iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70309

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Predictive Crime Rate Analysis System

Ishika Singh[1], Garima Bhandari[2], Sakshi Khatwani[3]
Dept. of Computer Science and Engineering SRM Institute of Science and Technology NCR

Abstract: Crime prediction plays a crucial role in enhancing urban safety and aiding law enforcement agencies in proactive decision-making. This study presents a predictive crime rate analysis system utilizing advanced machine learning techniques to analyse historical crime data and forecast future crime trends. The proposed approach integrates multiple models, including Long Short-Term Memory (LSTM), Support Vector Machines (SVM), Decision Trees, and Random Forest, to identify patterns and correlations across various crime types, time periods, and geographic locations. Feature engineering techniques are employed to preprocess data, handle missing values, and extract key insights. The system provides real-time crime risk assessments, helping law enforcement optimize resource allocation and strategic planning. Additionally, a web-based platform is developed to enhance accessibility, allowing users to visualize crime predictions based on time, location, and crime category. Experimental results demonstrate the effectiveness of the proposed model, achieving high accuracy in crime prediction and forecasting. This research contributes to improving urban safety by leveraging data-driven insights for crime prevention and law enforcement strategies.

Index Terms: Crime prediction, machine learning, LSTM, crime forecasting, law enforcement, urban safety.

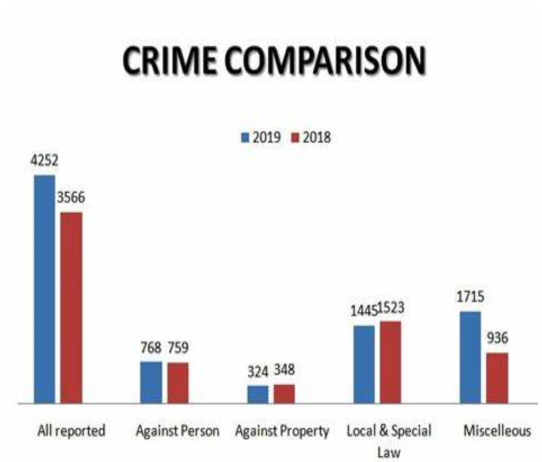
I. INTRODUCTION

Crime remains a persistent threat to urban safety, significantly impacting the quality of life and economic stability of communities. In 2019 alone, the United States witnessed 435 mass shooting events, resulting in 517 fatalities, 1,648 injuries, severe property losses, and widespread societal distress. Given these alarming statistics, crime risk assessment is essential for individuals and law enforcement agencies to prevent and mitigate potential criminal activities.

The availability of urban data in cities like Chicago has enabled researchers to explore various crime-related problems, including crime hotspot detection, classification, rate inference, and count prediction.

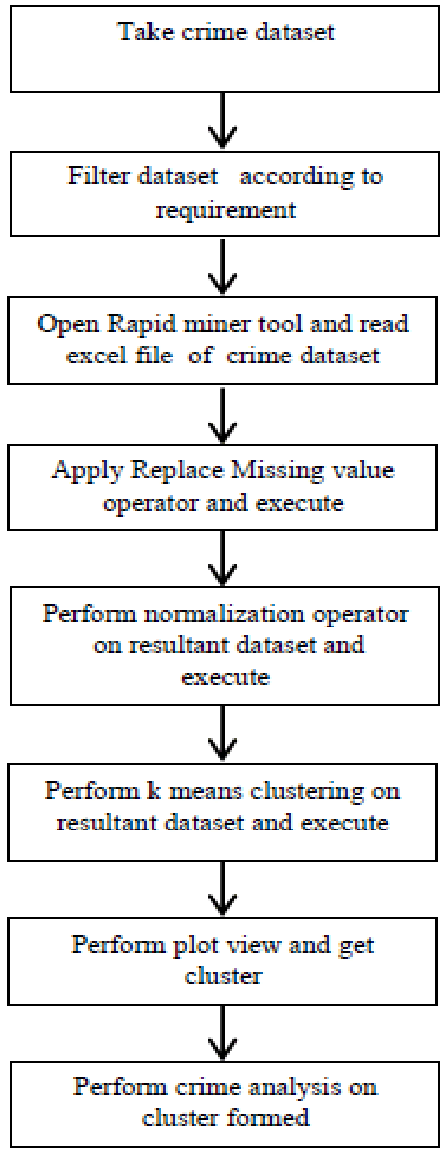
ThStudies indicate that crime occurrences often correlate with human mobility patterns, where higher population density may lead to increased incidents of larceny and other offenses. However, due to the uneven development of urban infrastructure and privacy concerns, many cities do not publicly disclose crime-related data. This lack of data creates challenges, especially for newcomers such as tourists, in assessing safety risks effectively.

