REST Journal on Data Analytics and Artificial Intelligence
Vol: 4(1), March 2025
REST Publisher; ISSN: 2583-5564
Website: http://restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/4/1/8

Web-Based Geographic Information System for Analyzing and Visualization Metropolitan Crime Data

***Sneha Satya Vardhan Tiruveedula, Yashwanth Kosuri, Vathsalya Taduru, M. Trupthi**
Department of Artificial Intelligence, School of Engineering Anurag University, Hyderabad, India.
**Corresponding Author Email:ssvardhant@gmail.com*

Abstract: *The hearing-impaired community relies heavily on sign language as a means of communication since it enables them to properly express their feelings and ideas. However, because it necessitates learning complex hand movements and postures, non-signers may find it difficult to comprehend sign language. In order to overcome this obstacle, we suggest a real-time sign language translation system that uses deep learning, computer vision, and picture categorization methods to close the communication gap between the public and the deaf and mute communities. This system uses Convolutional Neural Networks (CNNs) to recognize and decipher real-time sign language gestures recorded by a camera. The accuracy and resilience of the model are improved by the effective picture preprocessing made possible by the integration of OpenCV, which includes noise reduction, segmentation, and feature extraction. High identification rates across a variety of signs are ensured by training the model on an extensive dataset of hand gestures using supervised learning techniques. Through the integration of computer vision and deep learning with an easy-to-use interface, this project provides a scalable and accessible solution for sign language identification, increasing the efficiency and inclusivity of communication for the speech-impaired and hearing communities. Future research will concentrate on expanding the system to recognize dynamic gestures for full sentence translation, strengthening the dataset, and honing gesture classification algorithms.*

Keywords: *Computer Vision, Convolutional Neural Networks, Deaf and Mute Communication, Hand Gesture Recognition, Image Classification, Speech-to-Text Conversion*

1. INTRODUCTION

In recent years, the prevalence of drug-related crimes has become a significant challenge for communities, law enforce- ment agencies, and policymakers. Understanding the patterns, locations, and dynamics of these incidents is crucial for im- plementing effective interventions and improving public safety [1-2]. This project aims to develop a comprehensive web-based Ge- ographic Information System (GIS) that facilitates the tracking and analysis of drug-related crimes, providing a powerful tool for stakeholders involved in crime prevention and response [3]. The proposed GIS will leverage advanced web scraping tech- niques to gather real-time data from various sources, includ- ing news articles, government reports, and law enforcement databases. By integrating these data into an interactive mapping interface, users will be able to visualize crime hotspots, identify trends over time, and conduct in-depth analyses of drug-related incidents within specific geographic areas [4-5]. The system will not only aid in situational awareness, but also support strategic de- casino making by law enforcement and community organiza- tions. This project is designed with user experience in mind, fea- turing an intuitive interface that accommodates both technical and nontechnical users. Through robust data visualization and reporting capabilities, the GIS will empower stakeholders to make informed decisions and ultimately contribute to reducing drug-related crimes in their communities [5]. Some advanced systems incorporate Named Entity Recogni- tion (NER) and Conditional Random Fields (CRF) for a more

precise extraction of crime locations from text [6-7]. These methods allow systems to automatically detect and classify sentences as crime location data, which can then be plotted on a map. This approach is particularly useful for law enforcement agencies, as it automates the tedious process of manually locating crime spots from news reports and other textual sources [8]. For drug-related crime, there are systems specifically de- signed to extract detailed information on drug tra*ffi*cking. Rule- based extraction tools have been developed to analyze how drugs are hidden, transported, and sold [9-10]. These systems gather data on the nationalities of tra*ffi*ckers, the types of drugs being traded, and other relevant patterns. This extracted data is then visualized through GIS, enabling authorities to gain deeper in- sights into the scale and distribution of drug-related crimes. Another notable system, CAINES, is based on a knowledge engineering approach. It relies on sub-language analysis to per- form crime data extraction from online sources. By employing knowledge-based framework, this system aids in the more accu- rate extraction and classification of crime incidents, making it a powerful tool for law enforcement agencies seeking to enhance crime detection and prevention strategies [11]. These existing systems demonstrate the potential of com- binning GIS with text mining and natural language processing techniques. They provide valuable insights into crime patterns, hotspot detection, and trends analysis, which can significantly improve crime prevention e*ff*ort [12].

