



SECURE GRAM: AI-POWERED SECURE MESSAGING PLATFORM FOR REAL TIME DRUG DETECTION AND PREVENTION

Manya B. M.¹, Impana M.P.², Laxmi K.³, Kruthika G.R.⁴

Department of Computer Science and Engineering , GM Institute of Technology
Davangere, 577003, Karnataka, India

ABSTRACT :

The increasing prevalence of digital messaging platform has made difficult to detect drug-related communications, especially slangs, abbreviations, and coded terms. Secure Gram, an AI-powered secure messaging platform intended to enhance online communication safety by detecting drug-related content in real time. Utilizing Natural Language Processing (NLP), Secure Gram identifies potentially concerning messages and includes features such as message detection, real-time communication, an administrative dashboard for monitoring flagged content, and an AI-driven wellness chatbot that offers supportive guidance to users. The platform aims to raise awareness and facilitate early intervention regarding drug-related issues. The research investigates current detection methods and suggests the integration of machine learning techniques to enhance classification accuracy. Evaluation metrics like precision, recall, F1-score, and interface latency will be employed to determine the feasibility of combining automated message detection with wellness support within the messaging framework.

Keywords: Secure gram, Drug-Related Message Detection, Natural Language processing, Machine learning, AI Wellness Chatbot

INTRODUCTION

The rapid evolution of digital communication technologies has markedly altered interpersonal interactions, providing convenience through messaging platforms. However, this shift has raised significant concerns regarding online safety, particularly regarding drug-related discussions and activities. Detecting such communication poses challenges due to users' informal language. Slang, and context-dependent terminology, leading to potential misclassification of harmless messages and overlooking suspicious ones.

Natural Language processing (NLP) techniques are essential for analyzing digital text, involving tasks like text cleaning and tokenization. While rule-based approaches allow identification of predefined keywords, their effectiveness is limited by their rigidity, especially with unfamiliar expressions. In response, Secure Gram has been developed as an AI-powered messaging platform, integrating automated detection of drug-related content. It employs an NLTK-based text-processing system to analyze messages for suspicious patterns and keywords, enabling administrators to review flagged content through an administrators dashboard. The platform also features an AI wellness chatbot to provide users with helpful guidance and information.

The design of Secureram emphasizes the importance of cautions interpretation of detection results, recognizing that a mention of drugs does not inherently signify illegal activity or intent. The system aims to balance effective monitoring with user privacy and the risk of incorrect classifications.

Current mechanisms rely primarily on rule-based detection; however, incorporating machine-learning techniques offers the potential for enhanced performance in classifying complex language patterns. This study aims to evaluate the effectiveness of replacing or augmenting the rule-based system with a machine-learning approach, utilizing metrics such as precision, recall, F1-score, and inference latency to measure success. The results will inform future improvements to Secure Gram, aiming to create a secure, supportive messaging environment that fosters wellness while maintaining robust monitoring capabilities.

II. AIM AND OBJECTIVE

A. Aim

Secure Gram is an AI-driven secure messaging platform that employs Natural Language Processing(NLP) to identify potential drug-related messages in real time. It includes features like secure user authentication, an administrative dashboard for monitoring flagged content, and an AI-powered wellness chatbot that promotes safer digital communication and provides user with supportive guidance.

B. Objectives

The specific objectives are follows: (1) Developing a real-time messaging platform that prioritizes secure user authentication. (2) Implementing text preprocessing using NLTK and a rule-based system (3) Identifying and flagging suspicious messages for further review by administrators (4) Creating an administrative dashboard to monitor flagged messages and assess concerning content (5) Integration an AI powered wellness chatbot to provide supportive guidance and wellness information to users (6) Exploring the use of machine-learning techniques to improve the classification of drug related messages

C. Achievement of Objectives

The Secure Gram prototype effectively combines real-time messaging, user authentication, an administrative dashboard for monitoring flagged content, and an AI-powered wellness chatbot that promotes safer digital communication and provide users with supportive guidance.