Recent advancements in machine learning, particularly transfer learning, offer innovative solutions to these challenges. Transfer learning enables the use of knowledge from a data-rich source city to solve similar tasks in a data-scarce target city. For example, while New York City (NYC) has extensive urban datasets due to its long-standing open-data policies, cities like Los Angeles (LA) face data scarcity due to privacy regulations and high collection costs. The inconsistency in available context data between source and target cities hinders the performance of crime risk prediction models.



Crime Prediction Using Machine Learning

With the increasing integration of machine learning in various domains, crime analysis and prediction have emerged as critical applications. Traditional crime prevention methods often rely on historical data and expert knowledge but lack adaptability to evolving crime patterns. Machine learning algorithms, however, can efficiently process vast amounts of data, uncover hidden patterns, and develop accurate predictive models.This research focuses on leveraging machine learning algorithms to analyze historical crime data and develop a robust crime prediction model.



By utilizing diverse datasets encompassing temporal, spatial, demographic, and socioeconomic factors, the study aims to identify potential crime hotspots with high accuracy. Techniques such as classification, regression, and clustering are employed to extract meaningful patterns from the data.

The enhance accessibility and usability, the predictive model is integrated into a web-based platform, designed for real-time crime analysis and prevention.A straightforward approach to addressing this issue involves training a model using common contextual data from the source city while disregarding inconsistent, city-specific datasets.

This project focuses on predictive analysis of crime rates using historical data and machine learning techniques. By analyzing patterns and trends, the system aims to provide insights that can aid in decision-making and resource allocation. The model considers various factors such as location, time, and crime type to generate accurate forecasts. The ultimate goal is to enhance public safety by enabling proactive crime prevention strategies.

II. LITERATURE REVIEW

With the evolving nature of society, crime patterns are also shifting, and crime rates continue to rise. Traditional methods of crime analysis are becoming outdated. To address this, various deep learning-based approaches—supervised, semi-supervised, and unsupervised—are being utilized to predict crimes and analyze trends. Integrating modern technology into crime prediction enables authorities to take proactive measures and enhance public safety.

A. Baseline Networks

Feed-forward artificial neural networks (ANNs), particularly multi-layer perceptrons (MLPs), have proven effective in time-series forecasting as they do not require prior knowledge of data distribution [10]. Deep learning techniques enhance time-series forecasting by capturing temporal dependencies and structures more efficiently. The integration of recurrent neural networks (RNNs) in deep learning has significantly improved the handling of sequential data, particularly in forward-dependent networks. These models have demonstrated success in solving complex real-world problems.

Long Short-Term Memory (LSTM) networks can process entire sequences instead of individual data points, making them highly suitable for time-series analysis. Schusterfurther refined this concept by introducing Bidirectional LSTMs (Bi-LSTMs), which incorporate both forward and backward dependencies, improving contextual understanding. Said et al. and Kim et al. highlighted how Bi-LSTMs effectively capture variable correlations and changes in multivariate time-series data, enhancing predictive accuracy.

An alternative approach to reducing redundant data in sequential modeling is the use of attention-based models. These models focus on relevant parts of input data while filtering out unnecessary information. Attention mechanisms work well in sequence modeling because the distance between input and output layers does not affect dependency modelling. To address this, self-attention (or intra-attention) mechanisms have been introduced. These mechanisms establish relationships between different positions within a sequence, preserving long-range dependencies [29]. Unlike RNNs, self-attention layers enable faster execution in most cases [30]. Attention and self-attention models have been widely used in computer vision and natural language processing (NLP) applications, such as speech enhancement, text prediction, and summarization [25], [31]–[34]. Despite these advancements, a single approach is often insufficient for analyzing complex, feature-dependent data, necessitating model fusion strategies.

B. Crime Prediction And Classification

Big Data Analytics(BDA) has demonstrated remarkable success in criminology,helping to uncovertrends and relationshipswithin crime data.