2. LITERATURE SURVEY

Existing System: Existing systems in crime mapping and analysis have evolved significantly, with various approaches leveraging Geo- graphic Information Systems (GIS) and text mining techniques to provide actionable insights for law enforcement and the pub- lick [13-15]. One notable system implements web-based GIS for crime hotspot mapping using open-source software. Using techniques such as choropleth mapping, grid mapping, and kernel density estimation, these systems offer comprehensive visualizations of spatial crime patterns. The advantage of web-based platforms is that they allow crime data to be accessed in real-time by mul- tiple users, making it easier for small and medium-sized police departments in developing countries to analyses crime trends and allocate resources efficiently [16]. Other systems focus on web-based crime analysis by extract- ing data from online news articles. These systems utilize text mining techniques to classify crime and non-crime articles and extract relevant crime-related entities [17-18]. For example, through natural language processing (NLP) and information retrieval (IR), text mining systems identify crime hotspots and detect trends from unstructured data sources, such as online newspa- pers. By processing vast amounts of text data, they assist in generating a clearer picture of crime activity without relying on official reports alone [19-21].

Limitation of existing system: The current systems for crime mapping and analysis, though advanced, have several limitations that hinder their ability to fully address specific crime types. Most of these tools take a generalized approach, mapping a broad range of crimes such as burglary, assault, and theft, rather than focusing specifically on drug-related offenses. This broad scope can dilute the insights needed to track drug crimes, which often have unique patterns and dynamics. Additionally, many existing systems rely heavily on struc- turned data from official sources like police reports. While value- able, these sources can miss the timeliness and depth provided by unstructured data from online news articles. Some systems use text mining to extract crime-related data from such arty- cles, but they typically focus on crime as a whole and are not optimized for drug-specific analysis. Drug-related crimes re- quire specialized extraction of entities like drug types, traffick- ing routes, and key locations, which existing systems often fail to capture effectively. Another major limitation is the difficulty in handling the nat- ural variability of language in news reports. Current systems struggle with accurately identifying crime locations and fill- tering out irrelevant data, particularly when it comes to drug crimes. Issues like handling ambiguous terms, eliminating du- plicates, and cleaning noisy data make it challenging to consis- tently extract and map drug-specific information. Furthermore, while many systems use open-source GIS tools, they often lack real-time data integration and advanced features like personalized crime searches. This limits their ability to provide real-time, location-specific insights on drug crimes, which law enforcement agencies and public users may find crucial for timely decision-making. Unlike these existing solutions, your approach focuses ex- clusively on drug-related crimes by scraping online news data and using advanced natural language processing techniques like BERT and spaCy. By honing in on drug crimes, your system will fill a crucial gap, offering a more detailed and focused map- ping solution that leverages both structured and unstructured data sources to provide comprehensive and actionable insights. Some advanced systems incorporate Named Entity

Recognize- ton (NER) and Conditional Random Fields (CRF) for a more precise extraction of crime locations from text [6-7]. These methods allow systems to automatically detect and classify sentences as crime location data, which can then be plotted on a map. This approach is particularly useful for law enforcement agencies, as it automates the tedious process of manually locating crime spots from news reports and other textual sources [8]. For drug-related crime, there are systems specifically de- signed to extract detailed information on drug trafficking. Rule- based extraction tools have been developed to analyze how drugs are hidden, transported, and sold [9-10]. These systems gather data on the nationalities of traffickers, the types of drugs being traded, and other relevant patterns. This extracted data is then visualized through GIS, enabling authorities to gain deeper in- sights into the scale and distribution of drug-related crimes. Another notable system, CAINES, is based on a knowledge engineering approach. It relies on sub-language analysis to per- form crime data extraction from online sources. By employing knowledge-based framework, this system aids in the more accu- rate extraction and classification of crime incidents, making it a powerful tool for law enforcement agencies seeking to enhance crime detection and prevention strategies [11]. These existing systems demonstrate the potential of com- binning GIS with text mining and natural language processing techniques. They provide valuable insights into crime patterns, hotspot detection, and trends analysis, which can significantly improve crime prevention effort [12].

