

CHAPTER – 2

LITERATURE SURVEY

2.1 INTRODUCTION

Historically solving crimes has been the right of the criminal justice and law enforcement specialists. With the increase in the use of the computerized systems to work and track crimes and trace criminals, computer data analysts have started lending their hands in helping the law enforcement officers and detectives to speed up the process of solving crimes. Criminology is the process that is used to identify crime and criminal characteristics. The criminals and the crime occurrence possibility can be assessed with the help of criminology techniques. Crimes are social nuisance and cost our society dearly in several ways. Any research that can help in solving crimes faster will pay for itself. About 10% of the criminals commit about 50% of the crimes. People who study criminology will be able to identify the criminals based on the trace, characteristics, pattern and methods of crime which can be collected from the crime scene. Several Research papers working on similar ideas and areas were surveyed and the key points were noted in the implementation of our project.

2.2 LITERATURE SURVEY WORKS

Literature survey in a project report is that section which shows the various analyses and research made in the field related to our project work. The purpose of a literature survey is to gain an understanding of the existing research and relevant to our work and particular to the topic crime prediction and detection using historic data or area of study. From these literature works below mentions which is related to our work helps to build knowledge in our field of working. These are some of the literature survey works, which includes crime, crime analysis, methods, predicting the crime, machine learning, approach of crime detection and etc.

2.2.1 AN INTELLIGENT DOCUMENT TO DETECT CRIME PATTERNS

A wide number of reports and news on crimes, increasingly conducted almost every day, have resulted in making detection of such crimes more difficult if not complex. Therefore, the need for detecting and identifying such crimes emerges as a necessary way of detecting and identifying such crime patterns on the news. In addition, the following approach assists the law enforcement officers and detectives to enhance performance and speed up the process of solving crimes. Due to the rapid technological advances including the increasing application of the computerized systems to trace and track crimes, computer data analysts have taken practical steps in assisting the law enforcement officers and detectives to enhance the process of solving crimes and increase its performance.[1] It aims to enhance the k means algorithm for Document Clustering as well as the extracting of information which group topics/events of crimes can outperform the original Document Clustering and other Document Clustering based on two criteria of time and performance . Tracking the social crimes or events according to their time line is becoming a tedious task. The challenges and complexities in organizing the news of crime stories are generated from a huge dimensionality of crimes data, which usually refers to the highly diverse embedded modalities such as criminal data and weapon data.

2.2.2 A REVISED FREQUENT PATTERN MODEL FOR CRIME SITUATION RECOGNITION

Identifying related offences in a criminal investigation is an important goal for crime analysts. This can deliver evidence that can assist in apprehension of suspects and better attribution of past crimes. The use of pattern based approaches has the potential to assist crime experts in discovering new patterns of criminal activity. Hence, research in this area continues. This work revisits

frequent pattern growth models for crime pattern mining. Frequent pattern (FP) based approaches, such as the FP-Growth model, have been identified to be more effective than techniques proposed in the past, such as A priori. Therefore, this research proposes a descriptive statistical approach, based on a quartile (floor-ceil) function, for the minimum support threshold (MST) choice selection, which is a major decision step in the pruning phase of the Traditional FP-Growth (TFPG) model. Our revised frequent pattern growth (RFPG) model further proposes a Pattern-pattern (Pp) paradigm to identify tupelos of subtle crime pattern(s) sequences or recurring trends in criminal activity. We present empirical results in order to guide intended audience about future decisions or research regarding this model. Results indicate that RFPG is more promising than[2] TFPG and will always ensure the utilization of a reasonable percentage of the crime dataset, in order to produce more reliable and sufficiently informative patterns or trends. To produce more reliable and sufficiently informative patterns or trends. It is flexible and allows the user to choose from the available options for MST choice. Pattern-pattern paradigm is capable of identifying subtle frequent crime pattern trends in the generated frequent crime set. The challenge of crime is still being tackled by security agencies, while crime analysts also persist in discovering trends in crime data. Data mining approaches such as frequent crime pattern mining approaches (i.e. FP-Growth) have the potential of detecting actionable information that stakeholders or crime experts might inadvertently miss, since it does not lie within their expectations.