Several studies have applied clustering techniques for crime pattern analysis. Agarwal et al. [40] and Tayal et al. [41] employed K-means clustering to identify crime patterns based on yearly trends. However, like spatio-temporal systems, clustering models require a minimum threshold of data points per category, making them less effective for time-based analysis.

Rayhan et al. [9] developed an attention-based deep learning model that captures nonlinear spatial dependencies and temporal patterns for specific crime categories. By emphasizing feature dependencies, this model outperformed standard LSTMs in time-series crime prediction. Additionally, it maintained interpretability by dynamically establishing spatial-temporal relationships based on past crime occurrences and recurring trends. However, its performance is limited by the need for large training datasets for each crime category.

For crime classification, Kumar et al. [42] proposed a Naïve Bayes-based approach that combines historical crime data with incident-level reports to predict likely offenders. While this method worked well for specific crime types, its overall accuracy was around 50%. A similar approach was taken by ToppiReddy et al. [43], who used K-Nearest Neighbors (KNN) and Naïve Bayes classifiers to predict whether a crime would occur and classify its type based on location and time factors.

Advancements in clustering-based crime classification have also been made. Sivaranjani et al. [44] and Pednekar et al. [45] developed a hybrid approach combining K-means, Agglomerative, and DBSCAN clustering techniques. By merging insights from multiple clustering methods, their model improved classification accuracy. However, predicting the exact time of a crime remained a challenge.

Gradient-boosted decision trees have outperformed traditional classifiers such as KNN, Naïve Bayes, and Random Forests when applied to large datasets with feature correlations. Feng et al. [2] utilized forecasting methods to generate crime trend data and employed tree-based classification models to predict crime categories based on time and location. Their model outperformed KNN and Naïve Bayes but required merging smaller crime categories for better trend analysis.

Although various deep learning models excel in different aspects of crime forecasting, no single model has achieved optimal performance across all scenarios. A comprehensive review of existing state-of-the-art techniques is presented in Table 1, highlighting the need for further advancements to overcome current limitations in crime prediction and classification.

1) Existing Model

Traditional statistical models, such as linear regression and autoregressive integrated moving average (ARIMA), have been used to identify temporal trends in crime data. However, these models often struggle with complex, non-linear relationships inherent in crime datasets.Machine learning-based approaches have gained prominence in recent years, leveraging techniques such as decision trees, support vector machines (SVM), and k-nearest neighbors (KNN) for crime classification and prediction.

These models perform well in pattern recognition but may lack the temporal awareness required for sequential data analysis.Despite these advancements, existing models often face challenges related to data availability, real-time adaptability, and interpretability. Our proposed hybrid AI-driven model aims to address these limitations by integrating deep learning, real-time analytics, and geospatial intelligence.

2) Problem Statement

Traditional recommendation systems struggle to provide real-time, emotion-driven music suggestions. They rely on static data and do not account for dynamic emotional changes throughout the day. Additionally, existing facial expression-based models have limitations in accuracy due to varying lighting conditions, occlusions, and diverse facial structures. There is a need for a robust system that can adapt to real-time emotional variations while incorporating contextual awareness for better personalization.

3) Proposed Solution

We propose a hybrid AI-driven crime rate prediction model that integrates machine learning, real-time data analytics, and geospatial intelligence to improve forecasting accuracy and proactive crime prevention. The key components of our solution are:

a) Multi-Source Data Integration

Unlike traditional models that rely solely on past crime reports, our approach aggregates diverse datasets, including:

- Historical crime records for pattern recognition.
- Economic indicators (unemployment rate, GDP, inflation) to assess correlations between economic conditions and crime trends.
- Social media and news sentiment analysis using NLP and transformer models to detect emerging threats and crime-related discussions.
- Weather and environmental factors (e.g., heat waves, urban density) which have been linked to fluctuations in crime rates.
- Law enforcement reports and citizen complaints to capture real-time crime incidents and response efficiency.

b) Advanced AI & Machine Learning Techniques

To ensure high accuracy and adaptability, we incorporate:

- Deep Learning Models (LSTMs, Transformers) to process time-series data and predict crime hotspots based on past trends.
- Graph Neural Networks (GNNs) for analyzing spatial relationships between crime-prone areas and external influences like law enforcement presence or socio-economic shifts.
- AutoML-based Feature Selection to optimize prediction models by dynamically choosing the most relevant crime factors.

C. Geo-Spatial & Temporal Crime Mapping

We utilize Geographical Information Systems (GIS) and heatmap-based clustering for crime trend visualization.

