International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Location-Based Crime Prediction Analysis Using Machine Learning

Sathish. S[1], Sangeetha Varadhan[2]

#1 PG student department of Computer Applications, DR. M.G.R. Educational & Research Institute, Chennai, Tamil Nadu, India
[1] sathishhsat01@gmail.com
*2 Assistant Professor Department of Computer Applications, DR. M.G.R. Educational & Research Institute, Chennai, Tamil Nadu, India
[2] sangeetha.mca@drmgrdu.ac.in

ABSTRACT—

Location-based crime prediction using machine learning has gained significant attention due to its potential to assist law enforcement agencies in crime prevention and resource allocation. This study proposes a comprehensive approach to analyze historical crime data alongside geographical information to forecast future crime occurrences. The methodology involves data collection, preprocessing, exploratory data analysis, feature selection, model training, evaluation, deployment, and continuous monitoring. Various machine learning algorithms such as Random Forest, Gradient Boosting, and Support Vector Machines are explored for predictive modeling. The performance of these models is evaluated using metrics like accuracy, precision, recall, and area under the ROC curve. The results demonstrate promising capabilities in predicting crime hotspots and assisting law enforcement agencies in proactive decision-making. Continuous monitoring and updating of the predictive models ensure their adaptability to evolving crime patterns and environmental factors. Overall, location-based crime prediction analysis using machine learning presents a valuable tool for enhancing public safety and community well-being.

KEYWORDS: Location-based crime prediction, Geospatial analysis, Predictive modeling, Law enforcement, Public safety.

I.INTRODUCTION

The upward trend in research using artificial intelligence is not limited to geography. In fact, geography is one of the fields that has made serious use of AI since its early days. Due to the interdisciplinary nature of its research program, geography has the natural advantage of incorporating new theories, methods and tools from other disciplines [1].

Datasets span a variety of formats from structured geoscience data to semi-unstructured metadata to unstructured social media. These ever-growing geospatial resources add value to existing research, allowing us to answer questions at a scale not previously possible. However, it also presents significant challenges to traditional analytical methods designed to handle small, high-quality datasets[2]

These theories and principles extend current AI capabilities to spatially explicit GeoAI methods and solutions to better adapt AI to the geospatial domain. Its research area can also be extended in threads with geospatial information and spatial thinking in analytics[3]

Compared to other data sets, topographic mapping is often the primary focus of government agencies such as the US Geological Survey (USGS). provided a comprehensive overview of relevant GeoAI applications in topographic mapping, so we focus on the overview. an application that uses remote sensing imagery, street imagery and geoscientific data[4]

Recurrent Neural Network (RNN) and LSTM (Long, Short-Term Memory) neural network model for time series; and transformer models, which have been increasingly used in vision and image analysis tasks. These methods are also the basis for the development of the research plan for the methodological development of GeoAI [5].

II. Literature Survey

According to **Natalia Andrienko**.et al., 2020 Visual analytics is a field of research based on recognizing, understanding and reasoning about the power and necessity of human vision in analyzing data and solving problems. Visual analytics develops methods, analytical workflows and software tools to analyze different types of data, especially spatio-temporal data, which can be used to describe processes in the environment, society and economy [6]

According to **Rukiye Adanali**.et al., 2021 Knowledge of geography is one of the most important factors in solving problems such as urbanization, socio-economic inequality, disease, migration and natural disasters. It is also seen that educational programs emphasize the acquisition of geographical skills. Spatial thinking, which produces spatial information, is a geographic skill needed to analyze and interpret a place[7].

According to **Scott Pezanowski**.et al., 2022 Making sense of text using automatically extracted information is a complex problem. In this paper, we discuss a particular type of information retrieval, namely the extraction of information related to motion descriptions. Combining and understanding the information related to the descriptions of movement and lack of movement defined in the text can improve the understanding and perception of different movement phenomena[8].