2.2.3 Crime Prediction & Monitoring Framework Based on Spatial Analysis

Crimes are treacherous and common social problem faced worldwide. Crimes affect the quality of life, economic growth, and reputation of a nation. There has been an enormous increase in crime rate in the last few years. In order to reduce the crime rate, the law enforcements need to take the preventive

measures. With the aim of securing the society from crimes, there is a need for advanced systems and new approaches for improving the crime analytics for protecting their communities. Accurate real-time crime predictions help to reduce the crime rate but remains challenging problem for the scientific community as crime occurrences depend on many complex factors. In this work, various visualizing techniques and machine learning algorithms are adopted for predicting the crime distribution over an area. In the first step, the raw datasets were processed and visualized based on the need. Afterwards, machine learning algorithms were used to extract the knowledge out of these large datasets and discover the hidden relationships among the data which is further used to report and discover the crime patterns that is valuable for crime analysts to analyze these crime networks by the means of various interactive visualizations for crime prediction and hence is supportive in prevention of crimes. Supportive in prevention of crimes. The project helps the crime analysts to analyze these crime networks by means of various interactive visualizations. The interactive and visual feature applications will be helpful in reporting and discovering the crime patterns.[3] Accurate real-time crime predictions help to reduce the crime rate but remains challenging problem for the scientific community as crime occurrences depend on many complex factors.

2.2.4 Crime rate prediction in the urban environment using social factors

The aim of this study is to compare different approaches to the problem of forecasting the number of crimes in different areas of the city. During this research we studied three types of predictive models: linear regression, logistic regression and gradient boosting.[4] The predictive factors used in these models have been selected using the feature selection techniques. This approach allowed us to increase the accuracy of predictions and to avoid the model's over fitting. The obtained models were tested on criminal data of the city of Saint-Petersburg. We compared the results of model predictions and determined that

gradient boosting is the most appropriate method for the problem of crime rate prediction in certain urban areas. To compare different approaches to the problem of forecasting the number of crimes in different areas of the city. This approach allowed us to increase the accuracy of predictions and to avoid the model's over fitting. Urbanization creates a lot of social problems and these problems inherent in all cities of the world is crime. The analysis of criminal activity and prediction of number of crimes remains one of the most interesting problem for researchers.

2.2.5 SURVEY ON CRIME ANALYSIS AND PREDICTION USING DATA MINING TECHNIQUES

In this work the procedure which includes evaluating and examining large pre-existing databases in order to generate new information which may be essential to the organization. The extraction of new information is predicted using the existing datasets. Many approaches for analysis and prediction in data mining had been performed. But, many few efforts has made in the criminology field. Many few have taken efforts for comparing the information all these approaches produce.[5] The police stations and other similar criminal justice agencies hold many large databases of information which can be used to predict or analyze the criminal movements and criminal activity involvement in the society. The criminals can also be predicted based on the crime data. The main aim of this work is to perform a survey on the supervised learning and unsupervised learning techniques that has been applied towards criminal identification. This work presents the survey on the Crime analysis and crime prediction using several Machine Learning techniques.

The main aim of this work is to perform a survey on the supervised learning and unsupervised learning techniques that has been applied towards criminal identification. The motivation for proceeding with this survey work is to aid a helping hand to the young researchers who are performing their research in

criminal analysis and crime prediction areas. Solving crimes is a difficult and time consuming task that requires human intelligence. Crime analysis tasks can be a tedious process for the police or the investigation team to work with.

2.2.6 Knowledge Representation Model for Crime Analysis

The knowledge representation model is a particular way in representing knowledge by using the knowledge and reasoning mechanism. Ontology is a kind of knowledge representation model that represents knowledge as a set of concepts within a domain and the relationship between these concepts. It is important to solve the problem of large amount of data in the crime investigation domain which is not well defined in a proper relation. The problems exist currently are how to develop an ontology model to represent the crime investigation information and how to make good use of the information represented by the model. These problems could be solved by developing an ontological-based case matching model, named Crime Analysis as a study prototype. In this study, an ontology model is developed using the selected semantic modeling tool, named Top Braid Composer Standard Edition in order to represent the crime information with the well-defined classes and relationships. This would help to save the investigation officer's effort in aiming and targeting the possible suspect within the shortest time interval. Save the investigation officer's effort in aiming and targeting the possible suspect within the shortest time interval.[6] Crime Analysis can be used to match the newly reported cases with the similar past case records and display the related case summary report to assist the IO in the investigation process.

