IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Multilayer Perceptron Approach For Crime Detection In Social Media

1 Viriyala Jaya 2 Ch Anusha 3 V Anil Santhosh

1 M.Tech Scholar, Department of CSE, International School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India, jayaviriyala90@gmail.com.

2 Associate professor, Department of CSE, International School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India.

3 Associate Professor and HoD, Department of CSE, International School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, lndia.

ABSTRACT: Criminals have been increasingly utilizing Social Media Platforms (SMPs) to achieve a number of criminal goals. These goals can range from the establishment of criminal virtual groups to the sharing of user information and data breaches. Such criminal groups often use SMPs to share confidential information and coordinate criminal activities. The ease of access and the ability to remain anonymous makes SMPs a powerful tool for criminals. Data breaches are a major concern when it comes to SMPs. Criminals can use SMPs to gain access to personal and financial information, including credit card numbers and passwords. They can also use SMPs to spread malicious software and malware, which can be used to infect computers and mobile devices. The sharing of user information is also a concern for SMPs. Criminals can use SMPs to obtain personal information about individuals, which can be used for identity theft and other malicious activities. Furthermore, criminals can use SMPs to spread false information about individuals or organizations, which can lead to reputational damage. SMPs can be a powerful tool for criminals. Additionally, users should be careful when sharing personal information and be aware of any suspicious activity on SMPs.

Hence, we suggested an ontology-based multilayer perceptron (MLP) classifier a feed forward artificial neural network algorithm (MLP-NN) for criminal intention detection in SMPs, which creates program concepts for the choice of social network postings containing criminal slang terms and automatically categorizing these posts in line with illocutionary categories. The system uses trained models from previously published articles to accurately classify published posts with criminal purpose. The suggested method is examined and contrasted with different existing technologies. The results show that the suggested framework is effective in identifying crimes on social media.

1. INTRODUCTION: Crime detection on social media is the Multilayer Perceptron (MLP). MLP is a type of artificial neural network that is able to detect patterns in data and make predictions. MLP has been used for crime detection by analyzing the text, images, and videos posted on social media platforms. By using MLP [1], it is possible to automatically detect criminal activities including drug trafficking, terrorism, and cyber bullying.

Overall, social media platforms provide a powerful means for crime detection. By leveraging the data available on these platforms and applying Artificial Intelligence technologies such as MLP, it is possible to detect criminal activities and take preventative measures. Gathering data from social media sources such as Twitter, Facebook, and Instagram is the first step to using an MLP for crime detection in social media platforms. The data collected from these sources can include text, images, and videos related to criminal activities [2]. After gathering this data, it must be preprocessed and converted into a suitable format for training the MLP. Depending on the type of data gathered, this can involve cleaning the data, normalizing it, and converting it into a numerical format that the MLP can understand and process. Once the data has been preprocessed, it can then be fed into the MLP, which can be optimized to detect crime related activities with a high degree of accuracy.

The MLP is then trained on the preprocessed data using a supervised learning algorithm. The training process involves adjusting the weights of the connections between the nodes in the MLP to minimize the error between the predicted output and the actual output. Once the MLP is trained, it can be used to detect criminal activities in real-time by analyzing the data from social media platforms. The detection of criminal activities employing a multilayer perceptron (MLP) in social media platforms is an emerging field of research that involves the use of artificial intelligence (AI) and machine learning (ML) techniques in order to identify and prevent criminal activities on social media platforms [3]. Social media platforms have become a hub for people to connect, communicate, and share information, but they have also become a breeding ground for criminal activities. MLP can be used to detect and prevent criminal activities on social media, by analyzing user data and providing insights into criminal behavior. MLP can be used to identify and classify criminal activities, such as fraud, cyber bullying, and online harassment. This early detection capability can help law enforcement agencies detect potential criminal activities and respond appropriately to prevent them from occurring. By using MLP to detect criminal activities, law enforcement agencies can proactively take corrective measures to prevent crimes from occurring. Additionally, MLP can be used to detect and identify other types of data trends such as financial fraud and cybercrime. By utilizing MLP, law enforcement agencies can better protect the public from potential criminal activities and other data-related threats. The use of artificial intelligence (AI) to detect and prevent criminal activities on social media platforms has the potential to revolutionize the way companies protect their users.

This technology can be used to rapidly identify criminal activity, such as fraud and other malicious activities, enabling social media companies to respond quickly and effectively. However, there are also concerns that must be addressed before this technology can be responsibly used. The remaining of the section is divided into other sections, such as part 2 which refers to related works, part 3 which refers to a method, part 4 which refers to the results and discussion, and part 5 which refers to the conclusion.