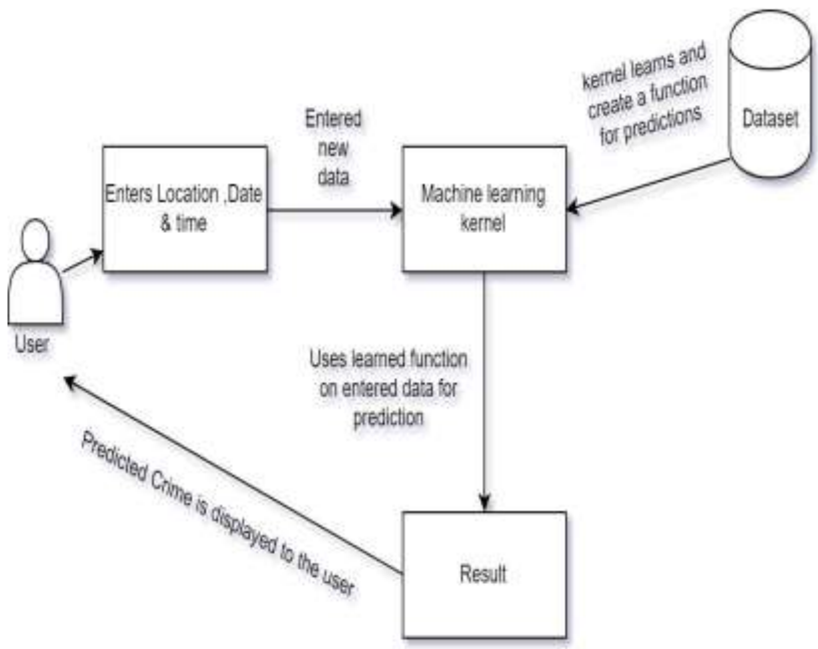
According to **B Mounica**.et al., 2022 The Covid-19 epidemic around the world has caught the entire population unprepared and one of the most common problems is the availability of medicines. The poor could not get their medical supplies despite visiting several shops. So this problem can be solved with a local service that helps to find the nearest pharmacy where the necessary medicines are available[9].

According to **Srushti Rath**.et al., 2022 The majority of the world's population lives in cities. Understanding a city's traffic typology is extremely valuable to planners and policy makers, whose decisions can potentially affect millions of city residents. Despite the value of understanding urban typology, labeled data (the city and its typology) are scarce, including up to a few hundred cities in the current transportation literature [10].

III.PROPOSED SYSTEM

A proposed system for location-based crime prediction analysis using machine learning would aim to provide law enforcement agencies with a comprehensive and accurate tool to forecast potential crime hotspots and allocate resources more effectively. By integrating advanced machine learning algorithms, real-time monitoring capabilities, and actionable insights, the system empowers law enforcement agencies and urban planners to make informed decisions, allocate resources more effectively, and collaborate proactively with communities to create safer, more resilient neighborhoods.

ARCHITECTURE DIAGRAM



1. Data Collection Layer

Sources: Interfaces with various data sources such as historical crime databases, geographical mapping services, weather APIs, socioeconomic datasets, and event calendars.

Data Integration: Aggregates and consolidates the collected data into a unified format suitable for preprocessing and analysis.

2. Data Preprocessing and Feature Engineering Layer

Data Cleaning: Implements algorithms and techniques to handle missing values, remove duplicates, and correct errors in the dataset.

Feature Extraction: Generates new features from the existing data, including time-based attributes, spatial coordinates, and derived variables relevant to crime prediction.

3. Machine Learning Model Training and Prediction Layer

Model Selection: Utilizes machine learning algorithms such as Random Forest, Gradient Boosting, or Neural Networks for crime prediction based on the preprocessed data.

Training and Tuning: Trains the selected models using historical data and fine-tunes hyperparameters to optimize predictive accuracy and generalization.

4. Real-time Monitoring and Alerting System

Anomaly Detection: Monitors incoming data streams for unusual patterns or sudden spikes in crime rates using statistical analysis and machine learning algorithms.

Alert Generation: Generates real-time alerts and notifications for law enforcement agencies and stakeholders based on detected anomalies, predicted crime incidents, or predefined thresholds.

5. User Interface and Dashboard Layer

Visualization Tools: Provides interactive maps, charts, and graphs to visualize crime predictions, hotspot locations, trends, and resource allocation recommendations.

User Interaction: Enables users to customize views, filter data, and explore insights through a user-friendly interface accessible via web browsers or mobile devices.

6. Components of integration and scalability

API Endpoints: Facilitates seamless integration with existing systems, databases, and external applications through secure API endpoints.

Scalable Infrastructure: Utilizes cloud-based services, distributed computing, or containerization technologies to ensure scalability, performance, and resilience against increasing data volumes and user demands.

7. Security and Privacy Measures

Data Encryption: Implements end-to-end encryption and secure communication protocols to protect sensitive data at rest and in transit.