The problems exist currently are how to develop an ontology model to represent the crime investigation information and how to make good use of the information represented by the model. The problem of large amount of data in the crime investigation domain which is not well defined in a proper relation. The current system of a recording nature only keeps the criminal records, card

files and databases. It does not construct the information meaningful and therefore it is insufficient in providing the information for crime investigation.

2.2.7 From data to disruption CRIME INVESTIGATION

Organized crime groups and law enforcement agencies are caught in complex system similar to a continuous game of cat-and-mouse, in which the latter frequently remains two or more steps behind. Law enforcement[7] agencies are therefore seeking for more proactive strategies in targeting these criminal network structures more effectively. This starts with a better understanding of the way they operate and adapt over time. A key element to developing this understanding remained largely unexploited: big data and big data analytics. This provides novel insight into how criminal co operations on a micro- and meson level are embedded in small-world criminal macro-networks and how this fosters its resilience against disruption. This work discusses the opportunities and the limitations of this data-driven approach and its implications for both law enforcement practice and scientific research. It provides novel insight into how criminal co operations on a micro and meson level are embedded in small-world criminal macro-networks and how this fosters its resilience against disruption. The aim of this work is to provide a short overview of the current developments in the field of big data analysis in the study of organized crime. Law enforcement agencies are way behind in solving complex crimes. Organized crime groups impose a continuous threat to global society, are undetected by government.

2.2.8 Crime Analysis Through Machine Learning

This work investigates machine-learning-based crime prediction. In this work, Vancouver crime data for the last 15 years is analyzed using two different data-processing approaches. Machine-Learning predictive models, K-nearest neighbor and boosted decision tree, are implemented and a crime prediction accuracy between 39% to 44% is obtained when predicting crime in Vancouver.

The prediction accuracy can be improved by tuning both the algorithm and the data for specific applications.[8] It provides a preliminary framework for further analyses. Crime affects life quality and economic growth. Increased crime rate.

2.2.9 Examining Deep Learning Architectures for Crime Classification and Prediction

In this work, a detailed study on crime classification and prediction using deep learning architectures is presented. We examine the effectiveness of deep learning algorithms on this domain and provide recommendations for designing and training deep learning systems for predicting crime areas, using open data from police reports. Having as training data time series of crime types per location, a comparative study of 10 state-of-the-art methods against 3 different deep learning configurations is conducted. In our experiments with five publicly available datasets, we demonstrate that the deep learning-based methods consistently outperform the existing best-performing methods. Moreover, we evaluate the effectiveness of different parameters in the deep learning architectures and give insights for configuring them in order to achieve improved performance in crime classification and finally crime prediction. Provide recommendations for designing and training[9] deep learning systems for predicting crime areas, using open data from police reports. To achieve improved performance in crime classification and finally crime prediction. The most challenging issues of police departments is to have accurate crime forecasts to dynamically deploy patrols and other resources so as to improve deterring of crime occurrence and police response times.

2.2.10 Crime data mining: A general framework

A major challenge facing all law enforcement and intelligence gathering organizations is accurately and efficiently analyzing the growing volume of crime data. Data mining is a powerful tool that enables criminal investigators who may lack extensive training as data analysts to explore large database

quickly and efficiently computer can also process thousands of instructions in seconds saving processing time.[10] Complex conspiracies are difficult to unravel because the information can be geographically diffuse and span long period of time. Growing volumes of crime data. By increasing efficiency and reduced errors. Crime data mining techniques can facilitate police work enable investigators to allocate then time to other values tasks.