2. SYSTEM ANALYSIS:

2 1 EXISTING SYSTEM:

Existing systems in this domain generally follow a similar structure:

Data Collection:

Gathering data from social media platforms, either through public APIs or web scraping.

Extracting relevant features, such as text content, timestamps, and user information.

Preprocessing:

Cleaning and preparing the data, including tasks like text normalization, tokenization, and feature engineering.

Machine Learning Model:

Designing and training a machine learning model, often a neural network like an MLP.

Using a labeled dataset to train the model for crime detection or sentiment analysis.

Evaluation:

Assessing the model's performance using various metrics, including accuracy, precision, recall, and F1 score.

Deployment:

Integrating the trained model into a real-time system for ongoing monitoring and analysis.

Ethical Considerations:

Ensuring that the system adheres to ethical guidelines, respects user privacy, and complies with the terms of service of the social media platforms.

LIMITATIONS OF EXISTING SYSTEM

Data Bias:

If the training data used to train the machine learning model is biased, the model may learn and perpetuate these biases, leading to skewed predictions.

Limited Generalization:

Models trained on specific datasets may struggle to generalize well to new or diverse data, making them less effective in detecting emerging or uncommon forms of criminal behavior.

Data Privacy Concerns:

Analyzing social media data for crime detection raises privacy concerns, especially if the data contains sensitive information about individuals.

Contextual Ambiguity:

Social media posts often contain slang, abbreviations, or ambiguous language that can be challenging for models to accurately interpret, leading to potential misclassifications.

Dynamic Nature of Social Media:

Social media platforms evolve rapidly, with changes in user behavior, language trends, and platform features. Models may struggle to keep up with these dynamic changes.

Imbalanced Datasets:

If the dataset used for training is imbalanced, with a disproportionately high number of one class compared to others, the model may be biased toward the majority class.

Adversarial Attacks:

Machine learning models, including neural networks, can be susceptible to adversarial attacks where slight modifications to input data can lead to misclassifications.

Computational Complexity:

Training and deploying complex neural network models, such as Multi-Layer Perceptrons, can be computationally expensive and may require substantial resources.

Legal and Ethical Issues:

Employing technology for crime detection on social media must comply with legal and ethical standards. Unintended consequences, misuse, or violation of user privacy can lead to legal and ethical challenges.

User Intent Understanding:

Determining user intent solely based on social media posts can be challenging, and models may struggle to differentiate between genuine expressions and sarcasm, humor, or other nuanced language.

2.2 PROPOSED SYSTEM:

In the proposed system, data collection would be focused on obtaining diverse and representative datasets from various social media platforms. Advanced web scraping techniques or the use of platform-provided APIs would facilitate the extraction of textual content, metadata, and user information. Preprocessing steps would involve thorough cleaning and transformation of the data, incorporating techniques such as text normalization and feature engineering to enhance the model's ability to discern relevant patterns.

The core of the system would feature a Multi-Layer Perceptron (MLP), a type of artificial neural network, designed to analyze the complex and dynamic nature of social media content. Training the MLP would involve utilizing labeled datasets, encompassing instances of criminal and non-criminal activities, to enable the model to learn and generalize patterns indicative of criminal behavior. Continuous refinement and optimization of the model would be essential to adapt to the evolving landscape of social media language and user interactions.

To address potential limitations, the proposed system could incorporate mechanisms to mitigate bias in the training data, ensure interpretability of model predictions, and regularly update the model to account for

emerging trends and new forms of criminal activity. Ethical considerations, user privacy, and compliance with legal standards would be integral parts of the system's design, emphasizing responsible and transparent use of technology for crime detection in the social media domain.

ADVANTAGES OF PROPOSED SYSTEM

Advanced Machine Learning Techniques:

Leveraging Multi-Layer Perceptron (MLP) neural networks allows for the extraction and learning of complex patterns in social media data, enhancing the system's ability to detect subtle indicators of criminal behavior.

Diverse and Representative Data:

The system aims to collect diverse and representative datasets from various social media platforms, ensuring a comprehensive understanding of different types of criminal activities and user behaviors.

Real-time Monitoring and Alerting:

The inclusion of real-time monitoring features enables the system to promptly detect and respond to potential threats, allowing for timely intervention by law enforcement or relevant authorities.

Adaptability to Dynamic Environments:

Continuous refinement and optimization of the model ensure adaptability to the dynamic nature of social media, accommodating shifts in user behavior, language trends, and emerging forms of criminal activity.

Ethical Considerations and Privacy Protection:

The proposed system emphasizes ethical guidelines, user privacy, and compliance with legal standards, addressing concerns related to responsible and transparent use of technology for crime detection on social media platforms.

Bias Mitigation:

Mechanisms are incorporated to mitigate bias in the training data, promoting fair and unbiased model predictions across diverse user groups and demographics.

Interpretability of Model Predictions:

Efforts are made to enhance the interpretability of the model's predictions, allowing users and stakeholders to understand how and why certain decisions are made, contributing to increased trust in the system.

Swift Response to Emerging Trends:

Regular updates and model retraining enable the system to adapt to emerging trends and new forms of criminal activity, ensuring that it remains effective in identifying the latest online threats.

Collaboration with Law Enforcement:

The system facilitates collaboration with law enforcement agencies, providing valuable insights and supporting their efforts in investigating and addressing criminal activities on social media platforms.

Holistic Approach to Crime Detection:

By addressing various challenges, including context ambiguity, imbalanced datasets, and adversarial attacks, the proposed system takes a holistic approach to crime detection, enhancing its overall reliability and performance.

3. SYSTEM IMPLEMENTATION:

The overall functionality of the project can be explained in 4 modules.

MODULES

Data Collection Module:

Objective: Gather diverse and representative datasets from social media platforms.

Functionality*:* Utilize web scraping techniques or social media platform APIs to collect text data, metadata, and user information.

Tasks: Define data collection protocols, establish connections with social media platforms, and implement mechanisms for continuous data updates.

Preprocessing Module:

Objective: Prepare and clean the collected data for effective model training.

Functionality: Perform text normalization, tokenization, and feature engineering to enhance the quality of input data for the machine learning model.

Tasks: Develop algorithms for data cleaning, handle missing values, and convert textual data into numerical features suitable for the Multi-Layer Perceptron.

Machine Learning Module (MLP Design and Training):

Objective: Design and train the Multi-Layer Perceptron for crime detection.

Functionality: Create the neural network architecture, define input and output layers, and train the model using labeled datasets.

Tasks: Implement back propagation for weight adjustments, optimize hyper parameters, and monitor the training process for convergence.

Real-Time Monitoring and Alerting Module:

Objective*:* Enable the system to monitor social media activities in real-time and provide alerts for potential criminal behavior.

Functionality: Implement a continuous monitoring mechanism for incoming data, integrate alerting systems based on model predictions, and facilitate rapid response to identified threats.

Tasks: Develop algorithms for real-time data processing, establish communication channels for alerts, and integrate the module with external notification systems.

Ethical Compliance and Privacy Protection Module:

Objective: Ensure the system adheres to ethical guidelines, user privacy, and legal standards.

Functionality*:* Implement measures to mitigate bias, enhance model interpretability, and incorporate privacy-preserving techniques.

Tasks: Define ethical considerations, conduct bias assessments, and establish protocols for handling and protecting user data in accordance with relevant regulations.

3.1 HARDWARE REQUIREMENTS:

MINIMUM (Required for Execution)		MY SYSTEM (Development)
System	Pentium IV 2.2 GHz	i3 Processor 5th Gen
Hard Disk	20 Gb	500 Gb
Ram	1 Gb	4 Gb

3.2 SOFTWARE REQUIREMENTS:

Operating System	Windows 10/11
Development Software	Python 3.10
Programming Language	Python
Integrated Development Environment (IDE)	Visual Studio Code
Front End Technologies	HTML5, CSS3, Java Script
Back End Technologies or Framework	Django
Database Language	SQL
Database (RDBMS)	MySQL
Database Software	WAMP or XAMPP Server
Web Server or Deployment Server	Django Application Development Server
Design/Modeling	Rational Rose

4. SYSTEM TEST

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

4.1 TYPES OF TESTS:

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is

functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they

actually, run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as

specified by the business and technical requirements, system documentation, and user manuals. Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or

special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

CONCLUSION: The suggested Crime Detection System MLP (Multilayer Perceptron) model is a powerful tool to help law enforcement agencies detect criminal activity at an early stage. It can be trained using various features such as user behavior, keywords, and geo location data to classify posts or messages as potentially related to criminal activity. This approach enables law enforcement to take appropriate action as soon as possible. However, it is important to be aware of the ethical and legal considerations of using an MLP model to detect and classify potentially criminal activity. It is essential to ensure that the data used to train the model is secure and safe, and that the model is not used to unfairly target specific populations or individuals. Furthermore, the model must be regularly tested and updated to ensure it is accurate and effective.

Finally, the results of the model must be carefully analyzed to ensure that any decisions taken by law enforcement are valid and within the law. Adam Optimizer provided excellent results with 96% accuracy after we tried modifying settings such as various learning rates, Optimizers, and a learning rate of 0.0001. As compared to other criminal detection methods already in use, the suggested system produces excellent results. In conclusion, crime detection using an MLP model on social media platforms can be a promising approach to enhance public safety. However, as with any technology, it is important to ensure that the system is developed and used in a responsible and ethical manner. This includes taking into account data privacy and security, as well as ensuring that the AI model is regularly monitored and updated with the latest information. It is also important to consider the potential implications of using such a system and take appropriate steps to mitigate any risks. With the right approach and the right safeguards in place, AI-based crime detection on social media platforms could prove to be a powerful tool in the fight against crime.