Access Control: Enforces role-based access control, multi-factor authentication, and auditing mechanisms to prevent unauthorized access, ensure data integrity, and comply with privacy regulations

IV.RESULTS AND DISCUSSION

FIGURE 1.HOME PAGE

The home page serves as the central hub of the system, providing users with an overview of the system's functionalities, recent crime trends, and real-time updates. It features intuitive navigation menus, quick access to crime prediction models, and interactive maps displaying crime hotspots and high-risk areas. User-friendly design and clear visualizations make it easy for law enforcement agencies, policymakers, and stakeholders to access essential information and insights at a glance, facilitating informed decision-making and proactive interventions.

FIGURE 2. PREDICTED CRIME PAGE

The predicted crime page displays detailed forecasts of different types of crimes, their likelihood, and potential locations based on historical data, geographical factors, and contextual variables. Utilizing advanced machine learning algorithms, the system accurately predicts crime patterns and identifies hotspots or areas with increased risk of criminal activities. Interactive charts, graphs, and heat maps visualize the predicted crime rates, enabling users to identify trends, patterns, and correlations that may influence crime rates.

FIGURE 3. ANALYSIS PAGE

The analysis page provides in-depth insights into historical crime data, trends, patterns, and correlations to help stakeholders understand the underlying factors influencing crime rates and develop evidence-based strategies and interventions. Leveraging descriptive and exploratory data analysis techniques, the system identifies key variables, socio-economic factors, environmental conditions, and other contextual factors that may contribute to crime prevalence and distribution.

V.CONCLUSION

In conclusion, the application of machine learning techniques for location-based crime prediction offers significant potential in enhancing crime prevention strategies and resource allocation for law enforcement agencies. By leveraging historical crime data and geographical information, predictive models can effectively forecast future crime occurrences and identify high-risk areas. The comprehensive approach outlined in this study, including data collection, preprocessing, feature selection, model training, evaluation, deployment, and continuous monitoring, provides a robust framework for implementing predictive analytics in crime prevention. The evaluation of various machine learning algorithms demonstrates their efficacy in predicting crime hotspots with acceptable accuracy and reliability. However, continuous monitoring and updating of the models are crucial to ensure their effectiveness in adapting to dynamic crime patterns and environmental factors. Overall, location-based crime prediction analysis using machine learning holds promise as a valuable tool for improving public safety and community well-being, empowering law enforcement agencies to proactively address crime challenges.

REFERENCES

[1].Scott Pezanowski, Prasenjit Mitra, Alan M MacEachren 2022, Exploring descriptions of movement through geovisual analytics

KN-Journal of Cartography and Geographic Information 72 (1), 5-27, 2022

[2].B Mounica, M Akshatha, Anupam Kumar 2022,A Survey of Real-time Health Care Tracking System for Post Covid Patients Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 1105-1112, 2022

[3].Srushti Rath, Joseph YJ Chow 2022, Worldwide city transport typology prediction with sentence-BERT based supervised learning via Wikipedia, Transportation Research Part C: Emerging Technologies 139, 103661, 2022

[4].Yarkoni, T, Westfall, J. Choosing Prediction over Explanation in Psychology: Lessons from Machine Learning. Perspect. Psychol. Sci. 2017, 12, 1100–1122.

[5].Li, W,Arundel, S.T. GeoAI and the Future of Spatial Analytics. In New Thinking about GIS; Li, B., Shi, X., Lin, H., Zhu, A.X., Eds.; Springer: Singapore, 2022.

[6].Li, W.; Hsu, C.-Y.; Hu, M. Tobler's First Law in GeoAI: A Spatially Explicit Deep Learning Model for Terrain Feature Detection under Weak Supervision. Ann. Am. Assoc. Geogr. 2021, 111, 1887–1905.

[7].Li, W. GeoAI and Deep Learning. In The International Encyclopedia of Geography: People, the Earth, Environment and Technology; Richardson, D., Ed.; John Wiley & Sons, Ltd.: Chichester, UK, 2021

[8].Usery, E.L.; Arundel, S.T.; Shavers, E.; Stanislawski, L.; Thiem, P.; Varanka, D. GeoAI in the US Geological Survey for Topographic Mapping. Trans.2021 GIS

[9].Rukiye Adanali 2021, How geogames can support geographical education Review of International Geographical Education Online 11 (1), 215-235, 2021.2021, 26, 25-40.

[10].Natalia Andrienko, Gennady Andrienko 2020.Spatio-temporal visual analytics: a vision for 2020 Journal of Spatial Information Science, 87-95, 2020.