III. LITERATURE REVIEW

A. Natural Language Processing for Drug-Related Text Detection

Natural Language Processing(NLP) is essential for detecting drug-related information in digital communications, where informal language and slang complicate automated detection. Research text classification models for identifying nonmedical prescription medication use on social media, comparing traditional machine learning and transformer models.

B. Rule-Based and Keyword-Based Detection Techniques

Rule-Based and Keyword-Based detection is a method for identifying specific patterns in textual data using predefined keywords, phrases, and rules. While simple to implement and effective for detecting drug-related terms, it struggles with unfamiliar slang and contextual nuances.

C. Machine Learning for Drug-Related Message Classification

Machine learning techniques, including Naïve Bayes, Logistic Regression, Support Vector Machines (SVM), and transformer-based models, are effective for classifying drug-related texts by identifying patterns in labelled training data. Assessed various methods for spotting nonmedical prescription medication use on social media.

D. Challenges in Slang, Context, and False-Positive Detection

Detecting harmful content in informal messages is difficult due to contextual-dependent meanings, misspellings, abbreviations, and coded language. Research highlights informal language and drug-related expression variations as key challenges. Keyword-based systems may wrongly flag benign educational or recovery discussions, necessitating a cautious approach where messages are seen as potentially concerning rather than definitive proof of illegality.

E. Secure Messaging and Administrative Monitoring

Secure messaging platforms ensure secure communication through user authentication, message exchange, and access control. Secure Gram, for example uses JWT-based authentication and a MongoDB database to facilitate real-time messaging along with an administrative dashboard for monitoring flagged content through automated detection. While such monitoring aids in identifying concerning messages.

F. AI-Powered Wellness Support and Research Gap

AI-powered conversational systems can offer supportive guidance and accessible information on drug-related concerns. Secure Gram introduces a wellness chatbot to enhance messaging and detection features by connecting automated identification with support services. While previous research has explored NLP and machine learning.

IV. METHODOLOGY

A. System Overview

Secure Gram architecture facilitates user interactions via a messaging interface while employing Natural Language Processing (NLP) techniques for message analysis. Currently, the system utilizes NLTK for text preprocessing and a rule-based detection framework to identify concerning keywords and message patterns. Detected messages can be flagged for review, while the wellness chatbot delivers supportive resources to users.

B. Development Methodology

The development process is iterative and modular, starting with a requirement analysis phase followed by design, frontend and backend development, database integration, and thorough testing. The frontend enables user authentication and messaging, while the backend oversees application logic and database communications. The message detection component is integrated into the messaging workflow to identify content related to drug use. Future advancements will explore machine-learning classification alternatives and evaluate their efficacy against the current rule-based system.

C. System Architecture

Secure Gram operates on five-layer architecture (1) Presentation Layer : Web interface for user registration, login, and messaging. (2) Application Layer : Backend services that handle logic processing and communication. (3) Message Detection Layer: Responsible for text preprocessing and rule-based message identification. (4) Data Layer : Employs MongoDB for storing user and message data, consistent with the scheme. (5) Support and Monitoring Layer : Houses the administrative dashboard for message review and the wellness chatbot.

D. Tools and Technologies

TABLE 1. Tools and Technologies Employed

CATEGORY	TECHNOLOGY	PURPOSE
Frontend	HTML, CSS, JavaScript	User Interface and Interaction
Backend	Node.js, Express.js	Server-side logic and REST API Handling
Database	Mongo DB	Storage and retrieval of application access
Authentication	JSON Web Token	User authentication and protected access
NLP Processing	Python NLTK-based text processing	Text Preprocessing and linguistic processing
Detection	Rule-based keyword and pattern matching	Identification of potentially drug-related messages
Administrative Interface	Web Based dashboard	Monitoring and reviewing flagged messages
AI Support	Wellness Chatbot	Providing supportive guidance to users
Development Tools	Visual Studio Code, Git, Git hub	Code development, version control, and collaboration