2.2.11 Crime detection using data mining

As crime rates spirally each day, new challenges are faced by law enforcement agencies. Looking out for any signs of criminal activity will cause more burden on their resource. Therefore they should be able to predict such increase or decrease or trends in crime. The system work is to extract data from crime record on which data mining is performed. The final end product could be system where some future predictions would made by training crime sets output could be visualized.[11] The work of this project is to help ascertain the incidence and pattern of crime occurrence in future which include forecasting and predicting the occurrence of it. It can help law enforcement officials gauge the crime pattern which are prevalent in a particular area or region or in a interval of time. This would save time, money and effort. Criminals becoming technologically gained so law enforcement agencies have to keep up with them. Traditional policies methodologies haven't been successful in reducing the crime rates.

2.2.12 Crime Analysis and Prediction Using Data Mining

Crime analysis and prevention is a systematic approach for identifying and analyzing patterns and trends in crime. Our system can predict regions which have high probability for crime occurrence and can visualize crime prone areas. With the increasing advent of computerized systems, crime data analysts can help the Law enforcement officers to speed up the process of solving crimes. Using the concept of data mining we can extract previously

unknown, useful information from an unstructured data. Here we have an approach between computer science and criminal justice to develop a data mining procedure that can help solve crimes faster. Instead of focusing on causes of crime occurrence like criminal background of offender, political enmity etc we are focusing mainly on crime factors of each day. Increase in crime[12] information that has to be stored and analyzed. Analysis of data is difficult since data is incomplete and inconsistent. Limitations in getting crime record data from law enforcement department. Speedup the process of solving crimes. Instead of focusing on the causes of crime occurrence it focuses on crime factors. The application helps in reducing crime rate to a certain extent by providing security in crime sensitive area.

2.2.13 Learning to Detect Patterns of Crime

Our goal is to automatically detect patterns of crime. Among a large set of crimes that happen every year in a major city, it is challenging, time-consuming, and labor-intensive for crime analysts to determine which ones may have been committed by the same individual(s). If automated, data-driven tools for crime pattern detection are made available to assist analysts, these tools could help police to better understand patterns of crime, leading to more precise attribution of past crimes, and the apprehension of suspects. To do this, we propose a pattern detection algorithm called Series Finder,[13] that grows a pattern of discovered crimes from within a database, starting from a “seed” of a few crimes. Series Finder incorporates both the common characteristics of all patterns and the unique aspects of each specific pattern, and has had promising results on a decade’s worth of crime pattern data collected by the Crime Analysis Unit of the Cambridge Police Department.

It is challenging, time-consuming, and labor-intensive for crime analysts to determine which ones may have been committed by the same individual(s) from a huge dataset.[13] The goal of crime data mining is to understand patterns in

criminal behavior in order to predict crime, anticipate criminal activity and prevent it. Series Finder is designed to detect patterns of crime committed by the same individual(s).

2.2.14 An Enhanced Algorithm to Predict a Future Crime using Data Mining

Concern about national security has increased after the 26/11 Mumbai attack. In this work we look at the use of missing value and clustering algorithm for a data mining approach to help predict the crimes patterns and fast up the process of solving crime. We will concentrate on MV algorithm and A priori algorithm with some enhancements to aid in the process of filling the missing value and identification of crime patterns.[14] We applied these techniques to real crime data. We also use semi supervised learning technique in this work for knowledge discovery from the crime records and to help increase the predictive accuracy.

Increase in the size of crime information that has to be stored and analyzed. Problem of identifying techniques that can accurately and efficiently analyze this growing volumes of crime data Different methods and structures used for recording crime data. The data available is inconsistent and are incomplete thus making the task of formal analysis a far more difficult. Investigation of the crime takes longer duration due to complexity of issues All the above challenges. [14] The research work focus on providing solutions that can enhance the process of crime analysis for identifying and reducing crime in India. The main aim of this research work consist of developing analytical data mining methods that can systematically address the complex problem related to various form of crime. Detecting crime patterns and perform crime analysis. Provide information to formulate strategies for crime prevention and reduction. Identify and analyze common crime patterns to reduce further occurrences of similar incidence.