1) Spatio-Temporal Crime Prediction identifies high-risk zones based on past patterns and real-time data.
2) Dynamic Crime Hotspot Detection leverages anomaly detection to flag unusual crime surges.
3) Predictive Risk Assessment for Urban Planning, aiding policymakers in resource allocation for high-risk neighbourhoods.

a) .Adaptive Learning & Real-Time Alerts

Unlike static models, our system dynamicallyadapts to real-time inputs, guaranteeing:

- Continuous Model Retraining using reinforcement learning to improve accuracy.
- Automated Alerts & Crime Forecast Reports for law enforcement, enabling data-driven decision-making.
- Explainable AI (XAI) for transparent model interpretations, assisting authorities in understanding crime drivers.

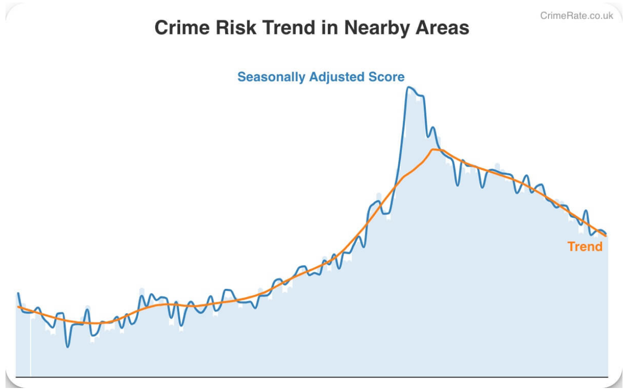
b) Law Enforcement & Policy Implications

- Predictive Policing Strategy: Assists law enforcement in resource deployment based on high-risk predictions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

- Community Awareness Dashboard: Provides crime insights to local authorities and citizens for increased vigilance.
- Ethical & Privacy Considerations: Ensures compliance with data privacy laws and mitigates algorithmic bias.

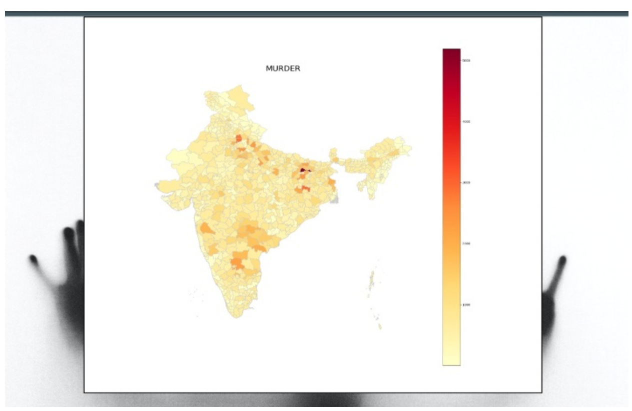
By integrating real-time data streams, deep learning, and geospatial analytics, our proposed solution aims to revolutionize crime prediction with enhanced precision, aiding in proactive law enforcement and crime prevention measures.

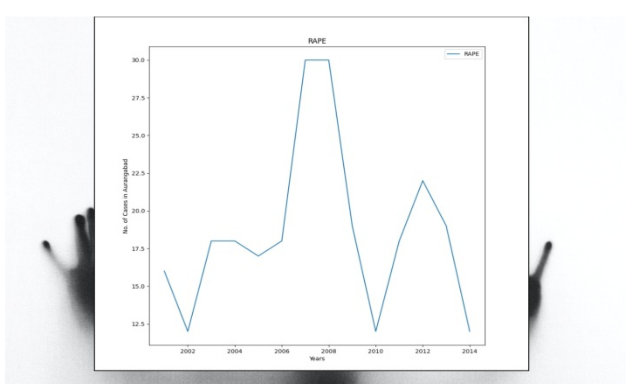


III. CONCLUSION

Our proposed AI-driven crime rate prediction model offers a data-driven approach to proactive crime prevention by integrating machine learning, geospatial analytics, and real-time data processing. Unlike traditional methods that rely solely on historical crime records, our model incorporates multi-source data such as socio-economic factors, social media sentiment, and environmental conditions to enhance predictive accuracy. By leveraging deep learning, graph neural networks, and real-time adaptive learning, the system can identify crime patterns, detect emerging hotspots, and assist law enforcement agencies in strategic decision-making.

