. (2017). Comparison of t-test ranking with PCA and SEPCOR feature selection for wake and stage 1 sleep pattern recognition in multichannel electroencephalograms. *Biomedical Signal Processing and Control*, 31, 499-512.

Shrivastav, A. K. and Ekata, D. (2012). Applicability of Soft computing technique for Crime Forecasting: A Preliminary Investigation. *International Journal of Computer Science and Engineering Technology*, 415-421.

Song, A., Wenzel, S. L., Kim, J. Y. and Nam, B. (2017). Experience of domestic violence during childhood, intimate partner violence, and the deterrent effect of awareness of legal consequences. *Journal of interpersonal violence*, 32, 357-372.

Song, J., Wang, J. and Lu, H. (2018). A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. *Applied Energy*, 215, 643-658.

Soundarya, V., Kanimozhi, U. and Manjula, D. (2017). Recommendation System for Criminal Behavioral Analysis on Social Network using Genetic Weighted K-Means Clustering. *JCP*, 12, 212-220.

Souza, A. J., Borges, A. P., Gomes, H. M., Barddal, J. P. and Enembreck, F. (2015). Applying Ensemble-based Online Learning Techniques on Crime Forecasting. ICEIS, 17-24.

Stansfield, R. (2014). Safer cities: A macro-level analysis of recent immigration, Hispanic-owned businesses, and crime rates in the United States. *Journal of Urban Affairs*, 36, 503-518.

Stansfield, R., Williams, K. R. and Parker, K. F. (2017). Economic disadvantage and homicide: estimating temporal trends in adolescence and adulthood. *Homicide studies,* 21, 59-81.

Stergiou, K., Christou, E. and Petrakis, G. (1997). Modelling and forecasting monthly fisheries catches: comparison of regression, univariate and multivariate time series methods. *Fisheries Research,* 29, 55-95.

Stowell, J. I., Messner, S. F., Mcgeever, K. F. and Raffalovich, L. E. (2009). Immigration and the recent violent crime drop in the united states: a pooled, cross-sectional time-series analysis of metropolitan areas. *Criminology,* 47, 889-928.

Strobl, C., Boulesteix, A.-L., Zeileis, A. and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics,* 8, 25.

Su, L.-Y. (2010). Prediction of multivariate chaotic time series with local polynomial fitting. *Computers and Mathematics with Applications*, 59, 737-744.

Sun, I. Y., Chu, D. C. and Sung, H.-E. (2011). A cross-national analysis of the mediating effect of economic deprivation on crime. *Asian Journal of Criminology*, 6, 15-32.

Sun, S., Peng, Q. and Shakoor, A. (2014). A kernel-based multivariate feature selection method for microarray data classification. *PloS one*, 9, e102541.

Sun, Y., Li, J., Liu, J., Chow, C., Sun, B. and Wang, R. (2015). Using causal discovery for feature selection in multivariate numerical time series. *Machine Learning*, 101, 377-395.

Tan, C., Wang, T., Yang, W. and Deng, L. (2020). PredPSD: A Gradient Tree Boosting Approach for Single-Stranded and Double-Stranded DNA Binding Protein Prediction. *Molecules*, 25, 98.

Tantithamthavorn, C., Mcintosh, S., Hassan, A. E. and Matsumoto, K. (2016). Automated parameter optimization of classification techniques for defect prediction models. IEEE/ACM 38th International Conference on Software Engineering (ICSE), 321-332.

Tiwari, S., Naresh, R. and Jha, R. (2013). Comparative study of backpropagation algorithms in neural network based identification of power system. *International Journal of Computer Science and Information Technology*, 5, 93.

Tomin, N., Zhukov, A., Sidorov, D., Kurbatsky, V., Panasetsky, D. and Spiryaev, V. (2015). Random forest based model for preventing large-scale emergencies in power systems. *International Journal of Artificial Intelligence*, 13, 211-228.

Tsai, C.-F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22, 120-127.

Tuncer, T. and Ertam, F. (2020). Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma. *Physica A: Statistical Mechanics and its Applications*, 540, 123143.

Vaquero, B. M. (2016). *Machine learning applied to crime prediction*. Bachelor Thesis, Universitat Politècnica de Catalunya.

Verikas, A., Kalsyte, Z., Bacauskiene, M. and Gelzinis, A. (2010). Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: a survey. *Soft Computing*, 14, 995-1010.

Vijayakumar, M., Karthick, S. and Prakash, N. (2013). The Day-To-Day Crime Forecasting Analysis of Using Spatial temporal Clustering Simulation. *Int J Sci Eng Res*, 4, 1-6.

