Crime Prediction Using Big Data Analysis

Hussam Hashim Hussein[1], Ahmed Talib Abdulameer[2]
{dac0019@mtu.edu.iq[1], Dr.ahmed.talib@mtu.edu.iq[2]}

Information Technology dept., Technical College of Management-Baghdad, Middle Technical University, Baghdad, Iraq[1,2]

Abstract. Crime is one of the most dangerous phenomena around the world. It is impossible to find a country or society free of crime, where criminals use all modern techniques and advanced technological methods in committing crimes, for example, cybercrime. In recent years, the generation of raw big crime data has been increased, therefore big data analysis become mandatory. Deep Learning is employed by using (LSTM) which is one of the Recurrent Neural Network (RNN) types. 'Criminal data in Chicago' is used as a dataset. An intelligent model is proposed for type, time, and place crime prediction to address the problem of high crime rates and reaching to help (police agencies - law enforcement institutions) to predict crime, reduce the spread of crime, as well as to make the best usage of security sources of (police elements - resources).

Keywords: Crime, Big Data Analysis, Deep Learning, Machine Learning, Recurrent neural network (RNN)/Long Short-Term Memory (LSTM).

1 Introduction

Crime is one of the most dangerous phenomena for any country. It is impossible to find a country or society free of Crime because Crime goes beyond the limits of harm or threatening the safety of others but rather affects the government and society in economic, cultural, and other terms. Crime prediction is an attempt to identify and limit future crime; crime prediction uses past data. After analysing the data, we notice that previous data related to criminal activities help expect crime Also, crime is one of the main problems in any society or country in this era. Even criminals are using all modern technologies and advanced technological methods to commit crimes, for example, cybercrime. Therefore, crime has become the focus of a challenge for both the police department and law enforcement agencies to reduce the spread of crime and use modern techniques to predict and reduce the spread of crime.

Many police departments use artificial intelligence algorithms and use big data tools to help them to predict where crimes are occurring. For example, police departments in places like Seattle, Los Angeles, and Atlanta have experimented with predictive police programs which try to identify the geographic area where crime is likely to occur in the future. At the same time, the Chicago Police Station used an algorithm based on a heat list that tries to identify people as criminals for violent crimes or repeated abuse.

To reduce the spread of crime, new technologies can deal with the vast amount of data, where traditional analysis techniques cannot analyse the data; therefore, artificial techniques and algorithms use big data to analyse crime data.

2 Related Works

Big data analytics (BDA) has been extensively applied and studied in data science. Londhe and Rao [1] presented various software frameworks available for BDA and discussed some widely used data mining algorithms. Grady et al. [2] illustrated the implications of an agile process for the cleansing, transformation, and analytics of data in BDA. Liu et al. [3] exemplified the utilization of BDA in the tourism domain in terms of the datasets, data capture techniques, analytical tools, and analysis results to provide insights for Destination Management. Fisher et al. [4] explained the conception of big data in BDA, its analytics, and the associated challenges when interacting among them. In [5] is to explores a methodology for reliably predicting the location, time, or likelihood of future residential burglaries. First, suitable data structure designs stored spatial/temporal information and aggregated counts of crime-related events categorized by the city's police department. Second, an ensemble of data mining classification algorithms is applied to perform residential burglary forecasting. In addition, it explores whether the crime rate will emerge or increase at specific locations (called' heating up'). Finally, in [6], a multivariate time series clustering technique based on dynamic time wrapping (DTW) and parametric Minkowski model is proposed to discover similar trends in crime data and subsequently use this information for future crime trends prediction. The algorithm has been tested on real-world datasets (provided by the Indian National Crime Records Bureau), performing separate analyses for various crimes such as murder, kidnapping, and so on.

3 Big Data

Data is everywhere and affects almost all aspects of life. Recently, many private institutions and national governments such as the United States have paid great attention to Big Data where these institutions and governments have become interested in the high potential of Big Data. For example, in July 2012, the United Nations released the Additional Data for Development Report, explicitly examining how governments have used additional data [7].
Many government agencies have announced master plans to accelerate Big Data research and its applications, such as research initiatives and projects. For example, the US National Science Foundation (NSF), as part of the Big Data Initiative launched by the Obama administration, announced BIG DATA in March 2012. The initiative resulted in several winning projects to investigate the foundations of Big Data Management (led by Washington University), analytical approaches for computing Big Data, and machine learning technologies. These projects seek to develop methods, algorithms, frameworks, infrastructure, and research that allows for analyzing this Big Data and explore data and extract information or knowledge to take advantage of them in the future, as well as to reduce this large amount of data to a manageable range and its interpretation by the human[8]. As for private, non-governmental organizations, over the past few years, almost all major companies, including Oracle, IBM, Microsoft, Google, Amazon, Facebook [8]. However, people who work in the Big Data field still have different opinions about defining Big Data:

1. Big Data: originally the volume of data that cannot be processed (efficiently) by traditional methods and database methods and tools [9].
2. Big Data: it refers to data sets that are both large and high in variety and speed, making them difficult to handle by using traditional tools [10].