Furthermore, the model's explainability and ethical considerations ensure responsible AI implementation, addressing concerns related to bias and data privacy. With its predictive policing capabilities, community awareness dashboard, and policy-driven insights, our solution has the potential to transform crime prevention strategies, making cities safer, more secure, and better prepared to combat future crime trends.





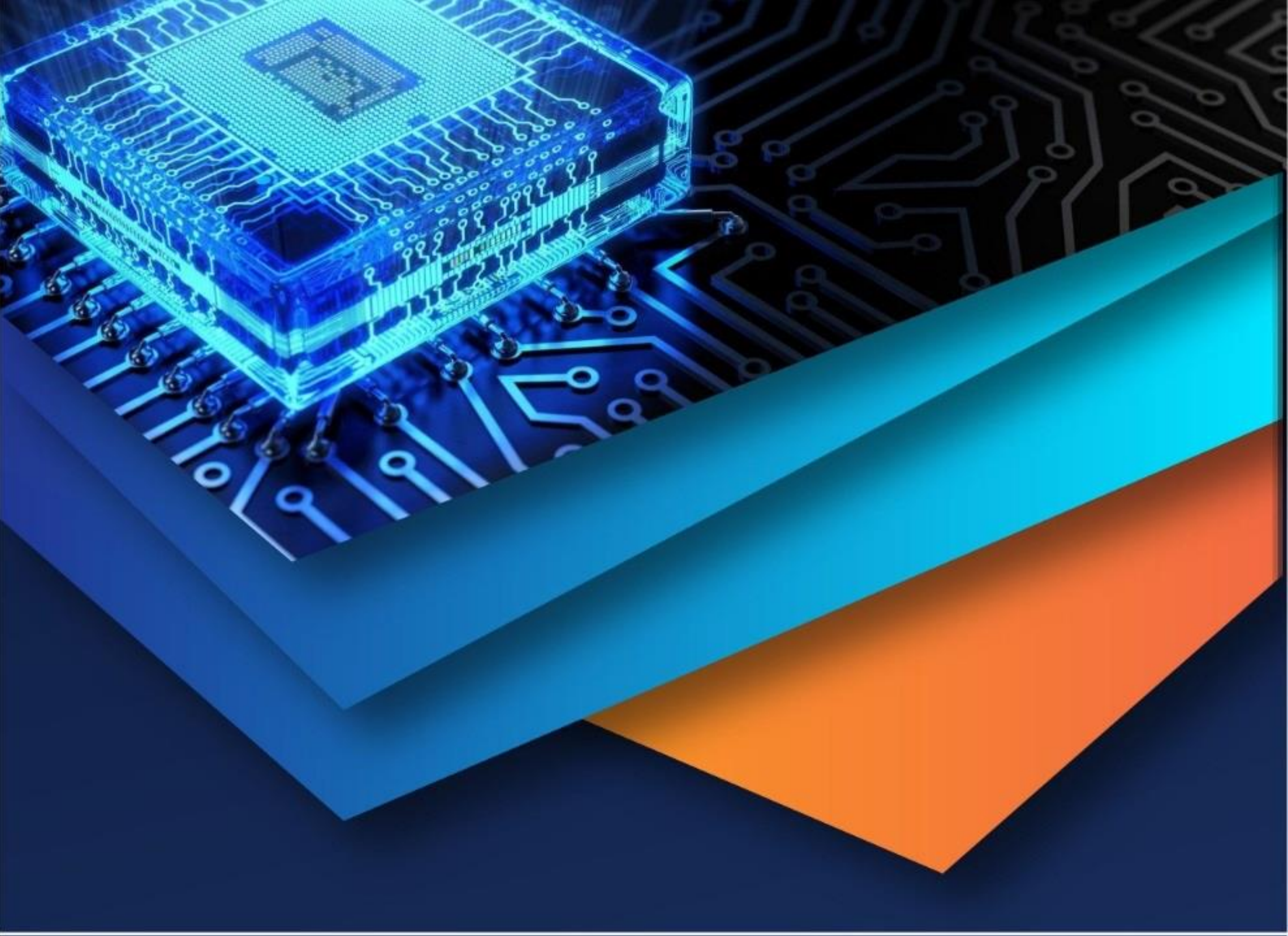
IV. ACKNOWLEDGMENT

Wesincerely appreciate, our project guide Mr. Chiranjeet Dutta,for his invaluable mentorship,and unwaveringassistance during this research. His insightful feedback and expertise have been instrumental in shaping our work.