2.2.15 Crime Prediction and Analysis Using Machine Learning

Crime is one of the biggest and dominating problem in our society and its prevention is an important task. Daily there are huge numbers of crimes committed frequently. This require keeping track of all the crimes and maintaining a database for same which may be used for future reference. The current problem faced are maintaining of proper dataset of crime and analyzing this data to help in predicting and solving crimes in future. The objective of this project is to analyze dataset which consist of numerous crimes and predicting the type of crime which may happen in future depending upon various conditions. In this project, we will be using the technique of machine learning and data science for crime prediction of Chicago crime data set. [15] The crime data is extracted from the official portal of Chicago police. It consists of crime information like location description, type of crime, date, time, latitude, longitude. Before training of the model data pre processing will be done following this feature selection and scaling will be done so that accuracy obtain will be high. The K-Nearest Neighbor (KNN) classification and various other algorithms will be tested for crime prediction and one with better accuracy will be used for training.

Visualization of dataset will be done in terms of graphical representation of many cases for example at which time the criminal rates are high or at which month the criminal activities are high. The soul purpose of this project is to give a jest idea of how machine learning can be used by the law enforcement agencies to detect, predict and solve crimes at a much faster rate and thus reduces the crime rate. It not restricted to Chicago, this can be used in other states or countries depending upon the availability of the dataset. The aim of this project is to make crime prediction using the features present in the dataset. The objective would be to train a model for prediction. The training would be done using the training data set which will be validated using the test dataset.

Crime prediction and criminal identification are the major problems to the police department as there are tremendous amount of crime data that exist. Maintaining of proper dataset of crime and analyzing this data to help in predicting and solving crimes in future.

2.2.16 AN APPROACH TO CRIME DATA ANALYSIS: A SYSTEMATIC REVIEW

In the current era, number of crimes occurs in the society and this criminal rate increase day by day. There is tremendous growth of criminal data. Crime has negatively influenced the societies. Crime control is essential for the welfare, stability and development of society. Law enforcement agencies are seeking for the system to target crime structure efficiently. The intelligent crime data analysis provides the best understanding of the dynamics of unlawful activities, discovering patterns of criminal behavior that will be useful to understand where, when and why crimes can occur.[16] There is a need for the advancements in the data storage collection, analysis and algorithm that can handle data and yield high accuracy. This work demonstrates the data mining technologies which are used in criminal investigation. The contribution of this work is to highlight the methodology used in crime data analytics. This work summarizes the challenges arising during the analysis process, which should be removed to get the desired result. Criminal data are available in different formats, thus tackling the variety of data formats from multiple data sources and transforming the data into a desirable form so that result can be obtained, is also a challenging task. There is another challenging aspect of storing the law enforcement data which are in large amount of size.

As the size of the data is huge, so there is need of storage devices which have large capability to store the data. [16] There are different analytical models are available, but appropriate analytical model is to be selected for data analyzing purpose and this is a challenging task. Some other challenging factors

are also exists, such as matching data mining technique and methodology, exploring proper integration methods to tackle complicated investigation problem. Complexity is another great challenge as the data reduction techniques are prior to the analyzing task. Review the use of data mining techniques and tools for identifying crime patterns. To highlight the methodology used in crime data analytics

2.2.17 CRIME PATTERN RECOGNITION AND PREDICTION USING OPTIMISED K-MEANS AND SVM

An act done by any person which is against the laws of the particular country or region is called crime. A person who does this is called a criminal. Almost without exception, people in America think that the crimes are increasing at all-time high levels. Such worries are linked to anxiety about drugs, permissive childrearing, hedonism, declining academic standards, and the collective inability to compete with the others. In Chicago the police maintain a criminal record which is accessible to the masses. [17] The basic idea of what things are called "crimes" is that they are thought to be things that might cause a problem for another person. In this work, we cluster the different parameters or attributes that are available with us in the dataset. After that we use SVM to do the prediction that were can the next crime happen. The output will be in the form of clusters mapped on a box plot and the prediction will also be on a box plot in the form of a graph. These results will help the police departments to plan their strategies according to the results obtained and prevent further crimes happening. The clustering and prediction approaches of data mining play vital role for the crime analysis which indeed helps to approach criminology in a different manner. Optimized k-means algorithm is used so that the running time can be reduced for the useful clusters to be generated and by using SVM for predicting crime pattern and drawing useful information from that. Criminals have maximum utilization of the novel practices and the modernized

innovation. For humans to process all the available data into a systematic manner needs computing powers.