E. System Module

Secure Gram includes numerous functional modules: (1) User Authentication Module : Manages user login and registration, employing JWT for protection. (2) Real-Time Messaging Module : Facilitates message exchanges. (3) Text Preprocessing Module : Prepares text using NLTK for detection. (4) Drug-Related Message Detection Module : identifies messages based on keywords. (5) Flagging and Review Module : Flags messages for future monitoring. (6) Administrative Dashboard Module : Displays flagged messages. (7) AI Wellness Chatbot Module : offers user support and guidance. (8) Database Management Module : Controls data handling and retrieval. (9) Machine-Learning Enhancement Module : Planned for future exploration of improved performance through classification models.

F. Data Flow

Users initiate the workflow by logging in and accessing the messaging interface. Upon submitting a message, the backend processes it, utilizing the detection component to apply text preprocessing and evaluate it against keyword criteria. Messages matching the rules are flagged for administrative review. The system also allows interaction with the wellness chatbot for additional support. Future enhancements may incorporate a trained machine-learning classifier to refine detection performance and compare it against the existing rule-based methodologies.

G. System Diagrams

Illustrative diagrams, including a high-level architecture and a message-processing flowchart, depict system components and interactions. They detail user interface actions, backend processing with Node.js, JWT authentication, MongoDB integration, NLP detection components, administrative review processes, and potential future workflows introducing machine-learning classifiers.

V. RESULT AND DISCUSSION

A. Observation

The Secure Gram prototype features a web-based messaging interface with contact lists and an AI Counselling Chatbot for users seeking drug addiction guidance. It includes registration and login interfaces and demonstrates user communication workflows. The message detection component employs NLTK for text preprocessing and rule-based matching.

