



SOFTWARE SOLUTION TO IDENTIFY USERS BEHIND TELEGRAM, WHATSAPP AND INSTAGRAM BASED DRUG-TRAFFICKING

Miss. Amrapali B Baviskar¹, Prof. Dr. Dinesh D Puri²

Research Scholar, Department of Computer Applications, SSBT's COET, Jalgaon Maharashtra, India¹

Professor, Department Of Computer Applications, SSBT's COET, Jalgaon Maharashtra, India²

Abstract: Drug trafficking is a serious social and security problem, and social media and instant messaging applications can be misused for illegal drug-related communication. Telegram, WhatsApp and Instagram provide features such as private communication, groups, temporary accounts, coded language, emojis and multiple online identities, which can make manual identification of suspicious activities difficult. This research proposes a software-based decision-support solution for identifying and analyzing potential drug-trafficking indicators using only publicly available or lawfully obtained digital data. The proposed approach combines Natural Language Processing (NLP), Machine Learning (ML), Cybersecurity, data analysis and Blockchain-based integrity mechanisms. NLP techniques can analyze text, keywords, coded terms, captions and comments, while Machine Learning (ML) techniques can be used to classify activities as Normal, Suspicious and High-Risk. Communication-pattern and cross-platform analysis can provide additional indicators of possible relationships among accounts. A risk score can help authorized investigators prioritize cases for further human review. Cybersecurity controls such as authentication, access control, encryption and secure storage are proposed to protect sensitive investigation data. Blockchain and hashing support evidence-integrity verification. The proposed solution is not intended to automatically identify a person as guilty; rather, it provides investigative leads and risk indicators that must be verified by authorized investigators. The research therefore focuses on supporting lawful, secure and human-verified analysis of potential drug-trafficking activities across multiple digital platforms.

Keywords: Drug Trafficking Detection, Social Media Analysis, Natural Language Processing (NLP), Machine Learning (ML), Cybersecurity, Blockchain, Risk Score, Digital Evidence, Telegram, WhatsApp, Instagram.

INTRODUCTION

Social media and instant messaging applications such as Telegram, WhatsApp and Instagram are widely used for communication, information sharing and online interaction. At the same time, these platforms may be misused for illegal activities, including the promotion, distribution and communication associated with drug-trafficking. Private groups, coded language, temporary accounts, multiple identities and rapidly changing online content can make suspicious activity difficult to recognize through manual investigation alone.

The increasing volume of digital communication creates a major challenge for cybersecurity and law-enforcement investigations. Investigators may need to examine large numbers of messages, captions, comments, usernames, timestamps and communication relationships. Manual analysis of such information can be time-consuming and may make it difficult to identify patterns across different sources. A software-assisted approach can help organize lawfully obtained information and highlight indicators that deserve further investigation.

The proposed research focuses on a software solution that combines Natural Language Processing, Machine Learning, cybersecurity techniques and data analysis. NLP can be used to clean and analyze text, identify relevant keywords or coded terms, and extract useful features. The proposed framework can use Machine Learning techniques to classify activities into Normal, Suspicious or High Risk categories.

Communication-pattern analysis can examine repeated interactions, frequency of communication and relationships among accounts. Where legally available and technically possible, relevant indicators may also be compared across Telegram, WhatsApp and Instagram. A central feature of the proposed system is risk-based decision support. Instead of



making an automatic accusation, the system generates a risk score from multiple indicators. Authorized investigators can use this score to prioritize cases and then verify the underlying evidence.

Investigation data should be protected through authentication, encryption, access control and secure storage. Blockchain-based records and hashing can also support the integrity of important investigation records.

Research Objective:

To design a secure software-based decision-support framework that analyzes legally obtained social-media data using NLP, Machine Learning, cybersecurity and evidence-integrity techniques to identify potential indicators of drug-trafficking activity across Telegram, WhatsApp and Instagram.

LITERATURE REVIEW

The literature review focuses on existing research related to drug trafficking and illicit drug-related activities through Instagram, WhatsApp, and Telegram. Previous studies have used Machine Learning, Deep Learning, multimodal analysis, encrypted messaging applications, and Social Network Analysis to identify drug-related content, users, communication patterns, and trafficking networks. These studies provide a foundation for the proposed software solution for analyzing potential users and drug-trafficking indicators across multiple social media and messaging platforms.

Li et al. (2019) proposed a Machine Learning approach for detecting and characterizing illicit drug dealers on Instagram. The researchers collected 12,857 Instagram posts and manually annotated the data to identify drug-dealing-related content. They compared an LSTM-based Deep Learning model with Random Forest, Decision Tree, and Support Vector Machine models. The LSTM model achieved an F1-score of approximately 95% and performed better than the other models. The study demonstrated that Deep Learning techniques can be used to detect illegal drug-selling posts and identify patterns in drug-related conversations. This research is relevant to the proposed system because it demonstrates how Machine Learning and NLP-based techniques can be used to analyse drug-related content and identify potential drug-trafficking indicators on Instagram.

Hu et al. (2021) developed a multimodal approach to identify potential illicit drug dealers on Instagram. They created the IDDIG dataset with nearly 4,000 user accounts, including more than 1,400 drug-dealer accounts. The researchers analysed post comments, images, user biographies, and profile images. Their approach reported an accuracy of approximately 95% and also used hashtags to identify geographical and drug-related patterns. The study showed that combining text, images, and profile information can provide multiple indicators of potential drug-related activities. This research is relevant to the proposed system because it shows the importance of analysing different types of information instead of relying on a single keyword or message.

Moyle et al. (2019) studied how social media and encrypted messaging applications, including WhatsApp and Instagram, are used for drug supply and access. The researchers conducted an international survey of 358 drug users along with interviews to understand digital drug-supply practices. The study found that these applications can provide a quick and convenient way for buyers and sellers to communicate and arrange drug-related transactions. It also showed that messaging applications can connect traditional street-level drug dealing with online drug markets. This research is relevant to the proposed system because it shows that WhatsApp and Instagram can be used as communication channels for drug-related activities.

Barratt et al. (2022) studied Televend, a digital drug market that used darknet infrastructure and Telegram messaging technology. The study examined how Telegram was used for communication and purchasing within the digital drug market. It showed how messaging-app technology can be connected with online drug trading activities. The research is relevant to the proposed system because it demonstrates that Telegram can be used as a communication and transaction platform for illicit drug markets.

Loggen et al. (2026) studied illicit products and services advertised in public Telegram groups. The researchers analysed more than 26 million messages from over 88,000 users across 31 public Dutch Telegram groups over a two-year period. The study identified different types of illicit products and services, with narcotics being the dominant category. The research showed that Telegram public groups can be used for communication related to illicit markets and drug-related



activities. This research is relevant to the proposed system because it highlights the importance of analysing messages, users, and communication patterns when identifying potential drug-trafficking indicators on Telegram.

Bichler, Malm and Cooper (2017) conducted a systematic review of research using Social Network Analysis to study the organizational structure of illicit drug-trafficking networks. The researchers examined how relationships between individuals and groups involved in drug trafficking can be represented and analysed as networks. The study showed that drug trafficking commonly involves interconnected actors rather than isolated individuals. Therefore, analysing relationships and network structures can provide additional information for investigations. This research supports the proposed system's communication-pattern and network-analysis components for identifying potential relationships between suspicious users.

Hughes, Bright and Chalmers (2017) analysed Australian poly-drug trafficking networks using Social Network Analysis. The researchers examined three drug-trafficking syndicates and studied their management structures, relationships, and methods of handling multiple illicit drugs. The study found that different trafficking syndicates can have different organizational structures and approaches to drug distribution. This research demonstrates the importance of analysing relationships, roles, and network structures when studying drug-trafficking activities. It supports the proposed system's use of communication patterns and network relationships as additional indicators for identifying potential drug-trafficking users.

Research Gap:

Existing studies show that Instagram, WhatsApp and Telegram can be used for drug-related communication and online drug activities. Studies have used Machine Learning, Deep Learning, NLP and network analysis to identify suspicious activities. However, most studies focus on only one platform or one type of data.

There is a need for an integrated approach that considers Telegram, WhatsApp and Instagram together. The proposed research combines NLP, Machine Learning, communication-pattern analysis, cross-platform indicators, cybersecurity and evidence integrity. It aims to provide risk indicators and useful information for authorized investigators. The system does not automatically prove a person's identity or guilt.

Approach	Strength	Limitation	Role
Machine Learning	Detects suspicious patterns	Needs training data	User activity detection
Deep Learning	Handles complex patterns	Needs large datasets	Advanced content detection
NLP	Analyses textual data	Coded language difficulty	Message analysis
Multimodal Analysis	Combines multiple data	Processing complexity	Text-image analysis
Social Network Analysis	Finds user relationships	Incomplete network data	Network identification
Digital Forensics	Examines digital evidence	Requires lawful access	Evidence analysis

Table 1. Comparison of Major Approaches

METHODOLOGY

The proposed research follows a system-design and analytical methodology. The study focuses on analyzing legally obtained digital information, identifying relevant indicators, classifying suspicious activities, and providing risk-based information for authorized investigation.