3. Big Data: it is the amount of data that exceeds the ability of technology to store, manage and process efficiently [11].

4 Recurrent Neural Network

It is a type of deep learning network used for the first time in the 1980s and 1980s. It is used to solve deep learning network problems. It contains recurrent loops within the network, which gives it memory effect operations by using a state-based structure that uses previous information to estimate the next value. This network is beneficial in recognizing the sequence dependence of specific data (predicting the next value from the chain). It retains some essential features of the sequential data and forecasts future data.

4.1 Long Short-Term Memory(LSTM)

Long-term memory networks, commonly called "LSTM", are a particular type of repetitive neural network, introduced by Hochreiter & Schmidhuber 1997, which works or can learn long-term by using the Cell State where it maintains the state of the cell and passes it to the next cell to link previous data to current data to ensure that data is not lost .on the other hand, the LSTM is explicitly designed to solve the problem of addressing the problem of gradient fading when updating the weights with the best expectations which are reached of the model in the recurring network model . LSTM (Long Range Memory) is an advanced RNN that aims to solve the fading gradient problem and achieve excellent results on longer sequences. The memory determines the most information to pass in each step of the time. Each cell in LSTM contains:

1. RNN Cell
2. Cell State= long term memory (Ci)
3. Forget gate (ft)
4. Input gate (it)
5. Output gate (Ot)

5 Problem Definition and Goal

Crime has always been among the phenomena that have attracted the attention of scholars due to its connection with the social structure of society. The research problem can be summarized in the high crime rates in addition to the existence of huge raw crime data for the Chicago-US research community for the period (2001-2017) [12]. Our goal is to build an intelligent model developed to help the competent authorities (police agencies - law enforcement institutions) for predicting the type, time, and place of crime.

5.1 The Proposed Developed Intelligent Model

To understand how the proposed developed intelligent crime prediction model works using big data analysis, it is necessary to start with the proposed model architecture, as in Fig 1, to understand the model clearly. The proposed intelligent model design to predict different types of predictions, such as predicting the rate of a specific crime type, as well as predicting a crime rate in a particular district in Chicago city, additionally predicting the rate of crime at times.
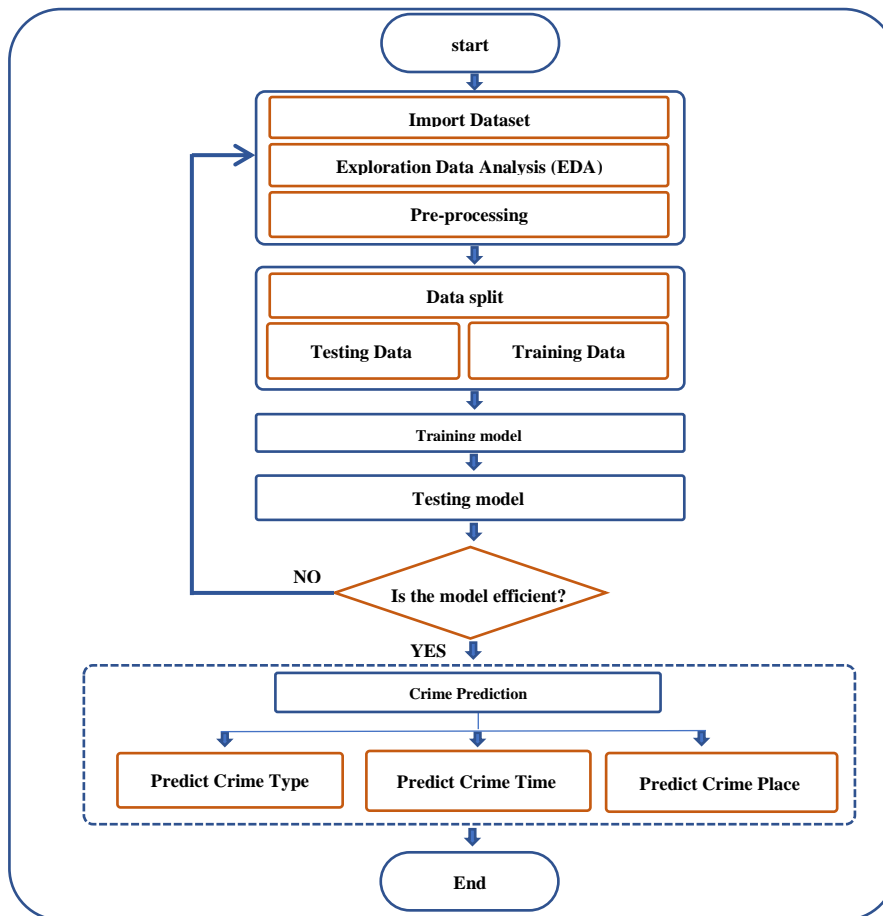
Fig 1. The architecture of the proposed developed intelligent model.