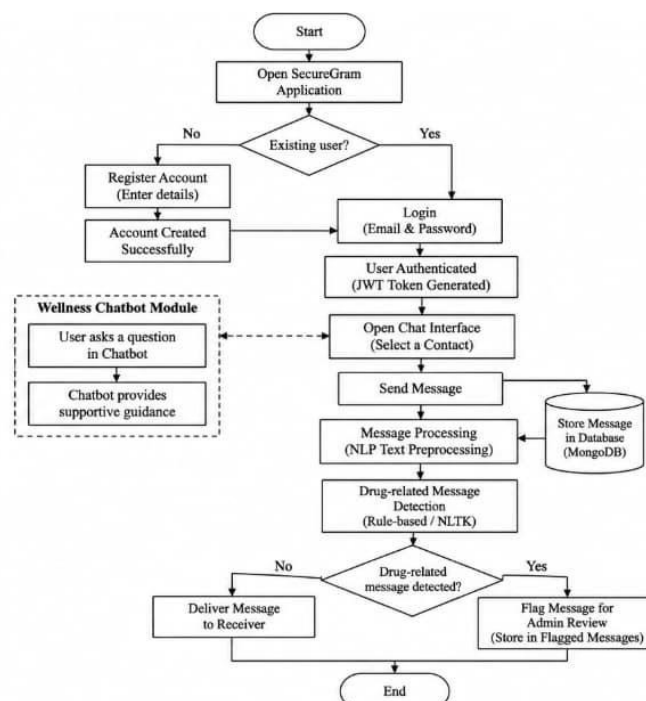
B. Test Results

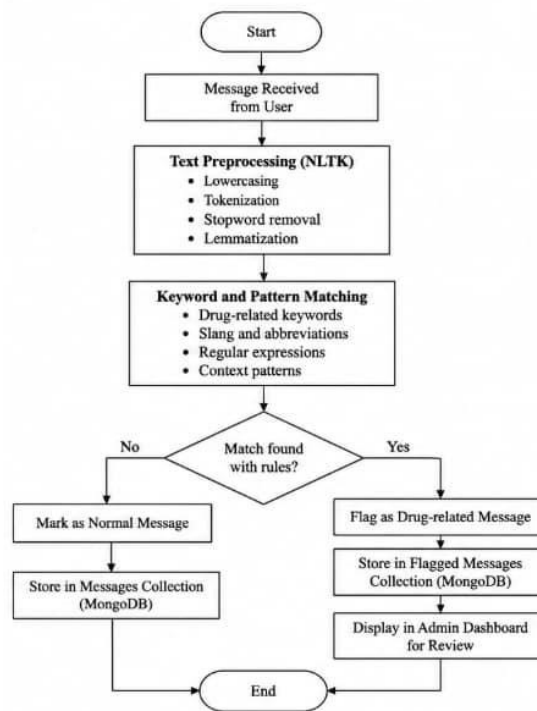
TABLE II. Functional Test case Result

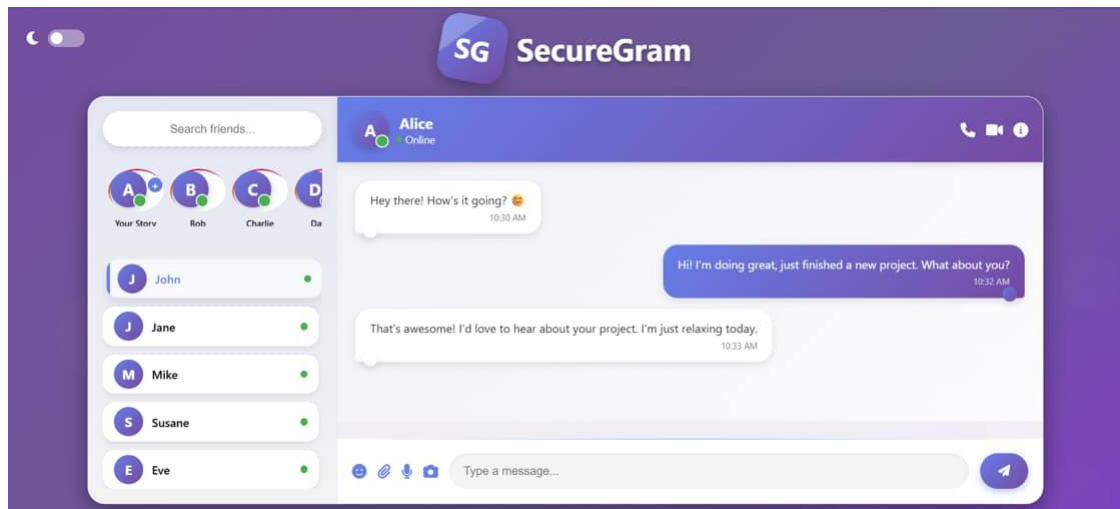
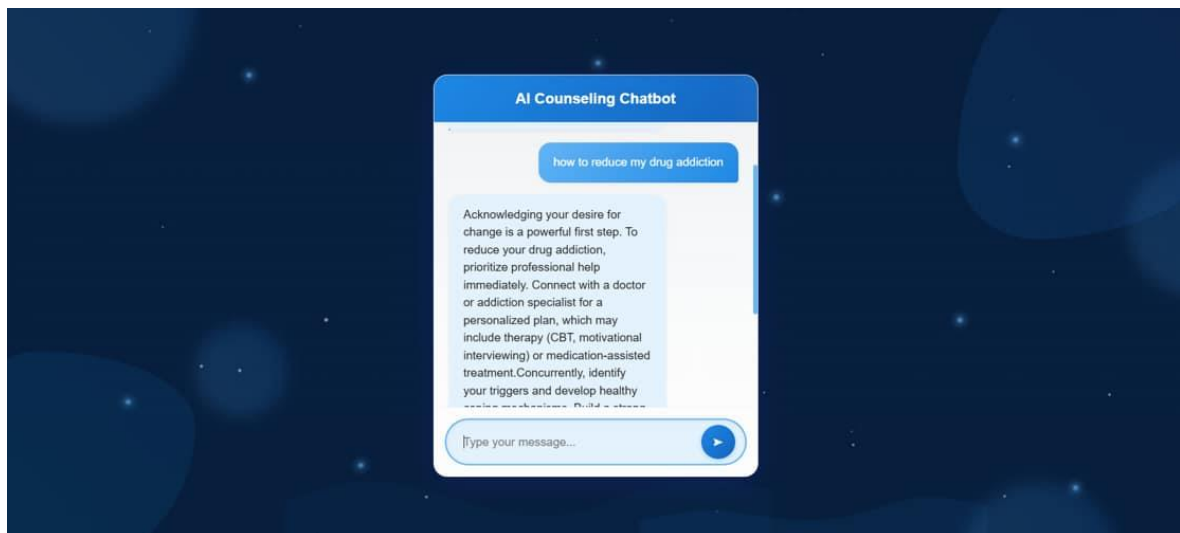
Test ID	Module	Test Scenario	Expected result	Status
01	Registration	Open the register page	Registration form displayed	Interface shown
02	Login	Open Login Page	Login form displayed	Interface shown
03	Messaging	Open the chat interface	Contact list and conversation displayed	Interface shown
04	Message exchange	Send and Receive a message	Messages are exchanged	Verified
05	Text Preprocessing	Process message through NLP	Text is processed	Verified
06	Drug-related detection	Submit message containing a keyword	Matching message is detected according to rules	Verified
07	Flagged message review	Check a message that matches the detection rules	Flagged message is available for review	Verified
08	Administrative Board	Open the admin Panel	Authorized administrator can review flagged messages	Verified
09	Wellness Chatbot	Submit drug-addiction-related question	Chatbot return a supportive response	Verified
10	Error Handling	Submit invalid input	System handles the input appropriately	Verified