Vineeth, K. S., Pandey, A. and Pradhan, T. (2016). A novel approach for intelligent crime pattern discovery and prediction. IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 531-538.

Vlahogianni, E. I., Karlaftis, M. G. and Golias, J. C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach. *Transportation Research Part C: Emerging Tech*, 13, 211-234.

Wan, J. D., Zhang, Y. B., Yu, W. J. and He, C. (2010). A New Model for Forecasting the Locations of Next Crime. *Trans Tech Publication of Applied Mechanics and Materials*, 1116-1121.

Wang, B., Yin, P., Bertozzi, A. L., Brantingham, P. J., Osher, S. J. and Xin, J. (2019a). Deep learning for real-time crime forecasting and its ternarization. *Chinese Annals of Mathematics Series B*, 40, 949-966.

Wang, J., Yang, W., Du, P. and Li, Y. (2018). Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. *Energy*, 148, 59-78.

Wang, K., Zhu, P., Zhu, H., Cui, P. and Zhang, Z. (2017). An Interweaved Time Series Locally Connected Recurrent Neural Network Model on Crime Forecasting. International Conference on Neural Information Processing, 466-474.

Wang, L., Wang, Z. and Liu, S. (2016). An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm. *Expert Systems with Applications*, 43, 237-249.

Wang, P., Mathieu, R., Ke, J. and Cai, H. (2010). Predicting criminal recidivism with support vector machine. IEEE International Conference on Management and Service Science (MASS), 1-9.

Wang, Q., Jin, G., Zhao, X., Feng, Y. and Huang, J. (2019b). CSAN: A neural network benchmark model for crime forecasting in spatio-temporal scale. *Knowledge-Based Systems*, 189, 105120.

Wang, Z. and Bessler, D. A. (2004). Forecasting performance of multivariate time series models with full and reduced rank: An empirical examination. *International Journal of Forecasting*, 20, 683-695.

Winfree Jr, L. T., Taylor, T. J., He, N. and Esbensen, F.-A. (2006). Self-control and variability over time: Multivariate results using a 5-year, multisite panel of youths. *Crime and Delinquency*, 52, 253-286.

Witten, I. H., Frank, E., Hall, M. A. and Pal, C. J. (2016). Data Mining: Practical machine learning tools and techniques, *Acm Sigmod Record*, 31(1), 76-77.

Worrall, J. L. and Pratt, T. C. (2004). On the consequences of ignoring unobserved heterogeneity when estimating macro-level models of crime. *Social Science Research*, 33, 79-105.

Wu, C.-H., Tzeng, G.-H., Goo, Y.-J. and Fang, W.-C. (2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert systems with applications*, 32, 397-408.

Wu, C.-H., Tzeng, G.-H. and Lin, R.-H. (2009). A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. *Expert Systems with Applications,* 36, 4725-4735.

Wu, J. and Lu, Z. (2012). A novel hybrid genetic algorithm and simulated annealing for feature selection and kernel optimization in support vector regression. IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI), 999-1003.

Wu, Z., Efros, A. A. and Yu, S. X. (2018). Improving generalization via scalable neighborhood component analysis. Proceedings of the European Conference on Computer Vision (ECCV), 685-701.

Xiao, H., Pei, W., Dong, Z., Kong, L. and Wang, D. (2018). Application and Comparison of Metaheuristic and New Metamodel Based Global Optimization Methods to the Optimal Operation of Active Distribution Networks. *Energies,* 11, 85.

Xu, H. and Niimura, T. (2004). Short-term electricity price modeling and forecasting using wavelets and multivariate time series. IEEE Power Systems Conference and Exposition (PES), 208-212.

Xu, M., Han, M. and Wang, X. (2016). Hierarchical neural networks for multivariate time series prediction. IEEE 35th Chinese Control Conference (CCC), 6971-6976.

Yang, F., Wu, C., Xiong, N. and Wu, Y. (2018). Prediction of criminal tendency of high-risk personnel based on combination of principal component analysis and support vector machine. *Internatioanl Journal of Software and Hardware Research in Engineering*, 6, 1-10.

Yang, K. and Shahabi, C. (2007). An efficient k nearest neighbor search for multivariate time series. *Information and Computation,* 205, 65-98.

Yang, K., Yoon, H. and Shahabi, C. (2005). CLeVer: A Feature Subset Selection Technique for Multivariate Time Series. PAKDD, Springer, 516-522.

Yang, S., Wu, J., Du, Y., He, Y. and Chen, X. (2017). Ensemble Learning for Short-Term Traffic Prediction Based on Gradient Boosting Machine. *Journal of Sensors*, 2017.

Yang, W., Wang, K. and Zuo, W. (2012a). Fast neighborhood component analysis. *Neurocomputing*, 83, 31-37.