2.3 RELATED WORKS

Saving the time and the effort of investigation officers by aiming and targeting the possible suspect within stipulated interval of time span. Crime Analysis can be used to match the newly reported cases with the similar past case records and display the related case summary report to assist the IO in the investigation process. The problems exist currently are how to develop and generate the model to represent the crime investigation information and how to make good use of the available information represented by the model. [1] The problem is huge amount of data and record in the crime investigation domain which is not well defined in a proper relation. The system only keeps the criminal records, card files and database for the development of the criminal records. A huge number of crime news and reports are increasing very day, have resulted in making detection of such crimes more difficult if not complex. Therefore, the need for detecting and identifying such crimes emerges as a necessary way of detecting and identifying such crime patterns on the news based. [2]

Using clustering and mining techniques the reports are generated which also causes an error in the generated records, this work discusses two major sequential stages in extraction and clustering of the crime records. [3] Due to the rapid technological advances including the increasing application of the computerized systems to trace and track crimes, computer data analysts have taken practical steps in assisting the law enforcement officers and detectives to enhance the process of solving crimes and increase its performance. It aims to enhance the k-means algorithm for Document Clustering as well as the extracting of information which group topics/events of crimes can outperform the original Document Clustering and other Document Clustering based on two

criteria of time and performance. [4] To achieve the good performance and improvising the crime classification and finally coming out with the perfect crime prediction. The challenging issues faced by the police departments is to have accurate crime forecasts and the details of the criminals to dynamically deploy patrols and other resources so as to improve deterring of crime occurrence and police response times. [5] In this project a particular module includes the forecasting and predicting the crime to help out in the investigation process in which the pattern of crime occurrence in future will be found out. It can help law enforcement officials gauge the crime pattern which are prevalent in a particular area or region or in a planned period of time. [6] This would save time, money and effort. So law enforcement agencies have to keep up with them. Traditional policy methods are not successful in reducing the crime rates. Process of solving crimes to resolving the criminal rates patterns are increased. The main focus is on the factors of the crime and the criminals and the rest of background are not taken into the notification by the investigator for the analyzing and the prediction. [11] The application used will be helpful for the investigation process for the identification of the crime and the criminals involved. The process of analysis and the storage of the crime records and the criminal data information must be processed. There arises a complexity issues in the analysis of the data due to the inconsistent and incomplete of raw data. [5] Limitations in getting crime record data from law enforcement department. Using the data mining technologies the crime patterns are analyzed and processed. The crime data mining process involves the prediction and detection of the crime scene patterns. [8] Data mining technologies are empowered here for the crime to be analyzed for further verifications. Goal of the employed data mining techniques is to target the analyzing predictive process. Series finder is a method of finding the source individual. For the detecting purpose of the crime and the criminal this method is used. It identifies [9] the crime and the criminal pattern. [5] Tracking the social crimes or events according to their time line is