5.2 Data Description

The crime dataset used in this model was gathered from a publicly available dataset of the United States as mentioned in [12]. Each crime describes by 23 features (i.e., type of crime, location district, date, x-longitude, y-latitude, and others). The dataset contains (7,951,232) records of crime, the snapshot of the original dataset is depicted in Fig 2.

	Unnamed: 0	ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Description	Arrest	...	Ward	Community Area	FBI Code	X Coordinate	
0	879	4786321	HM399414	01/01/2004 12:01:00 AM	082XX S COLES AVE	0840	THEFT	FINANCIAL ID THEFT: OVER $300	RESIDENCE	False	...	7.0	46.0	06	NaN	
1	2544	4676906	HM278933	03/01/2003 12:00:00 AM	004XX W 42ND PL	2825	OTHER OFFENSE	HARASSMENT BY TELEPHONE	RESIDENCE	False	...	11.0	61.0	26	1173974.0	1.8
2	2919	4789749	HM402220	06/20/2004 11:00:00 AM	025XX N KIMBALL AVE	1752	OFFENSE INVOLVING CHILDREN	AGG CRIM SEX ABUSE FAM MEMBER	RESIDENCE	False	...	35.0	22.0	20	NaN	
3	2927	4789765	HM402058	12/30/2004 08:00:00 PM	045XX W MONTANA ST	0840	THEFT	FINANCIAL ID THEFT: OVER $300	OTHER	False	...	31.0	20.0	06	NaN	
4	3302	4677901	HM275615	05/01/2003 01:00:00 AM	111XX S NORMAL AVE	0841	THEFT	FINANCIAL ID THEFT:$300 &UNDER	RESIDENCE	False	...	34.0	49.0	06	1174948.0	1.8

5 rows × 23 columns

Fig 2. Snapshot of Chicago dataset.

5.3 Data Pre-Processing

Data pre-processing is a technique that involves converting raw data into an understandable format. Unfortunately, real-world data is often incomplete or inconsistent, and it will likely contain many errors. Therefore, the pre-processing of crime data is an essential step before starting to train the model on this data; a sequence of pre-processing steps is performed for data as follows:

1. Features Selection: select an important feature from a dataset.
2. Missing value: checking and handling missing values.
3. Data scaling: normalized data.
4. Reshape Input Data: to be adequate with LSTM network.

5.4 Exploratory Data Analysis (EDA)

Through this exploration, the researcher can know the valuable data that lead the model to predict the crime (the goal of the proposed model). We found the types of recorded crimes have (35) different types of crimes as in Fig 3, for example, the crime of theft, extortion, kidnapping, theft of private property, and others. In addition, it finds that the most common type of crime that occurred is a theft crime.
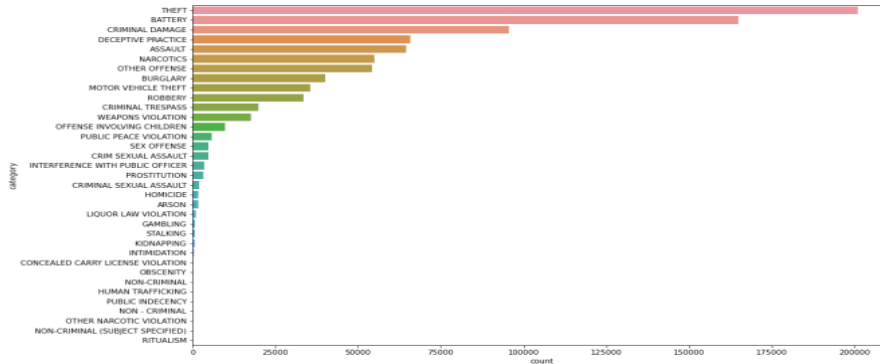
Fig 3. The types of crimes in the dataset.