FLOW DIAGRAM:

FIG 1: SYSTEM FLOWCHART OF SECURE GRAM



ER DIAGRAM:**FIG 2: MESSAGE DETECTION FLOWCHART****REGISTER PAGE:**
LOGIN PAGE:

MESSAGING PLATFORM:**AI WELLNESS CHATBOT:****C. Discussion :**

The Secure Gram prototype showcases the integration of real-time messaging and automated detection of drug-related content, alongside AI-enabled wellness support. It features a web-based interface with a Node.js and Express.js backend that utilizes MongoDB for data storage and JWT-based authentication for secure access. An NLTK-based text processing and rule-based detection system identifies messages with specific drug-related keywords. An administrative dashboard facilitates the review of flagged messages, while an AI counselling chatbot offers guidance on drug addiction. Nonetheless, the effectiveness of the detection system requires systematic testing and quantitative evaluation to ascertain its accuracy and reliability.

D. Limitations:

Several limitations of the Secure Gram implementation include reliance on predefined keywords, which may overlook unfamiliar slang and context-dependent language. Legitimate discussions can be wrongly flagged due to drug-related terms. The system lacks comparative performance data between rule-based and machine-learning detection. Detection quality is contingent on the maintenance of configured rules. The chatbot's responses may not always offer suitable guidance and should not substitute for professional support.

E. Future Enhancements:

Secure Gram include integrating machine-learning models like TF-IDF with Logistic Regression creating a labelled dataset for model training, improving contextual understanding of slang and misspellings, and evaluating performance metrics such as precision and recall. Additional improvements involve enhancing the administrative dashboard, upgrading the administrative dashboard, upgrading the wellness chatbot, strengthening privacy mechanisms, and conducting performance testing for scalability.

VI. AI-Based Drug-Related Message Detection and Wellness Integration

A. NLP-Based Drug-Related Message Detection

Natural language processing(NLP) allows computer systems to Analyze human language. In Secure Gram, NLP is applied for preprocessing messages to identify drug-related content using the Natural language Toolkit(NLTK). This prototype screens messages by utilizing predefined keywords and patterns. However, the performance can be hindered by variations in slang, Abbreviations , misspellings, and context, making the current rule-based implementation a preliminary step towards more sophisticated text classification methods.

B. Rule-Based Detection and Machine-Learning Approaches

Rule-based detection uses predefined keywords and patterns to identify messages, offering advantages like simplicity and transparency. However, it can struggle with unfamiliar expressions and may misidentify begin discussions with drug-related terminology. In contrast, machine-learning methods, which learn from labelled examples, may improve contextual analysis if provided with proper training data. Secure Gram suggests exploring models like TF-IDF with logistic regression or SVM to enhance detection. A assess if these methods outperform the rule-based baseline.

C. AI Wellness Chatbot Integration in Secure Gram

Secure gram incorporates an AI counselling chatbot designed addiction . This feature complements existing messaging and detection tools, linking troubling digital communications with accessible wellness information. While the chatbot offers supportive guidance, it is not a substitute for professional healthcare or addiction services, and its responses must be carefully assessed for relevance, clarity, safety, and appropriateness, especially regarding serious health concerns.

D. Benefits and Practical Considerations

The Integration of messaging, automated content screening, administrative monitoring, and chatbot-based support provides a unified approach to digital communication safety. Automated screening can help identify messages that may warrant further review, while the administrative dashboard supports monitoring of flagged content. The wellness chatbot adds a supportive interaction channel for users seeking guidance. Nevertheless, the practical effectiveness of Secure Gram depends on detection quality, appropriate handling of false positives, user Privacy , and secure access to sensitive information. Future evaluation should measure precision, recall, F1-score, false-positive rate, and inference latency. These results will help establish the effectiveness and limitations of the proposed system without assuming that automated detection can determine a user's intent.

VII. CONCLUSION

This paper presents Secure Gram, an AI-powered messaging platform designed to identify drug-related messages and offer wellness guidance. It features user registration, JWT authentication, real-time messaging, Mongo DB for data management, and an AI counselling chatbot. While the current system uses rule-based message detection, it may face challenges with slang and contextual meanings, leading to false positives. Future work will explore machine learning for better accuracy, focusing on precision, recall, F1- Score , and scalability.

Secure Gram aims to enhances communication and facilitate digital interventions for wellness. Further enhancements will target contextual comprehension, safety in chatbot responses, privacy measures, and scalability of the system. Secure Gram ultimately represents an initial step towards integrating real-time messaging with automated detection of drug-related content and supportive digital intervention on a unified platform.

REFERENCES

- [1] Ahmad M, Waqas M, Matyrshin I, and Sidorov G. "Leveraging large language" for Multi-Class and Multi-Label Detection of Drug Use and Overdose Symptoms on Social Media, 2025
- [2] Zhang X, "Developing Large language Models to Detect adverse Drug Events in posts on X", Journal of the American Medical Informatics Association , 2024
- [3] Devlin, J., Chang, M. W., Lee, K., and Toutanova, K., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proceedings of NAACL-HLT, , 2019.
- [4] Raj S, Tiwari A, "Bert" Based Transformers lead the way in extraction of Health Information from Social Media, 2021
- [5] Portelli, B, Passabi of Bio NLP Workshop, "Improving Adverse Drug Event Extraction with Span BERT on different text typologies", 2021
- [6] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V., "ROBERT a: A Robustly Optimized BERT Pretraining Approach," ar Xiv preprint arXiv:1907.11692, 2019.
- [7] Weissenbacher, D., et al., "Text Classification Models for the Automatic Detection of Nonmedical Prescription Medication Use from Social Media," Journal of the American Medical Informatics Association (JAMIA), vol. 28, no. 3, pp. 512–520, 2021.
- [8] D.Weissenbacher, "Text Classification Models for the Automatic Detection of Nonmedical Prescription Medication Use from Social Media", Journal of American medical Informatics Association, 2021
- [9] A. Fisher, M. Maclaren young, D. payer, K. Pacheco, C. Dubeau, and V. Mago, "Automating Detection of Drug-related harms on Social Media: Machine Learning Framework", Journal of Medical Internet Research
- [10] "Monitoring opioid-related Social Media chatter using natural language processing and large language models:" Temporal Analysis, "Journal of medical Internet research", 2025