3. LIMITATION OF EXISTING SYSTEM

The current systems for crime mapping and analysis, though advanced, have several limitations that hinder their ability to fully address specific crime types. Most of these tools take a generalized approach, mapping a broad range of crimes such as burglary, assault, and theft, rather than focusing specifically on drug-related offenses. This broad scope can dilute the insights needed to track drug crimes, which often have unique patterns and dynamics. Additionally, many existing systems rely heavily on struc- turned data from official sources like police reports. While value- able, these sources can miss the timeliness and depth provided by unstructured data from online news articles. Some systems use text mining to extract crime-related data from such arty- clues, but they typically focus on crime as a whole and are not optimized for drug-specific analysis. Drug-related crimes re- quire specialized extraction of entities like drug types, traffick- ing routes, and key locations, which existing systems often fail to capture effectively. Another major limitation is the difficulty in handling the nat- ural variability of language in news reports. Current systems struggle with accurately identifying crime locations and fill- tering out irrelevant data, particularly when it comes to drug crimes. Issues like handling ambiguous terms, eliminating du- plicates, and cleaning noisy data make it challenging to consist- tenthly extract and map drug-specific information. Furthermore, while many systems use open-source GIS tools, they often lack real-time data integration and advanced features like personalized crime searches. This limits their ability to provide real-time, location-specific insights on drug crimes, which law enforcement agencies and public users may find crucial for timely decision-making .Unlike these existing solutions, your approach focuses ex- clusively on drug-related crimes by scraping online news data and using advanced natural language processing techniques like BERT and spaCy. By honing in on drug crimes, your system will fill a crucial gap, offering a more detailed and focused map- ping solution that leverages both structured and unstructured data sources to provide comprehensive and actionable insights.

4. DATASET

Dataset description: The dataset used in this research consists of drug-related crime reports extracted from The Times of India. The data spans from 2021 to 2024 and contains approximately 1,200 records detailing various drug-related incidents. The dataset. Captures key information such as headlines, publication date, article URL, and a brief crime description, enabling a compare- pensive analysis of crime patterns.

Dataset sample: The primary data source is The Times of India, from which crime reports were collected using web scraping techniques with Selenium and Beautiful Soup.

	Headline	Date	URL	Description
2	Alarming drug tr	2024-09-02	https://timesofinc	In hyderabad, multiple drug addicts have
3	Hyderabad drug	2024-08-27	https://timesofinc	Three individuals transporting amphetami
4	Suppliers to ped	2024-09-13	https://timesofinc	In a significant crackdown on drug supplie
5	Rakul preet sing	2024-07-15	https://timesofinc	Actress rakul preet singh's brother aman
6	Drug bust case:	2024-07-16	https://timesofinc	According to officials, aman preet singh is
7	Drug bust in it hu	2024-07-08	https://timesofinc	Telangana anti-narcotics bureau raided 'fo
8	Rs 328 crore nar	2024-07-04	https://timesofinc	The kashimira crime unit has dismantled a
9	Actress hema re	2024-06-15	https://timesofinc	Telugu actress hema has been under cust
10	Hyderabad 'doct	2024-10-27	https://timesofinc	The mumbai crime branch which recently
11	Telugu students	2024-07-29	https://timesofinc	Two students from telangana studying in t
12	'bro daddy' assis	2024-09-12	https://timesofinc	Mansoor rasheed, assistant director of 'br

FIGURE 1. A subset of the cleaned dataset

5. METHODOLOGY

Selenium and Beautiful Soup are used to scrape drug-related crime data from online news sources. Selenium automates browser actions to navigate dynamic websites and load content, while Beautiful Soup parses the HTML to extract relevant information such as article headlines, publication dates, and crime details. Spa CY extracts entities like locations and drug types, while BERT identifies key drug-crime terms, ensuring data is analysis-ready. The collected data is cleaned by filtering out irrelevant con- tent and organizing it into a structured format, categorizing by location, date, drug type, and crime details. This cleaned data is used for visualization. Open Street Map and Leaflet map crime hotspots, allowing users to search locations, view incidents, and analyze drug crime distribution interactively.