We also extend our appreciation to SRM Institute of Science and Technology for providing us withessential facilities and a supportive academic setting to conduct this study.

REFERENCES

[1] GunViolenceArchive. List of Mass Shootings in the United States in 2019. [Online].Available:https://www.gunviolencearchive.org/reports/mass-shooting
[2] L. Xia et al., "Spatial–temporal sequential hypergraph network for crime prediction with dynamic multiplex relation learning," in Proc. 30th Int. Joint Conf. Artif. Intell., Aug. 2021, pp. 1631–1637.
[3] P. Das, A. K. Das, J. Nayak, D. Pelusi, and W. Ding, "Incremental classifier in crime prediction using bi-objective particle swarm optimization," Inf. Sci., vol. 562, pp. 279–303, Jul. 2021.
[4] F. Yi, Z. Yu, F. Zhuang, and B. Guo, "Neural network based continuous conditional random field for fine-grained crime prediction," in Proc. 28th Int. Joint Conf. Artif. Intell., Aug. 2019, pp. 10–16.
[5] Y. Liang et al., "Fine-grained urban flow prediction," in Proc. Web Conf., 2021, pp. 1833–1845.
[6] M. Fang, L. Tang, X. Yang, Y. Chen, C. Li, and Q. Li, "FTPG:A fine-grained traffic prediction method with graph attention network using big trace data," IEEE Trans. Intell. Transp. Syst., vol. 23, no. 6, pp. 5163–5175, Jun. 2021
[7] X. Wu, C. Huang, C. Zhang, and N. V. Chawla, "Hierarchically structured transformer networks for fine-grained spatial event forecasting," in Proc. Web Conf., Apr. 2020, pp. 2320–2330.
[8] B. Liu, S. Yan, J. Li, Y. Li, J. Lang, and G. Qu, "A spatiotemporal recurrent neural network for prediction of atmospheric PM2.5: A case study of Beijing," IEEE Trans. Computat. Social Syst., vol. 8, no. 3, pp. 578–588, Jun. 2021.
[9] B. Huang, H. Zhu, D. Liu, N. Wu, Y. Qiao, and Q. Jiang, "Solving last-mile logistics problem in spatiotemporal crowdsourcing via role awareness with adaptive clustering," IEEE Trans. Computat. Social Syst., vol. 8, no. 3, pp. 668–681, Jun. 2021.
[10] B. Zhou et al., "ESCORT: Fine-grained urban crime risk inference leveraging heterogeneous open data," IEEE Syst. J., vol. 15, no. 3, pp. 4656–4667, Sep. 2021.
[11] B. Zhou, L. Chen, F. Zhou, S. Li, S. Zhao, and G. Pan, "Dynamic road crime risk prediction with urban open data," Frontiers Comput. Sci., vol. 16, no. 1, pp. 1–13, Feb. 2022.
[12] Ni, C. Zhang, and X. Gu, "Transfer model collaborating metric learning and dictionary learning for cross-domain facial expression recognition," IEEE Trans. Computat. Social Syst., vol. 8, no. 5, pp. 1213–1222, Oct. 2020.
[13] L. Zheng, G. Liu, C. Yan, C. Jiang, and M. Li, "Improved TrAdaBoost and its application to transaction fraud detection," IEEE Trans. Computat. Social Syst., vol. 7, no. 5, pp. 1304–1316, Oct. 2020.
[14] X. Kong, H. Gao, G. Shen, G. Duan, and S. K. Das, "FedVCP: A federated-learning-based cooperative positioning scheme for social Internet of Vehicles," IEEE Trans. Computat. Social Syst., vol. 9, no. 1, pp. 197–206, Feb. 2021.
[15] Z. Chen, G. He, J. Li, Y. Liao, K. Gryllias, and W. Li, "Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery," IEEE Trans. Instrum. Meas., vol. 69, no. 11, pp. 8702–8712, Nov. 2020.
[16] C. Tian, X. Zhu, Z. Hu, and J. Ma, "A transfer approach with attention reptile method and long-term generation mechanism for few-shot traffic prediction," Neurocomputing, vol. 452, pp. 15–27, Sep. 2021.



10.22214/IJRASET

INDEX COPERNICUS

45.98

JISRAF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY