isrjournals

International Journal of Advanced Research in Computer Science Engineering and Information Technology

Volume: 6, Issue: 3,Special Issue: 3 ,Apr,2025 ,ISSN_NO: 2321-3337

AUTOMATED MAPPING OF DISASTER EVENT ACROSS LOCATION USING SOCIAL MEDIA A HYBRID MACHINE LEARNING APPROACH

Priyadharashini S[1], Swathy R[2], Revathi M[3],Varshadevi S[4]*

[123] UG Scholar –Dept CSE, GRT Institute of Engineering and Technology Tiruttani , India.

[4]*Assistant Professor- Dept CSE, GRT Institute of Engineering and Technology Tiruttani ,India.

Priya2004madhi@gmail.com,swathy1883@gmail.com,revathi0112m@gmail.com

[4*]Corresponding Author: varshamiyav13@gmail.com

Abstract **-** In Automated system mapping of disaster event across location using social media platform a Spam Comment Detection and User Blocking System for a social media web application, designed to enhance user experience and maintain a secure online environment. Users can register login and send friend requests, chat, and post text or images, which may receive likes, dislikes, and comments. This project employs an advanced classifier algorithm to detect and filter negative or spam comments in both the chat and post sections. If a user exceeds 05 spam attempts, their IP address is blocked, preventing further access to the platform. Users can also create and share local events, which are visible to other users. The admin has oversight capabilities, including viewing user activity, managing events, and monitoring time spent on the platform through graphical analysis and also intervene by sending warnings to users displaying addictive behavior. Integrates HAM algorithms and Bloom Filter data structures to improve spam detection efficiency and ensure optimal performance.

Keywords: Spam comment detection,User blocking, Advanced Classifier algorithm, Bloom filter data structure,The HAM Algorithm,Secure online Environment,Addictive behavior,Graphical Analysis, IP Address Blocking.

1.INTRODUCTION

As of April 2024, approximately 5.07 billion individuals globally are active on social media platforms. 3-7% of social media users engage in toxic or harmful behavior. over 210 million people, exhibit addictive behaviors related to social media usage. Some of the most popular social media platforms include Facebook, Instagram, Twitter (X), LinkedIn, TikTok, Snapchat, and YouTube, each offering unique features tailored to different types of users. Also removes billions of spam accounts from social media usage. While social media platforms offer various benefits, they also face challenges such as spam, cyber bullying, data privacy concerns, misinformation, an addictive usage patterns. To address these, this project are integrating classifier algorithm to automately detect the spam comments mechanisms to ensure a safer and more engaging experience. The remainder of the paper is structured as follows: Section 2 reviews the related literature and current solutions in the domain. Section 3 outlines the Methodology. Section 4 elaborates on the implementation and Proposed system. Section 5 presents performance Modules. Finally, Section 6 implements the outcome , Section 7 concludes the paper with potential directions for future enhancements.

2. RELATED WORK :

The increasing number of fake social media profiles , stress detection, Cyberbullying detection models and e-scams highlights the need for advanced detection methods using deep-transfer learning models. Spam comment detection research has introduced innovative techniques like emoji-text features and stacked post-comment pairs to enhance performance. Overall, these studies helps in emphasize the vital role of AI and machine learning in improving digital safety and user experience across various social media platforms.

2.1 SOCIAL MEDIA ANALYSIS FOR FAKE-PROFILE IDENTIFICATION[1]

The rise in e-scams, linked to approximately 30% of fake social media accounts, underscores the urgent need for effective detection methods. This study presents an **advanced deep-transfer learning** model designed to identify fraudulent profiles by analyzing diverse social media data types, including posts, likes, comments, multimedia content, user activity, and login behaviors. Each data type is processed separately to uncover

suspicious patterns, such as inconsistencies in user demographics**.** Audio signals are analyzed using various transforms (1D Fourier, Cosine, Convolutional, Gabor, and Wavelet), while image and video data are processed with 2D techniques. Text data is transformed through Word2Vec to enhance the capabilities of a **binary Convolutional Neural Network (bCNN)** in distinguishing between genuine and fake profiles.

2.2 DETECTION AND ANALYSIS OF STRESS-RELATED POSTS[2]

This study focuses on detecting and analyzing stress-related posts in **Reddit's academic communities**. Due to online education and remote work, these communities have become central for academic discussions and support. We classify text as stressed or not using **natural language processing and machine learning classifiers**, with Dreaddit as our training dataset containing labeled Reddit data.

Next, we collect and analyze posts from various academic subreddits. We identified that the most effective individual feature for stress detection is the **Bag of Words**, paired with the Logistic Regression classifier

2.3 CYBERBULLYING DETECTION AND SEVERITY DETERMINATION MODEL[3]

Many teenagers are unaware of the risks posed by cyberbullying, which can include **depression, self-harm, and suicide**. Because of the serious harm it can cause to a person's mental health, cyberbullying is an important problem that needs to be addressed. This research aimed to develop a technique to identify the severity of bullying using a deep learning algorithm and **fuzzy logic**.

which means a post having only image or video are not a part of this research, combining the image with text has been found in cyberbullying posts.
However, this study is limited to text oriented cyberbullying detection. The comments embedded by Keras were fed into a long short-term memory .In this task, Twitter data (47,733 comments) from Kaggle were processed and analyzed to flag cyberbullying comments. The comments embedded by Keras were fed into a long short-term memory network, for classification,the fuzzy logic was applied to determine the severity of the comments. Fuzzy rule sets are designed to specify the strength of different types of bullying.

2.4 ENHANCING HATE SPEECH DETECTION IN THE DIGITAL AGE[4]

This surge has also amplified the rapid spread of hate speech, prompting extensive research efforts for effective mitigation. These efforts have prominently featured advanced natural language processing techniques, particularly emphasizing deep learning methods that have shown promising outcomes.

This article presents a novel approach to address this pressing issue, combining a comprehensive **dataset of 18 sources.** There were two models utilized to address the diversity in the data and leverage distinct strengths found within deep learning frameworks: CNN and BiLSTM with an attention mechanism. These models were tailored to handle specific subsets of the data, allowing for a more targeted approach. The unique outputs from both models were then fused into a unified model.

2.5 SPAM COMMENT DETECTION ON SOCIAL MEDIA AND POST-COMMENT PAIRS APPROACH[5]

This research used the SpamID-Pair dataset derived from social media for Indonesian spam comment detection. After a comprehensive investigation, the emoji-text feature, the stacked post-comment pairs, and ensemble voting could **boost detection performance (in terms of accuracy and F1)**. Based on the experiment, the best stand-alone methods for spam comment detection are the SVM (RBF kernel) and the soft voting ensemble method for the best average performance.

This research employed two strategies to address issues in text spam detection on social media. The **first strategy was utilizing emojis** that had been frequently discarded in many studies. In fact, many social media users use emojis to convey their intentions. The **second strategy was utilizing stacked post-comment pairs**, which was different from many spam detection systems that solely focused on comment-only data.The post-comment pairs were required to detect whether a comment was relevant (not spam) or spam irrelevant to the post context. This research used the SpamID-Pair

dataset derived from social media for Indonesian spam comment detection. After a comprehensive investigation, the emoji-text feature, the stacked post-comment pairs and ensemble voting could boost detection.

3. METHODOLOGY

3.1 User Interaction Module

User Registration and Authentication: Users create accounts and log in securely using standard authentication protocols. **Social Features** :Users can send/accept friend requests, engage in chats, post text/images, and comment on posts. **Event Management**: Users can create and share local events that are visible to their network. **Content Engagement:** Users can like, dislike, and report posts or comments, allowing for dynamic interaction

3.1 Spam Detection System

Data Input: Real-time monitoring of user-generated content (posts, comments, and chats). **Feature Extraction**: Extract linguistic, behavioral, and temporal features from the content. **Spam Classification:** Employ a Hybrid HAM (Hierarchical Attention Mechanism) based Machine Learning model integrated with Bloom Filter data structures to: Detect spam or abusive content. Classify comments and chat messages as spam or non-spam. **Threshold Mechanism**: If a user crosses five spam attempts, their IP address is automatically blocked.

3.2 Admin Monitoring and Intervention

User Activity Visualization:The system tracks and graphically analyzes:Time spent by users.Number of posts, chats, and events created. **Addictive Behavior Detection**:Using activity patterns, the system flags potential addiction. **Admin Actions**:The admin can View flagged users.Send personalized warnings.Block users or events violating guidelines.

3.3 System Performance Enhancement

Efficiency via Bloom Filter: Integrate **Bloom Filter** structures to improve spam detection speed and memory efficiency, ensuring low false positives **Threshold Mechanism**: If a user crosses five spam attempts, their IP address is automatically blocked.

3.4 Security and Privacy Management

IP Blocking:Maintain a blacklist of IP addresses exceeding spam thresholds. **Data Privacy**:Ensure encrypted storage of sensitive user information. While the system blocks malicious users based on IP addresses, users' IPs are stored in an encrypted format. Users are given control over their data, including the ability to request account deletion, download their data, and manage their privacy settings

4. PROPOSED SYSTEM

This proposed system aims to enhance spam detection and user blocking within a social media web application by implementing an advanced classifier algorithm to identify and filter harmful or spam comments. The HAM Algorithm recognize the pattern of application.The Bloom Filter data structure used to filter bad comments then Blocks the user account who attempts the spam comment, so it reduces the unseen dangers of depression, self-harm and suicide. It uses HAM algorithms and the Bloom Filter data structure to optimize detection efficiency,It will ensure scalability, real-time performance, and a safe user environment by minimizing harmful content. Developed a Event management category in access with location helps in developing the knowledge with different domains.Additionally, it will promote responsible user behavior through automated interventions. The system will be continuously updated to stay ahead of evolving spam techniques.

Advantages: The system detects and filters harmful or spammy comments in both posts and chats instantly, ensuring a safer and more enjoyable user experience. By blocking users' IP addresses after a set number of spam attempts, the system prevents repeat offenders from disrupting the platform..

5. MODULES-SPLIT UP

The Module is systematically divided into multiple functional modules to enhance user experience and platform security, each associated with a specific

category to ensure structured organization and clarity with ease in functionality.

5.1 Friend Request and Interaction Module:

This module enables users to send and receive friend requests. Once a request is accepted, users can view each other's profiles and engage in private chats. It also allows users to like, dislike, and comment on posts. This module fosters social interaction and engagement within the platform**.**

5.2 Post and Comment Management Module:

This module enables users to send and receive friend requests. Once a request is accepted, users can view each other's profiles and engage in private chats. It also allows users to like, dislike, and comment on posts. This module fosters social interaction and engagement within the platform.

5.3 Spam Detection and Blocking Module:

This module employs classifier algorithms to detect negative or spam comments in both posts and chat sections. It filters out harmful content in real-time and tracks users who repeatedly attempt to post spam. Users who exceed a set limit of spam attempts will have their IP address blocked.

5.4 Admin Management Module:

This module allows admins to monitor and manage user activities, including approving user registrations and monitoring uploaded events. Admins can view graphical reports on platform usage and user behavior. The module also includes functionality for blocking users who repeatedly violate platform rules and managing flagged content.

5.5 Event Management Module:

This module allows users to create and share events happening at their location. Events can be viewed by other users on the platform, promoting engagement and community participation. Admins can manage the events, ensuring the content complies with platform guidelines. It supports event details like dates, times, and location.

6.EXPERIMENTAL RESULTS

This result discusses about the implementation of the User sign page and user activity, Admin management and Event management with responsive User activity.

- .The user sign page performs the login credentials with users details.
- The Admin management handles the user activity and monitors.
- The Event management provides the information of any event from any location.

6.1 SIGNUP PAGE

In this **User Sign Up** page, users create a new account by selecting their **user type** and filling in their **personal** and **address** details also **upload a file** (such as a profile photo).

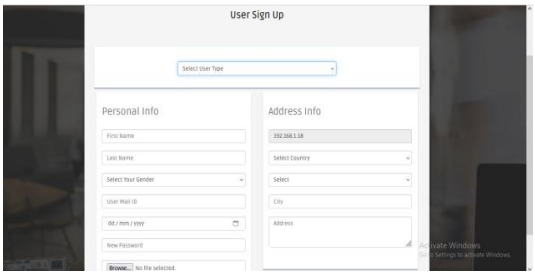

Fig 6.1 shows User Signup page

They enter information like **name, gender, email, birth date, and password**, and provide their **country, state, city**, and **address** users can **submit it** to register on the platform.

6.2 USER LOGIN PAGE

This is a **User Login** page,Users are asked to **sign in** by entering their **email ID** and **password.**If users don't have an account, there's a link to **create a new account**. The **login form** ensures that only **registered users** can access the system helping to maintain **security and personalied Access**.

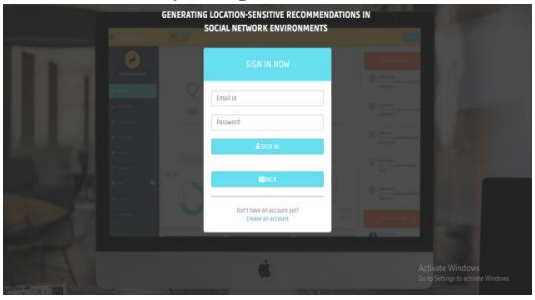

Fig 6.2 shows User Login page

They enter information like **name, gender, email, birth date, and password**, and provide their **country, state, city**, and **address** users can **submit it** to register on the platform.

isrjournals

International Journal of Advanced Research in Computer Science Engineering and Information Technology

Volume: 6, Issue: 3,Special Issue: 3 ,Apr,2025 ,ISSN_NO: 2321-3337

6.3 USER HOME PAGE

This page allows users to interact with posts by **liking, commenting,** or **reporting** them. It provides **navigation** to features like Home, Friends,Location, Events, and Recommendations.A friend list is shown on the right, promoting **social connectivity** within the network.

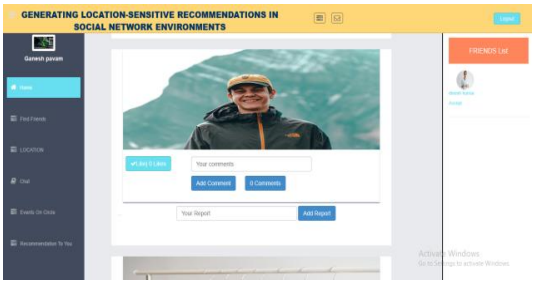

Fig 6.3 shows User Home page

Users can enter **comments or report** inappropriate content directly from the interface.The left sidebar provides navigation to features like Home, Friends, Location, Events, and Recommendations.

6.4 USER FIND FRIENDS PAGE

It displays a list of user profiles with their images, names, and an option to **add them as friends**.Each user card includes a visible "Add Friend" button for sending friend requests.

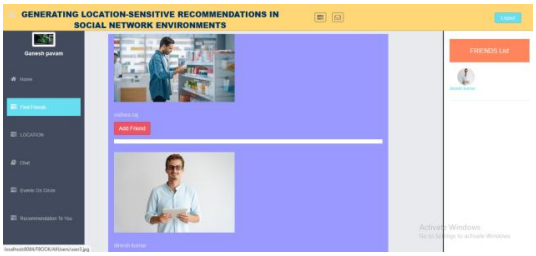

Fig 6.4 shows User Find friends page

The layout allows users to scroll through potential connections and expand their social network. Each user card includes a visible "Add Friend" button for sending friend requests.

6.5 USER CHAT WITH FRIENDS PAGE

Enabling real-time messaging between users.Users can type and send messages using the input field and "Send" button below the chat area. At the top, there is a dropdown menu to **select a user** to chat with. The chat window displays a conversation thread with sent and received messages styled distinctly. Enhances **direct communication** within the social network, promoting user interaction.

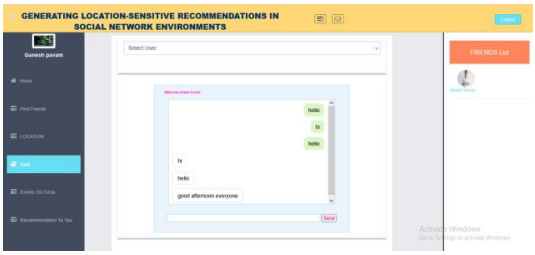

Fig 6.5 shows User Chat with Friends page

chat activities are monitored. Any form of spam or inappropriate comments will be traced, and the associated IP address may be blocked by the system administrator to ensure a safe and respectful environment for all users."

6.6 USER DETAILS PAGE

This page appears to be part of the **Admin Panel** for managing user activity within the application.
It includes sections for **Reported Users**, displaying post ID, reporting user, and the comment made.
This page helps administrators **moderate content, track user behavior**, and manage the platform's safety.

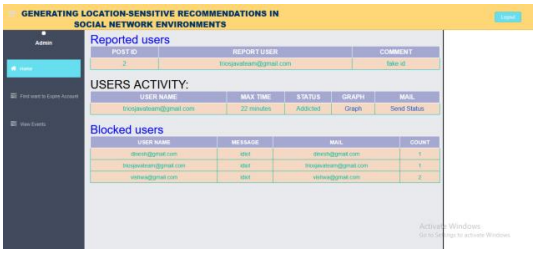

Fig 6.6 shows User Details page

The **Users Activity** section shows login details like email, time spent, and activity status for monitoring purposes.There's also a **Blocked Users** list that includes user details, blocked message, and block count.

6.7 USER VIEW EVENT PAGE

This page is part of the **"Events On Circle"** feature in the application, allowing users to **create and share events**.Users can input event details such as the **event name**, **date/time**, and optionally upload an image or file. This feature encourages users to **share local events** and invite others within

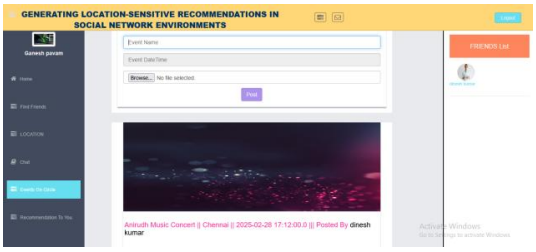

Fig 6.7 shows User View Event page

This feature encourages users to **share local events** and invite others within their network to participate. The showcased event example is a **music concert** in Chennai, posted by a user named Dinesh Kumar.

6.8 USERS IP BLOCKED PAGE

This page is an **Access Denied** screen displayed to users whose IP address has been **blocked by the administrator**.It features a **bold warning message** on a red background, indicating **restricted access** for security or policy reasons.The message clearly states, "Sorry, your IP has been blocked by the administrator." It serves as part of the application's **security and moderation system** to ensure safe usage.

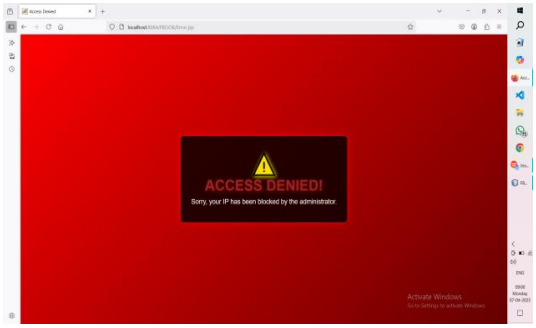

Fig 6.8 shows Users IP blocked page

7. CONCLUSION & FUTURE WORK

In conclusion, the proposed system offers a comprehensive solution to detect and prevent spam and negative comments within a social media platform. By utilizing advanced machine learning algorithms and real-time monitoring, it effectively filters harmful content in posts and chats, ensuring a safer and more positive user experience. The automated process of tracking user activity and blocking IP addresses after repeated offenses helps maintain the integrity of the platform. Additionally, the admin interface empowers platform administrators to oversee user interactions and content, while promoting community engagement through event sharing. This system enhances the platform's overall security, scalability, and user satisfaction, providing a proactive approach to managing harmful behavior online.

FUTURE WORK Future enhancements can incorporate deep learning models like BERT or GPT for improved spam detection and real-time sentiment analysis to differentiate between criticism and harmful content. An adaptive blocking mechanism can replace the fixed 10-attempt threshold Strengthening security with CAPTCHA, behavioral biometrics, and two-factor authentication will prevent automated spam attacks. Blockchain-based moderation can ensure transparency, while cross-platform spam detection can enhance accuracy. Advanced user analytics can track excessive usage, promoting healthier digital habits and a safer online environment.

REFERENCE

[1] Bhrugumalla L.V.S. Aditya And Sachi Nandan Mohanty. Heterogenous Social Media Analysis for Efficient Deep Learning Fake-Profile Identification. VOLUME 12,pp. 99339 - 99351, 08 September 2023.

[2] NAZZERE ORYNGOZHA, PAKIZAR SHAMOI AND AYAN IGALI. Detection and Analysis of Stress-Related Posts in Reddit's Acamedic Communities. VOLUME 12, pp. 14932 - 14948, 23 January 2024.

[3] MOHAMMED HUSSEIN OBAID, SHAWKAT KAMAL GUIRGUIS AND SALEH MESBAH ELKAFFAS. Cyberbullying Detection and Severity Determination Model. VOLUME 11,pp. 97391 – 97399, 07 September 2023.

[4] WAQAS SHARIF, SAIMA ABDULLAH, SAMAN IFTIKHAR, DANIAH AL-MADANI, AND SHAHZAD MUMTAZ. Enhancing Hate Speech Detection in the Digital Age: A Novel Model Fusion Approach Leveraging a Comprehensive Dataset. VOLUME 12,pp. 27225 – 27236, 19 February 2024.

[5] ANTONIUS RACHMAT CHRISMANTO, ANNY KARTIKA SARI AND YOHANES SUYANTO. Enhancing Spam Comment Detection

on Social Media With Emoji Feature and Post-Comment Pairs Approach Using Ensemble Methods of Machine Learning. VOLUME 11,pp.80246 – 80265, 28 July 2023.

[6] R. Shounak, S. Roy, V. Kumar, and V. Tiwari, ‘‘Reddit comment toxicity score prediction through BERT via transformer based architecture,’’ in Proc. IEEE 13th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON), Oct. 2022, pp. 0353–0358.

[7] A. Triantafyllopoulos, S. Zänkert, A. Baird, J. Konzok, B. M. Kudielka, and B. W. Schuller, ‘‘Insights on modelling physiological, appraisal, and affective indicators of stress using audio features,’’ in Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2022, pp. 2619–2622.

[8] B. A. H. Murshed, J. Abawajy, S. Mallappa, M. A. N. Saif, and H. D. E. Al-Ariki, ‘‘DEA-RNN: A hybrid deep learning approach for cyber bullying detection in Twitter social media platform,’’ IEEE Access, vol. 10, pp. 25857–25871, 2022.

[9] M. Mozafari, R. Farahbakhsh, and N. Crespi, ‘‘Cross-lingual few-shot hate speech and offensive language detection using meta learning,’’ IEEE Access, vol. 10, pp. 14880–14896, 2022.

[10] K. T. Mursi, M. D. Alahmadi, F. S. Alsubaei, and A. S. Alghamdi, ‘‘Detecting Islamic radicalism Arabic tweets using natural language processing,’’IEEE Access, vol. 10, pp. 72526–72534, 2022.