becoming a tedious task. The challenges and complexities in organizing the news of crime stories are generated from a huge dimensionality of crimes data, which usually refers to the highly diverse embedded modalities such as criminal data and weapon data. Identifying related offences in a criminal investigation is an important goal for crime analysts. [2] This is purely based on the pattern based reaches which includes the algorithms to find out the crimes. Hence, research in this area continues. [15] This work revisits frequency based tracking problem the algorithms applied are F P growth and A priori. [4] Therefore, this research proposes a descriptive statistical approach, based on a quartile (floor-ceil) function the repeated frequency patterns are observed and they are clustered accordingly based on there priorities. We present empirical results in order to guide intended audience about future decisions or research regarding this model. Results indicate that RFPG is more promising than TFPG and will always ensure the utilization of a reasonable percentage of the crime dataset, in order to produce more reliable and sufficiently informative patterns or trends. [10] To produce more reliable and sufficiently informative patterns or trends. It is flexible and allows the user to choose from the available options for MST choice. Pattern-pattern paradigm is capable of identifying subtle frequent crime pattern trends in the generated frequent crime set. The challenge of crime is still being tackled by security agencies, while crime analysts also persist in discovering trends in crime data. [12] Data mining approaches such as frequent crime pattern mining approaches (i.e. FP-Growth) have the potential of detecting actionable information that stakeholders or crime experts might inadvertently miss, since it does not lie within their expectations. The aim of this study is to compare different approaches to the problem of forecasting the number of crimes in different areas of the city [4] During this research we studied three types of predictive models: linear regression, logistic regression and gradient boosting. The predictive factors used in these models have been selected using the feature selection techniques.

This approach allowed us to increase the accuracy of predictions and to avoid the model's over fitting. The obtained models were tested on criminal data of the city of Saint-Petersburg. [15] We compared the results of model predictions and determined that gradient boosting is the most appropriate method for the problem of crime rate prediction in certain urban areas. To compare different approaches to the problem of forecasting the number of crimes in different areas of the city. This approach allowed us to increase the accuracy of predictions and to avoid the model's over fitting. [9] Urbanization creates a lot of social problems and these problems inherent in all cities of the world is crime. The analysis of criminal activity and prediction of number of crimes remains one of the most interesting problem for researchers. [17] The knowledge representation model is a particular way in representing knowledge by using the knowledge and reasoning mechanism. Ontology is a kind of knowledge representation model that represents knowledge as a set of concepts within a domain and the relationship between these concepts. [6] It is important to solve the problem of large amount of data in the crime investigation domain which is not well defined in a proper relation. These problems could be solved by developing an ontological-based case matching model, named Crime Analysis as a study prototype. [7] In this study, an ontology model is developed using the selected semantic modeling tool, named Top Braid Composer Standard Edition in order to represent the crime information with the well-defined classes and relationships. This would help to save the investigation officer's effort in aiming and targeting the possible suspect within the shortest time interval. [3].

Save the investigation officer's effort in aiming and targeting the possible suspect within the shortest time interval [14] Crime Analysis can be used to match the newly reported cases with the similar past case records and display the related case summary report to assist the IO in the investigation process. [11]

The problems exist currently are how to develop an ontology model to represent the crime investigation information and how to make good use of the information represented by the model. The problem of large amount of data in the crime investigation domain which is not well defined in a proper relation. The current system of a recording nature only keeps the criminal records, card files and databases. It does not construct the information meaningful and therefore it is insufficient in providing the information for crime investigation. this work investigates machine learning based crime prediction. [16] In this work, Vancouver crime data for the last 15 years is analyzed using two different data-processing approaches.

Machine-Learning predictive models, K-mean and boosted decision tree, are implemented and a crime prediction accuracy between 39% to 44% is obtained when predicting crime in Vancouver. The prediction accuracy can be improved by tuning both the algorithm and the data for specific applications. It provides a preliminary framework for further analyses. [5] Crime affects life quality and economic growth. Increased crime rate. A detailed study on crime classification and prediction using deep learning architectures is presented. We examine the effectiveness of deep learning algorithms on this domain and provide recommendations for designing and training deep learning systems for predicting crime areas, using open data from police reports. Having as training data time series of crime types per location, a comparative study of 10 state-of-the-art methods against 3 different deep learning configurations is conducted. [8] In our experiments with five publicly available datasets, we demonstrate that the deep learning-based methods consistently outperform the existing best-performing methods. Moreover, we evaluate the effectiveness of different parameters in the deep learning architectures and give insights for configuring them in order to achieve improved performance in crime classification and finally [1] crime prediction. Provide recommendations for designing and

training deep learning systems for predicting crime areas, using open data from police reports. To achieve improved performance in crime classification and finally crime prediction. [15] The most challenging issues of police departments is to have accurate crime forecasts to dynamically deploy patrols and other resources so as to improve deterring of crime occurrence and police response times. As crime rates are increasing, new challenges are faced by the police department. Traditional methodologies are no more successful in reducing the crime rates. Criminals are becoming technologically savvy. [7] The work of this project is to identify the crime pattern and predict the future occurrence. This helps law enforcement officials to identify crime pattern in a particular area and thus save time, money and effort. The final output will be system that make future predictions that could be visualized.