Additionally, to know which days are the most dangerous (there are more crimes in it than on other days), we found through the graphical chart in Fig 4 that the crimes occurred on all days, and this is a clear indication that there is a high rate of crimes on all days of the week and for most types of crimes.
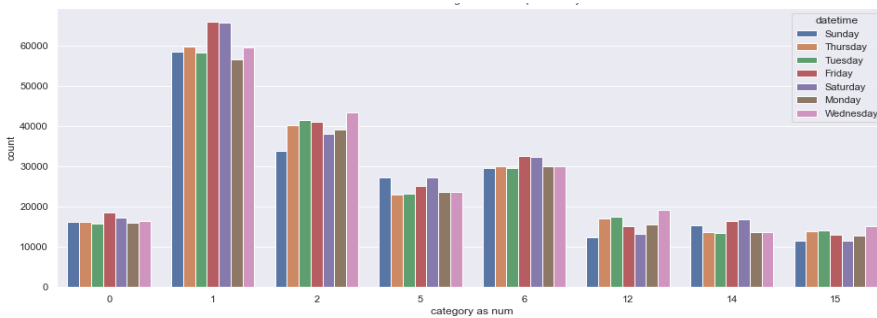

Fig 4. The number of crimes by days of the week.

During the exploration of the data, we found that the crime rate for each province is different from one to the other. The crimes that occur in District No. (11) are more incidence than in the rest of the provinces. District No. (20) recorded the lowest crime rate compared with other provinces.
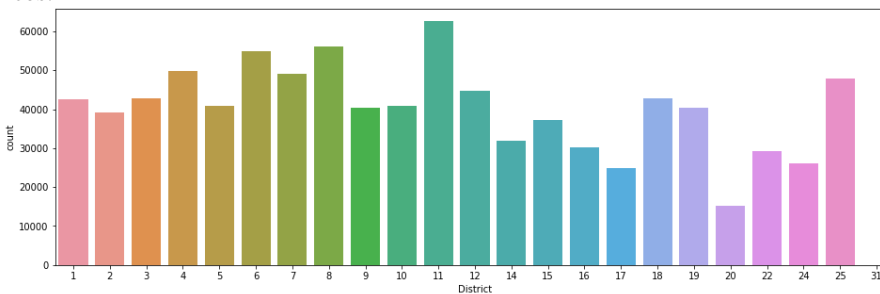

Fig 5. The number of crimes by District No.

5.4 Experiments Results

This section deals with the results of the proposed model. Experiments' results and trained model settings are shown in Table 1 and Table 2 respectively.

Table 1. Root Mean Square Error (RMSE) results for the Proposed Model

Prediction Model	RMSE
Crime Type prediction model	51.546612
Crime Time prediction model	11.25940
Crime Place prediction model	0.65185

Table 2. Hyper-parameters for the Proposed Model.

HYBER PARAMETER	VALUE
NODES OF LSTM 1st LAYER	50
NODES OF LSTM 2nd LAYER	50
NODES OF DENCE LAYER	1
LEARNING RATE	0.001
OPTIMIZER	Adam
ACTIVATION FUNCTION	Relu

Crime Type Prediction: The researcher tested the proposed model by (80%) of training data, from the original data. The original data are collected by month of each year from 2001 to 2017. The researcher reached RMSE (51.546612) to predict the type of crime.

Crime Time Prediction: The researcher tested the proposed model by (80%) of training data, from the original data. To predict a specific time (10:00 PM) for 'theft' crime, the researcher reached RMSE (11.25940).

Crime Place Prediction: The researcher tested the proposed model by (80%) of training data, from the original data. to select a specific place (a district in Chicago city) to predict 'ARSON ' crime, the researcher reached RMSE (0.65185).

Fig 6. Crime Type Prediction by the proposed model.