Yang, W., Wang, K. and Zuo, W. (2012b). Neighborhood Component Feature Selection for High-Dimensional Data. *JCP,* 7, 161-168.

Yang, Y. and Zou, H. (2015). Nonparametric multiple expectile regression via ER-Boost. *Journal of Statistical Computation and Simulation,* 85, 1442-1458.

Yang, Y., Che, J., Deng, C. and Li, L. (2019). Sequential grid approach based support vector regression for short-term electric load forecasting. *Applied energy*, 238, 1010-1021.

Yang, Y. and Zou, H. (2015). Nonparametric multiple expectile regression via ER-Boost. *Journal of Statistical Computation and Simulation*, 85, 1442-1458.

Ye, J., Chow, J.-H., Chen, J. and Zheng, Z. (2009). Stochastic gradient boosted distributed decision trees. Proceedings of the 18th ACM conference on Information and knowledge management, 2061-2064.

Yearwood, D. L. and Koinis, G. (2011). Revisiting property crime and economic conditions: An exploratory study to identify predictive indicators beyond unemployment rates. *The Social Science Journal,* 48, 145-158.

Yeh, W.-C. (2013). New parameter-free simplified swarm optimization for artificial neural network training and its application in the prediction of time series. *IEEE Transactions on Neural Networks and Learning Systems,* 24, 661-665.

Yin, Z. Y., Jin, Y. F., Shen, J. S. and Hicher, P. Y. 2018. Optimization techniques for identifying soil parameters in geotechnical engineering: Comparative study and enhancement. *International Journal for Numerical and Analytical Methods in Geomechanics,* 42, 70-94.

Yoon, H., Yang, K. and Shahabi, C. (2005). Feature subset selection and feature ranking for multivariate time series. *IEEE transactions on knowledge and data engineering*, 17, 1186-1198.

Yoshihara, I., Aoyama, T. and Yasunaga, M. (2000). GP-based modeling method for time series prediction with parameter optimization and node alternation. Proceedings of the IEEE Congress on Evolutionary Computation, 1475-1481.

Yu, C. H., Ward, M. W., Morabito, M. and Ding, W. (2011). Crime forecasting using data mining techniques. IEEE 11th International Conference on Data Mining Workshops (ICDMW), 779-786.

Yu, Z., Yousaf, K., Ahmad, M., Methods, M. Y., Gao, Q. and Chen, K. (2020). Efficient Pyrolysis of Ginkgo Biloba Leaf Residue and Pharmaceutical Sludge (mixture) with High production of Clean Energy: Process

Optimization by Particle Swarm Optimization and Gradient Boosting Decision Tree Algorithm. *Bioresource Technology*, 123020.

Zainal, N. A. and Mustaffa, Z. (2015). A literature review on gold price predictive techniques. IEEE 4th International Conference on Software Engineering and Computer Systems (ICSECS), 39-44.

Zatz, M. S. and Smith, H. (2012). Immigration, crime, and victimization: Rhetoric and reality. *Annual Review of Law and Social Science,* 8, 141-159.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing,* 50, 159-175.

Zhang, G. P. (2007). A neural network ensemble method with jittered training data for time series forecasting. *Information Sciences,* 177, 5329-5346.

Zhang, X., Chen, X. and He, Z. (2010). An ACO-based algorithm for parameter optimization of support vector machines. *Expert Systems with Applications,* 37, 6618-6628.

Zhang, X., Lu, X., Shi, Q., Xu, X.-Q., Hon-Chiu, E. L., Harris, L. N., Iglehart, J. D., Miron, A., Liu, J. S. and Wong, W. H. (2006). Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data. *BMC bioinformatics,* 7, 197.

Zhang, X., Nguyen, H., Bui, X.-N., Tran, Q.-H., Nguyen, D.-A., Bui, D. T. and Moayedi, H. (2019). Novel soft computing model for predicting blast-induced ground vibration in open-pit mines based on particle swarm optimization and XGBoost. *Natural Resources Research*, 1-11.

Zhang, Y. and Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies,* 58, 308-324.

Zhao, J., Wang, W., Pedrycz, W. and Tian, X. (2012). Online parameter optimization-based prediction for converter gas system by parallel strategies. *IEEE Transactions on Control Systems Technology*, 20, 835-845.

Zhao, X. and Tang, J. (2018). Crime in urban areas: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 20, 1-12.

Zhou, H., Chen, J., Dong, G., Wang, H. and Yuan, H. (2016). Bearing fault recognition method based on neighbourhood component analysis and coupled hidden Markov model. *Mechanical Systems and Signal Processing,* 66, 568-581.