REFERENCES

1. Kethineni, S. and Cao, Y., 2020. The rise in popularity of crypto currency and associated criminal activity. International Criminal Justice Review, 30(3), pp.325-344.

2. Lu, J.G., Lee, J.J., Gino, F. and Galinsky, A.D., 2018. Polluted morality: Air pollution predicts criminal activity and unethical behavior. Psychological science, 29(3), pp.340-355.

3. Navalgund, U.V. and Priyadharshini, K., 2018, December. Crime intention detection system using deep learning. In 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET) (pp. 1-6). IEEE.

4. Rahim, S., Muslim, M. and Amin, A., 2019. Red Flag And Auditor Experience Toward Criminal Detection Through Profesional Skepticism. Jurnal Akuntansi, 23(1), pp.47-62.

5. Prabakaran, S. and Mitra, S., 2018, April. Survey of analysis of crime detection techniques using data mining and machine learning. In Journal of Physics: Conference Series (Vol. 1000, No. 1, p. 012046). IOP Publishing.

6. Rahim, S., Muslim, M. and Amin, A., 2019. Red Flag And Auditor Experience Toward Criminal Detection Trough Profesional Skepticism. Jurnal Akuntansi, 23(1), pp.47-62.

7. Yadav, S., Timbadia, M., Yadav, A., Vishwakarma, R. and Yadav, N., 2017, April. Crime pattern detection, analysis & prediction. In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA) (Vol. 1, pp. 225-230). IEEE.

8. Babaei, M., Shirzad, J., Taghilou, M., Faghih Fard, P. and Ezazi Ardi, L., 2020. The efficiency of collected biological samples from crime scene on crime detection. Journal of Police Medicine, 10(1), pp.5-12.

9. Sikandar, T., Ghazali, K.H. and Rabbi, M.F., 2019. ATM crime detection using image processing integrated video surveillance: a systematic review. Multimedia Systems, 25, pp.229-251.

10. Pramanik, M.I., Lau, R.Y., Yue, W.T., Ye, Y. and Li, C., 2017. Big data analytics for security and criminal investigations. Wiley interdisciplinary reviews: data mining and knowledge discovery, 7(4), p.e1208.

11. Ram, N., Guerrini, C.J. and McGuire, A.L., 2018. Genealogy databases and the future of criminal investigation. Science, 360(6393), pp.1078-1079.

12. Suzumura, T., Zhou, Y., Baracaldo, N., Ye, G., Houck, K., Kawahara, R., Anwar, A., Stavarache, L.L., Watanabe, Y., Loyola, P. and Klyashtorny, D., 2019. Towards federated graph learning for collaborative financial crimes detection. arXiv preprint arXiv:1909.12946.

13. Rawat, R., Mahor, V., Chirgaiya, S., Shaw, R.N. and Ghosh, A., 2021. Analysis of darknet traffic for criminal activities detection using TF-IDF and light gradient boosted machine learning algorithm. In Innovations in Electrical and Electronic Engineering: Proceedings of ICEEE 2021 (pp. 671-681). Springer Singapore.

14. Hua, N., Li, B. and Zhang, T., 2020. Crime research in hospitality and tourism. International Journal of Contemporary Hospitality Management, 32(3), pp.1299-1323.

15. Arora, T., Sharma, M. and Khatri, S.K., 2019, October. Detection of cyber crime on social media using random forest algorithm. In 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC) (pp. 47- 51). IEEE.

Biography of authors:

Viriyala Jaya was a M.Tech Scholar in Department of CSE inInternational School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India. Her intersted research area is machine learning and artificial intelligence (AI) typically focuses on advanced computational techniques.

Ch Anusha was a Associate Professor (PhD) in Department of CSE inInternational School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India. Her current research work is is machine learning and artificial intelligence (AI) typically focuses on advanced computational techniques that enable machines to learn from data, identify patterns, and make decisions without being explicitly programmed.

V Anil Santosh was a Associate Professor(PhD) and Head of the Department of Department of CSE inInternational School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India. His current research work is a variety of AI subfields, including deep learning, neural networks, natural language processing, and reinforcement learning. Their work may involve developing new algorithms, applying AI to solve real-world problems (like forecasting, automation, or image recognition), and exploring ethical concerns related to AI deployment. Many such authors combine academic research with industry applications, publishing papers, books, or articles aimed at both technical and non-technical audiences.