Design:

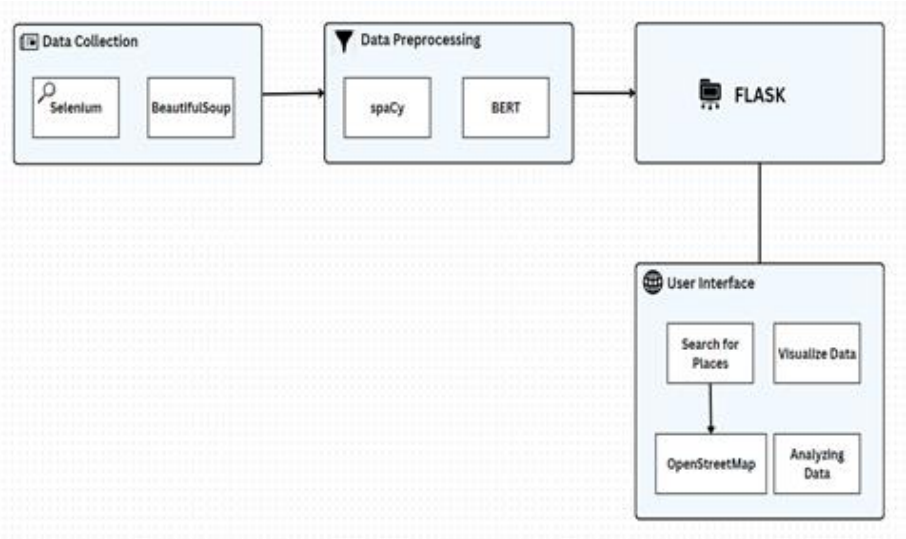

FIGURE 2. Architecture diagram

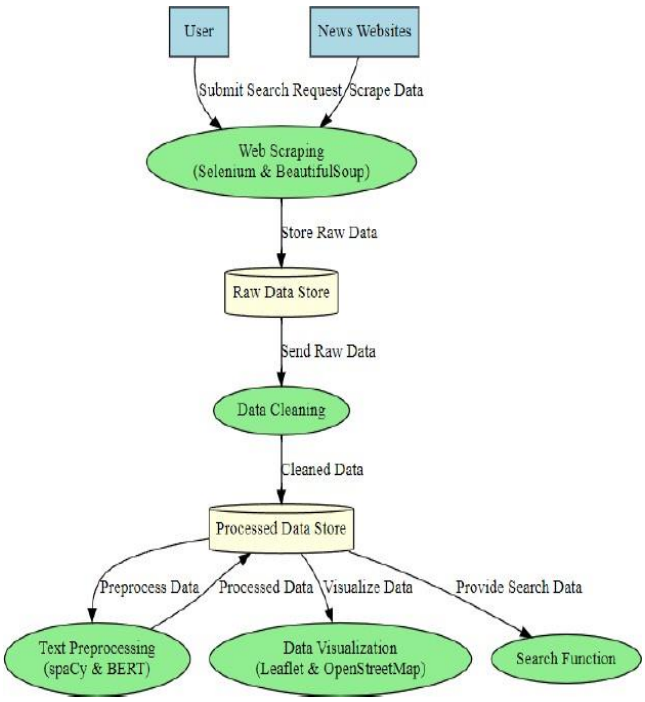

FIGURE 3. Data flow diagram

6. RESULTS AND DISCUSSIONS

Web Scraping and Data Collection: The system success- fully scrapes drug-related crime data from news articles on the Times of India website using Selenium and Beautiful Soup. The scraped data includes titles, dates, and content of articles, which are then stored for further analysis. A validation step was per- formed to confirm the accuracy of the scraped data, ensuring that the content gathered is relevant to the project scope. Text Preprocessing and Entity Detection: The text preprocessing pipeline, powered by spa Cy, successfully tokenized the scraped content and identified entities such as drugs and locations. BERT was employed to detect keywords, achieving significant accuracy in recognizing drug-related terms. During testing, the system demonstrated high precision in extracting the targeted information, such as place names and drug terms, which can be used to map crime hotspots. Geospatial Visualization: The data was mapped onto an interactive interface using Open Street Map and Leaflet. Crime hotspots were visualized successfully, allowing users to zoom in on specific areas and search for relevant crime data. The visualization was tested with sample data, and the results demonstrated the effectiveness of the map interface in providing users with an intuitive way to explore drug crime trends geographically. Search Functionality: The search function allowed users to retrieve crime data by location. This feature was tested using various search terms (city names, district names, etc.), and the system consistently returned relevant data points on the map. The search mechanism proved robust in handling multiple queries and displaying results quickly. Backend Performance: The Flask backend was validated for its ability to handle requests from the frontend. It communicated seamlessly with the database to fetch and serve data. The testing indicated that the system can handle multiple requests efficiently, though further scalability tests will be required when additional data sources are incorporated.

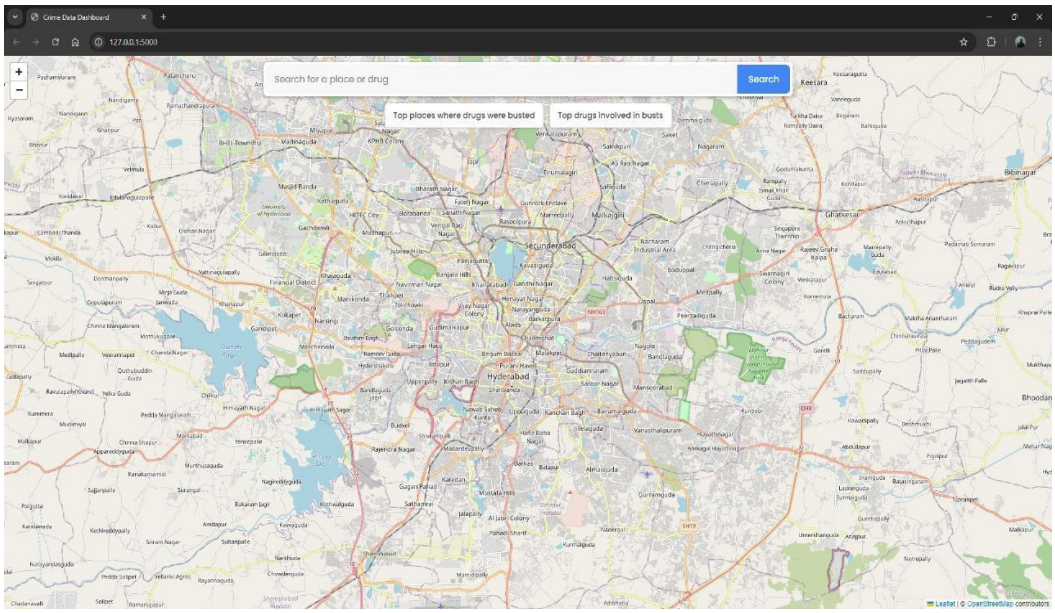

FIGURE 4. Search screen page

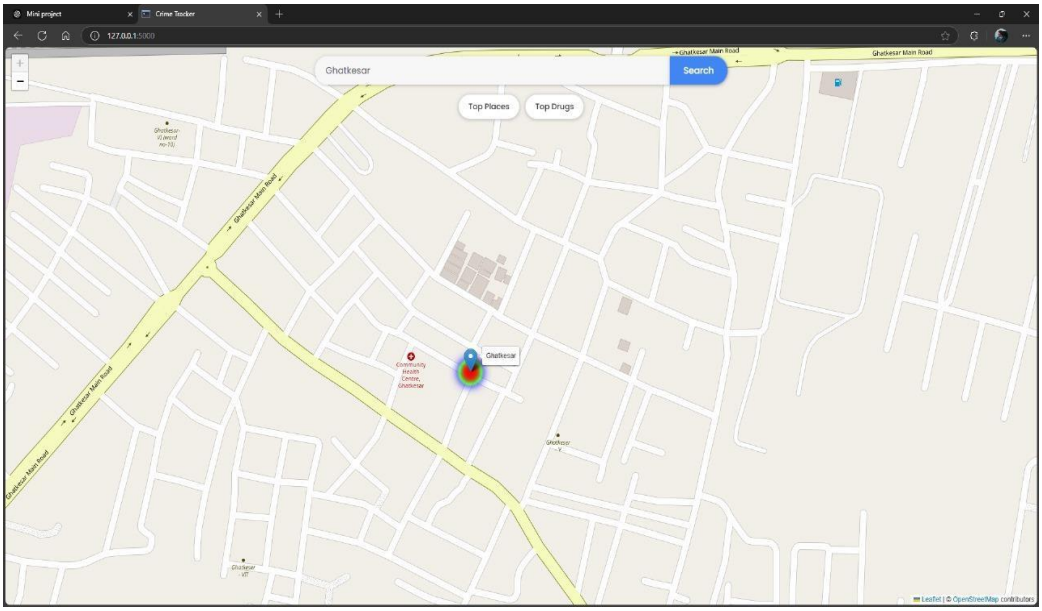

FIGURE 5. Result screen page

7. CONCLUSION

This project addresses the pressing issue of drug-related crimes by providing a robust GIS solution that combines web scraping, data preprocessing, and geospatial visualization. By utilizing advanced technologies like BERT for natural language processing and Open Street Map for mapping, the system not only enhances data accessibility but also empowers users to take proactive steps toward community safety. The findings from this project have the potential to

inform law enforcement strategies and community outreach programs, ultimately contributing to a safer environment for residents in Teagan

8. ACKNOWLEDGMENT

We owe our gratitude to Prof.Archana Mantri, Vice- Chancellor, Anurag University, for extending the University fa- cilities to the successful pursuit of our project so far and her kind patronage. We acknowledge our deep sense of gratitude to Prof. Balaji Utla , Registrar, Anurag University, for be- ing a constant source of inspiration and motivation. We wish to record our profound gratitude Dr. V. Vijay Kumar, Dean – School of Engineering, for his motivation and encouragement. We sincerely thank Dr. A. Mallikarjuna Reddy, Associate Pro- fessor and the Head of the Department of Artificial Intelligence, Anurag University, for all the facilities provided to us in the pursuit of this project. We are indebted to our project guide Dr. M. Trupthi, Associate Professor, Department of Artificial Intelligence, Anurag University. We feel it's a pleasure to be indebted to our guide for her valuable support, advice, and en- couragement and we thank her for superb and constant guid- ance towards this project**.**

REFERENCES

[1]. Guiyun Zhou, Jiayuan Lin and Wenfeng Zheng, "A web-based geograph- ical information system for crime mapping and decision support," 2012 International Conference on Computational Problem-Solving (ICCP), Le- shan, 2012, pp. 147-150, doi: 10.1109/ICCPS.2012.6384228.

[2]. D. V. Rohini and P. Isakki, "Crime analysis and mapping through on- line newspapers: A survey," 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), Kovilpatti, India, 2016, pp. 1-4, doi: 10.1109/ICCTIDE.2016.7725331.

[3]. M. A. Kaif, S. Samaiya, R. A. Purnima, T. Sivasankar, A. Roy and A. Poojitha, "Development of an interactive web-based geovisual analytics platform for analysing crime data," 2023 IEEE Symposium on Wireless Technology & Applications (ISWTA), Kuala Lumpur, Malaysia, 2023, pp. 158-162, doi: 10.1109/ISWTA58588.2023.10249483.

[4]. P. Das and A. K. Das, "Behavioural analysis of crime against women us- ing a graph based clustering approach," 2017 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2017, pp. 1-6, doi: 10.1109/ICCCI.2017.8117714.

[5]. Manoranjan Dash et al.," Detection of Psychological Stability Status Using Machine Learning Algorithms", International Conference on Intelligent Systems and Machine Learning, Springer Nature Switzerland, Pp.44-51, 2022.

[6]. G. Garc´ıa-Zanabria et al., "CriPAV: Street-Level Crime Patterns Analy- sis and Visualization," in IEEE Transactions on Visualization and Com- puter Graphics, vol. 28, no. 12, pp. 4000-4015, 1 Dec. 2022, doi: 10.1109/TVCG.2021.3111146.

[7]. Tobias C.J.B. Mwanza B. (2024). Spatial and Temporal Analy- sis: A GIS-Based Application Tool for Crime Monitoring and Clus- tering in Malawi. European Scientific Journal, ESJ, 20 (8), 167. https://doi.org/10.19044/esj.2024.v20n8p167

[8]. Al-Aboosi A, Sheikh Abdullah S, Ismail R, Abdul Maulud K, Nahar L, Zainol Ariffin K, Lam M, bin Talib M, Wahab S, Elias M A Geospa- tial Drug Abuse Risk Assessment and Monitoring Dashboard Tailored for School Students: Development Study With Requirement Analysis and Acceptance Evaluation JMIR Hum Factors 2024;11:e48139 URL: https://humanfactors.jmir.org/2024/1/e48139 DOI: 10.2196/48139

[9]. Vytla, V., Ramakuri, S. K., Peddi, A., Srinivas, K. K., & Ragav, N. N. (2021, February). Mathematical models for predicting COVID-19 pandemic: a review. In Journal of Physics: Conference Series (Vol. 1797, No. 1, p. 012009). IOP Publishing.