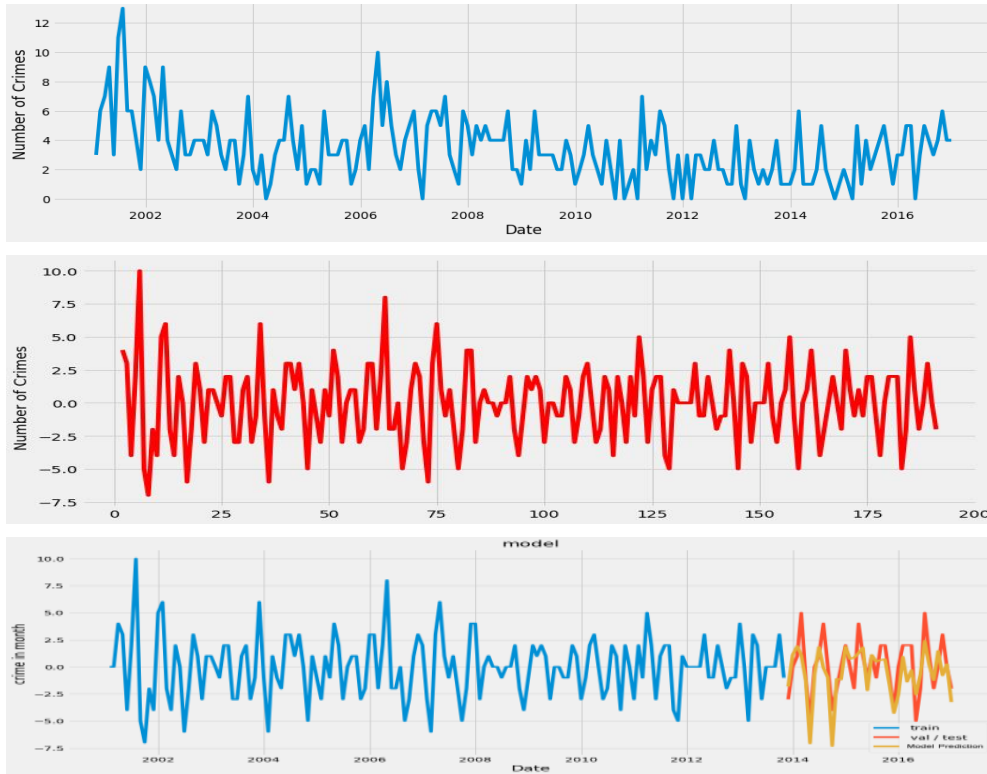**Fig 7.** Time Crime Prediction by the proposed model.

Fig 8. Crime Place Prediction by the proposed model.

From the results of the root mean square error (RMSE) and experiments' results, the researcher showed the results of different types of prediction, as well as it is considered that the model works on crime data in a way that can predict the crime efficiently.

6   Conclusion

Crime is one of the most dangerous phenomena for any country. To reduce and non-proliferation of crime, it requires new techniques that can deal with the vast amount of data, where the data cannot be analysed with traditional analysis techniques. Therefore, Deep Learning algorithms techniques were employed through using repeated neural network (RNN) and the employment of one of its types (LSTM). Our contribution is to build an intelligent crime prediction model, where the model predicts different types: crime type, time of the crime, the place of crime). In addition, many deep learning techniques have been employed such as dropout and early stopping for addressing the problem of high crime rates and helping the competent authorities (police agencies - law enforcement institutions) to predict crime, reduce the spread of crime as well as make the best use of security sources of (police elements - resources), and employ them properly.

References

[1] Londhe A, Rao PP, editors. Platforms for big data analytics: Trend towards hybrid era. 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS); 2017: IEEE.

[2] Grady NW, Payne JA, Parker H, editors. Agile big data analytics: Analytics Ops for data science. 2017 IEEE international conference on big data (big data); 2017: IEEE.

[3] Liu Y-Y, Tseng F-M, Tseng Y-H. Big Data analytics for forecasting tourism destination arrivals with the applied Vector Autoregression model. Technological Forecasting and Social Change. 2018;130:123-34.

[4] Fisher D, DeLine R, Czerwinski M, Drucker S. Interactions with big data analytics. interactions. 2012;19(3):50-9.

[5] Yu C-H, Ward MW, Morabito M, Ding W, editors. Crime forecasting using data mining techniques. 2011 IEEE 11th international conference on data mining workshops; 2011: IEEE.

[6] Chandra B, Gupta M, Gupta M, editors. A multivariate time series clustering approach for crime trends prediction. 2008 IEEE International Conference on Systems, Man and Cybernetics; 2008: IEEE.

[7] Janssen M, van der Voort H, Wahyudi A. Factors influencing big data decision-making quality. Journal of business research. 2017;70:338-45.

[8] Kaisler S, Armour F, Espinosa JA, Money W, editors. Big data: Issues and challenges moving forward. 2013 46th Hawaii international conference on system sciences; 2013: IEEE.

[9] Chen CP, Zhang C-Y. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information sciences. 2014;275:314-47.

[10] Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU. The rise of "big data" on cloud computing: Review and open research issues. Information systems. 2015;47:98-115.

[11] Williams ML, Burnap P, Sloan L. Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns. The British Journal of Criminology. 2017 Mar 1;57(2):320-40.

[12] Feng M, Zheng J, Ren J, Hussain A, Li X, Xi Y, Liu Q. Big data analytics and mining for effective visualization and trends forecasting of crime data. IEEE Access. 2019 Jul 22; 7:106111-23.