1.1 Data Collection

The research uses publicly available or lawfully obtained digital information from Telegram, WhatsApp and Instagram. The data may include permitted messages, posts, captions, comments, usernames, timestamps and relevant metadata. Unauthorized access to private accounts is excluded.

1.2 Data Pre-processing and NLP

Collected text is cleaned and normalized for analysis. NLP techniques such as tokenization, keyword analysis and TF-IDF feature extraction can be used to identify recurring terms, coded expressions, emojis and suspicious text patterns.

1.3 Machine Learning Classification

Machine Learning can help classify activities into Normal, Suspicious and High-Risk categories. In this research, no actual Machine Learning model is developed or trained. Model performance can be checked using accuracy, precision, recall and F1-score in future work.

1.4 Communication and Cross-Platform Analysis

The research considers repeated interactions, communication frequency and relationships between accounts. Information from Telegram, WhatsApp and Instagram can be compared using legally permitted identifiers. Similar accounts should not automatically be assumed to belong to the same person.

1.5 Risk Score Generation

A risk score can be generated using multiple indicators such as suspicious text, repeated interactions and account relationships. The score can help authorized investigators prioritize cases for further examination. It is not a legal finding.

1.6 Cybersecurity and Evidence Management

Investigation data should be protected using authentication, access control, encryption and secure storage. Access logs can support accountability, while SHA-256 hashing can help verify the integrity of important digital records.

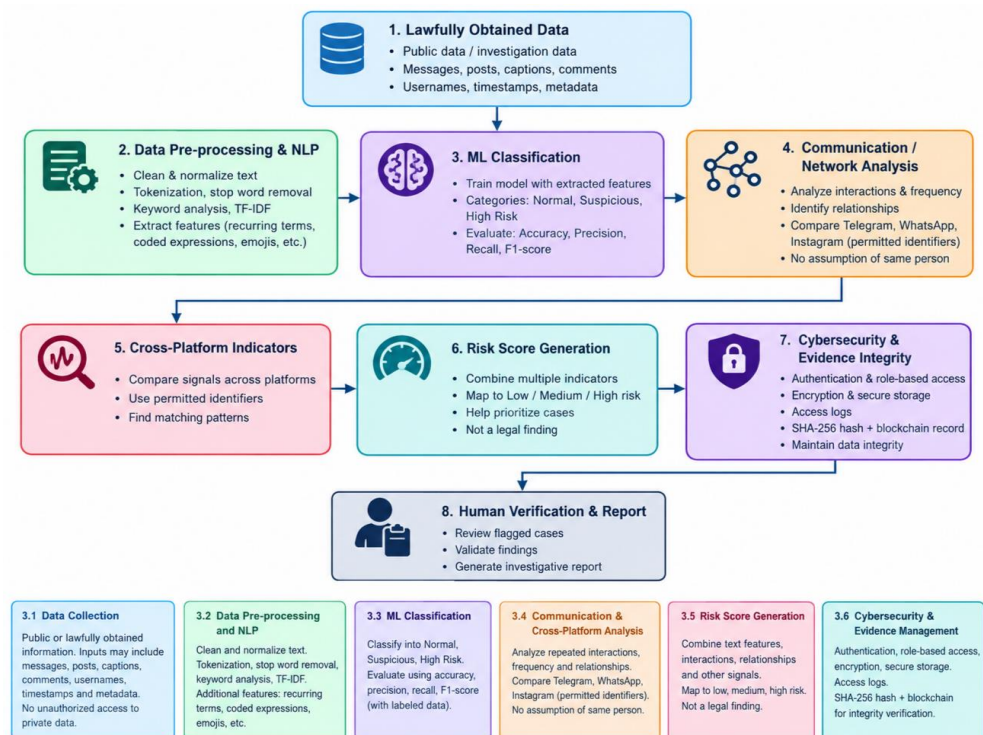


Figure 1. Methodology Workflow Diagram



RESULTS

The proposed methodology is expected to produce a structured and secure analysis framework for identifying potential drug-trafficking indicators. The processed data can be classified into Normal, Suspicious and High-Risk categories. Communication and cross-platform analysis can provide interaction patterns and relationship indicators across Telegram, WhatsApp and Instagram. A risk score can help prioritize suspicious cases, while cybersecurity and evidence-integrity measures can support secure handling of investigation data. The final output can be presented as an investigative report for authorized investigators and requires human verification.

The proposed methodology is expected to produce a structured and secure analysis framework for identifying potential drug-trafficking indicators. The results include processed data, classification of activities, network insights, risk assessment, secure evidence handling and a final investigative report.