[10]. Aziz, R.M., Hussain, A. Sharma, P. Cognizable crime rate predic- tion and analysis under Indian penal code using deep learning with novel optimization approach. Multimed Tools Appl 83, 22663–22700 (2024).https://doi.org/10.1007/s11042-023-16371-0

[11]. Baradel, M., Breuer, N. (2024). Mapping drug smuggling networks in Japan: a social network analysis of trial documents. Global Crime, 25(3–4), 220–241. https://doi.org/10.1080/17440572.2024.2375241

[12]. Manoranjan Dash et al.," Effective Automated Medical Image Segmentation Using Hybrid Computational Intelligence Technique", Blockchain and IoT Based Smart Healthcare Systems, Bentham Science Publishers, Pp. 174-182,2024

[13]. A.Mallikarjuna, B. Karuna Sree, " Security towards Flooding Attacks in Inter Domain Routing Object using Ad hoc Network" International Journal of Engineering and Advanced Technology (IJEAT), Volume-8 Issue-3, February 2019.

[14]. S. P, A. S M, K. M and S. Chitturi, "Evaluating Crime Type And Crime Pattern Analysis Using Machine Learning," 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India, 2024, pp. 1-6, doi:10.1109/I2CT61223.2024.10543291.

[15]. P. Shobha Rani, R. Dayana. K, R. B. S, P. Kumar S and K. S. Sumanth, "A Web Application to Report Drug Trafficking Crime Anonymously," 2024 5th International Conference on Recent Trends in Computer Sci- ence and Technology (ICRTCST), Jamshedpur, India, 2024, pp. 133-137, doi:10.1109/ICRTCST61793.2024.10578507.

[16]. M. Ozer, I. Onat, H. Akbas, N. Elsayed, Z. ElSayed and S. Varli- oglu, "Exploring the Journey to Drug Overdose: Applying the Journey to Crime Framework to Drug Sales Locations and Overdose Death Lo- cations," 2023 Congress in Computer Science, Computer Engineering, Applied Computing (CSCE), Las Vegas, NV, USA, 2023, pp. 504-510, doi:10.1109/CSCE60160.2023.00088.

[17]. D. Petrou, V. Martinez-Gil, F. Castillo, C. Tunc and R. Bryce, "Twitter Account Analysis for Drug Involvement Detection," 2023 3rd Intelligent Cybersecurity Conference (ICSC), San Antonio, TX, USA, 2023, pp. 9- 16, doi: 10.1109/ICSC60084.2023.10349992.

[18]. Manoranjan Dash, N.D. Londhe, S. Ghosh, et al., "Hybrid Seeker Optimization Algorithm-based Accurate Image Clustering for Automatic Psoriasis Lesion Detection", Artificial Intelligence for Healthcare (Taylor & Francis), 2022, ISBN: 9781003241409

[19]. I. Jayaweera, C. Sajeewa, S. Liyanage, T. Wijewardane, I. Perera and A. Wijayasiri, "Crime analytics: Analysis of crimes through newspa- per articles," 2015 Moratuwa Engineering Research Conference (MER- Con), Moratuwa, Sri Lanka, 2015, pp. 277-282, doi: 10.1109/MER- Con.2015.7112359.

[20]. R. Diouf, E. N. Sarr, O. Sall, B. Birregah, M. Bousso and S. N. Mbaye, "Web Scraping: State-of-the-Art and Areas of Application," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 6040-6042, doi: 10.1109/BigData47090.2019.9005594.

[21]. B. Bhardwaj, S. I. Ahmed, J. Jaiharie, R. Sorabh Dadhich and M. Gane- san, "Web Scraping Using Summarization and Named Entity Recognition (NER)," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 261- 265, doi: 10.1109/ICACCS51430.2021.9441888.

