Crime Analysis & Prediction Using Machine Learning

Prof.Nikita Joshi[1],Sandip Chavan[2] Yashraj Dalvi[2], Ashutosh Dange[2], Manthan Dimble[2]

[1]*Student, Department of Computer Engineering, Savitribai Phule Pune University, Pune, India.*

Zeal College of Engineering, Pune, Maharashtra, India.

Abstract: - Understanding crime patterns is a crucial step in preventing and combating crime, as it enables the identification of prevalent crime types and their occurrence times in specific areas. Support Vector Machines (SVM), a machine learning algorithm, can be employed to classify different crime types and enhance prediction accuracy. The utilization of crime data from sources such as Kaggle can provide valuable insights into crime patterns, aiding law enforcement strategies and resource allocation. Analyzing crime data enables the identification of crime hotspots, facilitating targeted resource allocation. However, it is essential to acknowledge that crime data might not always reflect the true extent of crime in an area due to underreporting or lack of recording. Moreover, relying solely on data analysis may overlook the social and economic factors that contribute to crime. Therefore, while SVM and machine learning algorithms can assist in predicting and identifying crime patterns, adopting a multifaceted approach that incorporates community engagement, social services, and crime prevention initiatives is crucial to effectively address crime.

Key Words— *Crime, Analyze, Crime patterns,Kaggle, Estimate, Support Vector Machine, Classification, Regression and Accuracy*

I. INTRODUCTION

Crimes have a significant impact on society, leading governments to allocate substantial resources to law enforcement agencies in an attempt to prevent them. In recent times, law enforcement bodies have amassed large volumes of crime-related data that require processing to extract valuable information. The continuously increasing crime rates and evolving crime patterns make it challenging to explain the behaviors observed in these patterns. This paper aims to explore how social development can contribute to crime prevention. It provides a comprehensive review of relevant theories and research pertaining to crime prevention in society and implements different data analysis algorithms to investigate the connections between crime and its patterns. The data for this project is sourced from legitimate government sources and has been converted to the .csv format for ease of preprocessing. Various technologies are employed for mining crime patterns and conducting analysis.

II. LITERATURE SURVEY

Sr.no	Author,Title,Published date	Approaches	Advantages
1.	Crime type and occurence prediction using Machine learning algorithm. Kanimozhi N , Keerthana N V , Pavithra G S , Ranjitha G ,Yuvarani S (May 15,2021 IEEE)	pattern analysis by using crime data obtained from Kaggle open source which in turn used for the prediction of most recently occurring crimes. 2] Some machine learning algorithms such as Naïve Bayes is implied in this work	the crime patterns. 2] Crime forecasting can help to prevent recurring crimes in an area by identifying the patterns of crimes
2.	Crime Analysis and Prediction Using Fuzzy C-Means Algorithm B. Sivanagaleela,M.T	1] Crime analysis is methodological approach for identify the crime areas.	1] It is very useful for investigators to solve the crime rate.

	ech, S.Rajesh, Asst.Professor (-8/19/$31.00 ©2019 IEEE)	2] With the increasing of computer systems the crime data analysts can help to the crime investigators to analyze the crime.	2] Based on the fuzzy clustering technique the crime prone areas are identified it takes a less time.
3.	Crime Prediction Using K-Nearest Neighboring Algorithm BAkash Kumar, Aniket Verma, Gandhali Shinde. (-n May 06,2020IEEE)	1] In order to avoid the unfortunate, we will try to observe crime rates by the KNN prediction method. 2] It will predict, tentatively, the type of crime, when, where and at what time it may take place.	1] It focuses on having a crime prediction tool that can be helpful to law enforcement. 2] Along the way, many patterns of criminal activities in various areas which will be helpful for criminal investigatio n were known.
4.	QoI-Aware Unified Framework for Node Classification and SelfReconfiguratio n Within Heterogeneous Visual Sensor Networks ANAS AMJAD, (Student Member, IEEE), ALISON GRIFFITHS, AND MOHAMMAD PATWARY,	1] In order to avoid the unfortunate, we will try to observe crime rates by the KNN prediction method. 2] It will predict, tentatively, the type of crime, when, where and at what time it	1] Focuses on having a crime prediction tool that can be helpful to law enforcement. 2] Many patterns of criminal activities in various areas which will be helpful for
	(Senior Member, IEEE. (-IEEE ACCESS, n May 6, 2021,)	may take place.	criminal investigatio n were known.
5.	Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques WAJIHA SAFAT , SOHAIL ASGHAR , (Member, IEEE), AND SAIRA ANDLEEB GILLANI. (-IEEE ACCESS, VOL. XX, NO. X, DECEMBER 2016)	1] The accurate estimation of the crime rate, types and hot spots from past patterns creates many computationa l challenges and opportunities 2] Accurate crime prediction and future forecasting trends can assist to enhance metropolitan safety computationa lly.	1] . Crimes are serious threats to human society, safety, and sustainable developmen t and are thus meant to be controlled. 2] Along the way, many patterns of criminal activities in various areas which will be helpful for criminal investigatio n were known.

It is a kind of network architecture for deep learning algorithms and is specifically used for image recognition and tasks that involve the processing of pixel data.

CNN is designed to automatically and adaptively learn spatial hierarchies of features through back propagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers.

III. TECHNOLOGY LIMITATIONS

1) Data Integration and Interoperability: Crime data is often sourced from various heterogeneous systems and platforms, making data integration and interoperability a challenge. Inconsistent data formats, incompatible databases, and data silos hinder the seamless integration of data for comprehensive analysis.

2) Algorithm Selection and Performance: Choosing the appropriate machine learning or data analysis algorithms for crime prediction can be complex. Different algorithms have varying strengths, limitations, and performance metrics. Selecting the most suitable algorithm for a specific crime analysis task requires careful evaluation and experimentation.

3) Resource and Computing Constraints: Analyzing large volumes of crime data can be computationally intensive and resource-demanding. Limited computational resources, such as processing power and storage, may impede the effectiveness and scalability of crime prediction and analysis projects.

4) Algorithm Selection and Performance: Choosing the appropriate machine learning or data analysis algorithms for crime prediction can be complex. Different algorithms have varying strengths, limitations, and performance metrics. Selecting the most suitable algorithm for a specific crime analysis task requires careful evaluation and experimentation.

5) Resource and Computing Constraints: Analyzing large volumes of crime data can be computationally intensive and resource-demanding. Limited computational resources, such as processing power and storage, may impede the effectiveness and scalability of crime prediction and analysis projects.

6) Interpretability and Explainability: Some advanced machine learning algorithms used in crime prediction, such as deep learning models, often lack interpretability. Interpreting and explaining the factors that contribute to crime patterns can be challenging, limiting the understanding and acceptance of the generated predictions.

IV. PROPOSED SYSTEM

This paper proposes a system for crime prediction and analysis utilizing the support vector machine (SVM) algorithm instead of the Naive Bayes algorithm. The system aims to enhance the accuracy and efficiency of crime prediction by leveraging the strengths of SVM. Compared to Naive Bayes, SVM classifiers offer improved performance and faster prediction. The proposed system utilizes SVM to create an optimal decision boundary, known as a hyperplane, to effectively separate crime data into different classes. By selecting support vectors, the system can accurately predict future crime incidents. The paper discusses the advantages of SVM over naïvebayes and presents the design and implementation of the proposed system for crime prediction and analysis. Experimental results and comparisons between SVM and naive Bayes are provided to demonstrate the superiority of the SVM algorithm in the context of crime prediction.

V. METHODOLOGY

Artificial Intelligence: Artificial intelligence (AI) is a broad field that encompasses a range of technologies, including machine learning, natural language processing, and computer vision. AI systems can be used to automate tasks, identify patterns in data, and make predictions or decisions based on that data.

Machine Learning: Machine learning is a subset of AI that involves training algorithms to learn from data. This process involves feeding large amounts of data into a machine learning algorithm, which then identifies patterns and correlations in the data. Once the algorithm has been trained, it can be used to make predictions on new data.

Programming Languages: Python, R, and Java are three programming languages that are commonly used in AI and machine learning. Each language has its own strengths and weaknesses, and the choice of language often depends on the specific task or project being undertaken.

One of the key advantages of AI and machine learning is their ability to process vast amounts of data quickly and accurately. This makes them well-suited to tasks such as image recognition, speech recognition, and natural language processing. With the right training data and algorithms, an AI system can learn to recognize objects in photographs, translate languages, and even carry on conversations with people.

As AI continues to evolve and become more sophisticated, it is likely to have a major impact on a wide range of industries and applications, from healthcare and finance to manufacturing and transportation.

VI. APPLICATIONS

1) Crime Classification: SVM can be used for crime classification by training a model on historical crime data, such as the type of crime, location, and time of occurrence. The model can then be used to predict the likelihood of a new crime being committed, and classify it based on its characteristics.

2) Criminal Profiling: Criminal profiling involves creating a profile of a suspect based on their behavior and characteristics. SVMs can be used to analyze patterns in criminal behavior and predict the characteristics of the suspect, such as their age, gender, and occupation.

3) Sentencing Analysis: Linear regression can be used to analyze the factors that affect sentencing outcomes, such as the defendant's criminal history, the severity of the crime, and the location of the offense. This information can be used to develop more equitable sentencing guidelines.

VII.   CONCLUSION

Machine learning algorithms, such as SVM, can be used to classify crime patterns accurately and efficiently, helping law enforcement agencies prevent and combat crime. By analyzing crime data, we can understand the contributing factors to crime and predict its occurrence, which can assist in creating strategies and policies for crime prevention.

As AI and machine learning continue to advance, it opens up new opportunities for crime analysis and prediction. By leveraging these technologies, we can improve public safety and help create a more secure and stable society.

In conclusion, crime analysis and prediction using data mining and machine learning techniques can provide valuable insights into crime patterns and help in preventing and combating crime.