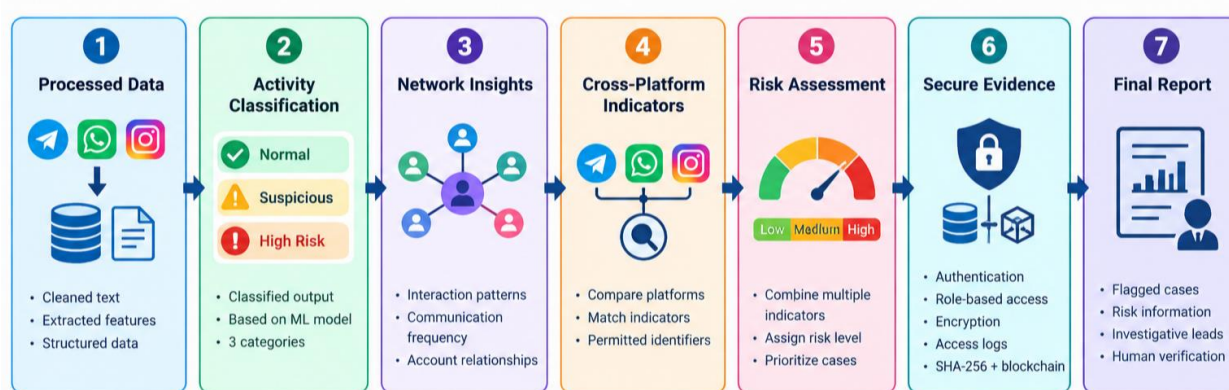


Figure 2. Expected Result From The Methodology

DISCUSSION

The proposed approach combines NLP, Machine Learning, network analysis and cybersecurity techniques to identify suspicious drug-related activities on social media. NLP helps analyze unstructured text, while Machine Learning supports activity classification. Keywords, emojis, communication patterns and account relationships can be used to generate a risk score. However, these indicators may produce false positives and cannot alone prove criminal behavior. Therefore, the risk score should be treated only as an investigative priority indicator and verified by an authorized investigator.

Privacy, lawful access and evidence security are important considerations. The system should use only publicly available or lawfully obtained information with authentication, authorization, encryption and secure storage. Cross-platform identity matching is also challenging because similar usernames, profile names or content patterns do not necessarily prove that accounts belong to the same person. Hence, possible connections should be presented for human verification rather than automatically confirming identity.

CONCLUSION

This research proposes a software system to identify possible drug-trafficking activities on Telegram, WhatsApp and Instagram using lawfully obtained digital information. The system uses NLP, Machine Learning, communication analysis, cybersecurity and Blockchain to analyze suspicious text, keywords, coded messages and account interactions. It can classify activities as Normal, Suspicious and High-Risk and generate a risk score for investigation.

The system is designed to help authorized investigators identify cases that need further checking. A risk score does not prove criminal activity, so human verification is necessary. The system also focuses on privacy, security and lawful use of data. Future implementation and testing with authorized datasets can help evaluate its actual performance.

**REFERENCES**

- [1]. Li, J., Xu, Q., Shah, N., & Mackey, T. K. (2019). A Machine Learning Approach for the Detection and Characterization of Illicit Drug Dealers on Instagram. *Journal of Medical Internet Research*, 21(6), e13803, DOI: 10.1016/j.chb.2026.109032.
- [2]. Hu, C., Yin, M., Liu, B., Li, X., & Ye, Y. (2021). Detection of Illicit Drug Trafficking Events on Instagram. *Proceedings of CIKM 2021*, 3838–3846, DOI: 10.1186/s40163-017-0063.
- [3]. Moyle, L., Childs, A., Coomber, R., & Barratt, M. (2019). Drugsforsale: Social Media and Encrypted Messaging Apps. *International Journal of Drug Policy*, 63, 101–110, DOI: 10.1016/j.socnet.2016.11.004.
- [4]. Barratt, M. J., et al. (2022). Exploring Televend: Cryptomarket and Messaging App Technologies. *Drug and Alcohol Dependence*, 231, 10924, DOI: 10.1016/j.drugalcdep.2021.109243.
- [5]. Loggen, J., et al. (2026). Online Criminal Markets on Telegram. *Computers in Human Behavior*, 182, 109032, DOI: 10.1016/j.chb.2026.109032.
- [6]. Bichler, G., Malm, A., & Cooper, T. (2017). Drug Supply Networks: A Systematic Review. *Crime Science*, 6, 2, DOI: 10.1016/j.expnephrol.2017.02.001.
- [7]. Hughes, C. E., Bright, D. A., & Chalmers, J. (2017). Social Network Analysis of Australian Poly-Drug Trafficking Networks. *Social Networks*, 51, 135–147, DOI: 10.1016/j.socnet.2016.11.04.