Crime analysis and prevention is a systematic approach for solving crime. Due to increase in crime data, the analysis of data becomes difficult [8] The data collected is incomplete and inconsistent and thus slows the process of solving crime. Our system predicts the area that has high probability in crime occurrence and visualize them. The approach between computer science and criminal justice to develop a data mining procedure help solve crimes faster by focusing on crime factors. [1] The application help in reducing crime rate to a certain extent by providing security in crime sensitive area. [15] It is challenging, time consuming for crime analysts to predict the criminal. The automated data driven tools could help police to understand the crime pattern. To do this, we propose a pattern detection algorithm called Series Finder, that grows a pattern of discovered crimes from within a database. [2] The goal of project is to understand patterns in criminal in order to predict crime, anticipate criminal activity and prevent it from occurring in future. People are concerned about the increase in crime rate that is happening in their society. The problem arises in the size of information that has to be stored and analyzed. The

inconsistency and incompleteness of data makes it difficult for the police to solve the crimes. [6] The aim of this research work is to develop analytical data mining methods to solve complex problem. MV algorithm and A priori algorithm are used to aid the process of filling the missing value and identification of crime patterns. It formulates strategies for crime prevention and reduction. It provides a solution for crime analysis for identifying and reducing crime in India. Crime has become a serious problem in today's society due to increase in crime rates. [4] The current problem faced is maintenance of proper datasets in the database. Crime prediction and criminal identification are the major problems to the police department. The objective of this project is to analyze the dataset and predict which type of crime can happen in future. We use the technique of machine learning and data science for prediction purpose. [13] The K-Nearest Neighbor (KNN) classification will be used for crime prediction and one with better accuracy will be used as a training model. Visualization is done in the form of graphs. The motive of this project is to help the law enforcement agencies to detect, [9] predict and solve the crime in much lesser time. In current situation, crime rates occurring in the society are increasing day by day. Criminal data are in different format thus making it difficult for tracking.

Another challenge is storage of large amount of data by law enforcement agencies. [11] This work underlines the data mining technologies used in criminal investigation. There is a need for improvement in data storage, analysis and algorithm. The paper highlights the methodology used in crime analytics. It summarizes the problems arising during the analysis process, which should be removed to get the desired result. A Criminal is a person who does an act which is against a country or a region. [5] Crime rates in America are increasing at high levels. Crimes are related to drug, childrearing, decline in academic standards etc. Nowadays, Criminals are using modernized

technologies and hence it becomes difficult for the police to solve the crimes. In this paper, we cluster different parameter and use support vector machine for future prediction. The clustering and prediction technique helps to approach criminology in a different manner. The output will be a box plot in the form of graph. [3] The results will help the police to solve the crimes and prevent them from happening.

2.4 SUMMARY

Various crime solving techniques were briefed in the survey related works. Above list of works were surveyed for several factors with respect to our proposed idea. The factors include cost efficiency, processing efficiency, time efficiency, system processing and about the machine learning methodologies and techniques applied for the prediction process. The efficiency of the model depends on how well it is used to solve the crime and help the police to track the criminal at a faster rate. Out of all the works that were referred, the one with crime pattern recognition and prediction using optimized algorithm will prove efficient. However, this proposal does have some disadvantage. Few works will have, the systematic review of crime analysis will be useful but it cannot be implemented for a larger set of data. The rest of the works were all similar with their approach which might sure bring a change when modeled. The most common demerit that were considered in all the above mentioned works were handling of huge amount of crime data and the complexity involved in solving the case. It doesn't develop a model that can predict or forecast the future crime rate effectively. So our model will try to implement all these features in our work.