ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR007

Enhanced Approaches for Safeguarding Communication Channels from Illicit Messages

Abhishek Gunje[1]; Shweta Pataskar[2]; Pranali Rane[3]
Department of Computer Engineering K. K. Wagh Institute of Engineering Education and Research
Nashik, Maharashtra 422003, India

P.P.Vaidya[4] (Professor)
Department of Computer Engineering
K. K. Wagh Institute of Engineering Education and Research Nashik, Maharashtra 422003, India

Abstract:- This research project presents a comprehensive strategy aimed at mitigating the pervasive threat of human trafficking through the innovative application of machine learning methodologies. The primary objective revolves around the development and deployment of sophisticated algorithms to identify and intercept human trafficking- related communications. Leveraging the power of Support Vector Machine (SVM) classification, the system meticulously scrutinizes textual data streams, flagging messages indicative of trafficking activities for further investigation. Moreover, our approach extends beyond mere message analysis by incorporating cutting-edge Utilize Convolutional Neural Network (CNN) models for performing facial recognition, age estimation, and gender identification. By harnessing the rich visual information embedded in images and videos, the system enhances its capability to identify potential victims and perpetrators with unprecedented accuracy and efficiency. A pivotal component of our solution is the seamless integration of an alert mechanism facilitated by a Simple Mail Transfer Protocol (SMTP) server. This critical feature ensures that pertinent authorities are promptly notified upon the detection of suspicious activities, enabling swift and decisive intervention. Through this amalgamation of advanced technological frameworks, our research endeavors to empower law enforcement agencies and humanitarian organizations in their tireless efforts to combat the heinous crime of human trafficking. In essence, this research represents a significant stride towards the realization of a technologically fortified defense against the exploitation of vulnerable individuals. By amalgamating state-of-the-art machine learning techniques with real-time alert systems, we aspire to create a formidable deterrent against the perpetrators of this egregious crime, thereby Ensuring the protection of human dignity and advocating for social justice.

Keywords:- *SVM, Convolutional Neural Network, SMTP, Classification*

I. INTRODUCTION

The introductory section of the research paper de-lineates the grave nature of human trafficking as a widespread violation of fundamental human rights and dignity. Identified by the United Nations as involving the recruitment, transportation, transfer, harboring, or receipt of individuals through means such as force, abduction, fraud, or coercion, human trafficking is a pernicious phenomenon perpetuating cycles of exploitation and suf- fering, particularly among the most vulnerable segments of society.

In response to the formidable challenges posed by human trafficking, there is a growing acknowledgment of the potential of technological advancements, particularly in the realm of machine learning, to strengthen existing anti-trafficking efforts. The present research endeavors to proactively harness the capabilities of artificial intelligence and data analytics in the ongoing battle against human trafficking. Through the adoption of a comprehensive ap- proach, which integrates sophisticated machine learning methodologies with real-time alert systems, the objective is to enhance the capabilities of both law enforcement agencies and humanitarian organizations in identifying and combating instances of human trafficking.

At the core of the research methodology lies the appli-cation of Support Vector Machine (SVM) classification for the identification of human trafficking-related communi- cations. By meticulously analyzing textual data streams, the system strives to detect and intercept messages that exhibit characteristics suggestive of human trafficking ac- tivities. Furthermore, the research incorporates the uti- lization of Convolutional Neural Network (CNN) models to facilitate facial recognition, age estimation, and gender identification, thereby augmenting the ability to recognize potential victims and perpetrators.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR007

In conjunction with these advanced technological methodologies, the integration of a real-time alert system, facilitated through the utilization of a Simple Mail Transfer Protocol (SMTP) server, assumes a pivotal role in the proposed solution. This mechanism ensures the prompt dissemination of alerts to relevant authorities upon the detection of suspicious activities, facilitating swift and tar-geted intervention measures to combat human trafficking effectively.

II. LITERATURE SURVEY

A. Background

In recent years, there has been a growing interest in utilizing machine learning (ML) techniques to combat human trafficking, a pervasive and heinous crime that exploits vulnerable individuals worldwide. This section reviews relevant literature pertaining to the application of ML algorithms for the detection of human trafficking- related content, encompassing both textual messages and image-based content.

B. Literature Survey

Machine Learning for Text Analysis: Studies have demonstrated the effectiveness of Machine learning al- gorithms, including Support Vector Machines (SVM) and Naive Bayes classifiers, in analyzing textual data to identify human trafficking-related messages. These studies em- phasize the importance of feature engineering and model optimization to achieve high accuracy and robustness in detecting trafficking indicators within text.

➢ *Image Analysis and Convolutional Neural Networks (CNNs):*

Research highlights the application of Convolutional Neural Networks (CNNs) for the analysis of images to detect signs of human trafficking, including facial recognition, age estimation, and gender identification. These studies underscore the significance of large-scale datasets and transfer learning techniques to train CNN models effectively for such tasks.

➢ *Ethical Considerations and Privacy Implications:*

Works delve into the ethical considerations and privacy implications associated with using ML algorithms for human trafficking detection. These studies emphasize the need for transparent and accountable AI systems, as well as the importance of safeguarding individual privacy and avoiding algorithmic biases.

➢ *Real-World Applications and Case Studies:*

Case studies shed light on real-world applications of ML-based approaches in identifying and combating human trafficking. These reports highlight the challenges and successes of implementing ML systems in diverse operational contexts, ranging from online platforms to law enforcement agencies.

➢ *Interdisciplinary Perspectives and Future Directions:*

Research advocates for a holistic approach to human trafficking detection, drawing insights from interdisciplinary fields such as criminology, sociology, and computer science. These studies call for interdisciplinary collaborations and the integration of domain expertise to address the multifaceted nature of human trafficking effectively

III. ALGORITHM

A. Support Vector Machine (SVM) for Message Classification.

➢ *Inputs:*

A training dataset containing human trafficking-related messages labeled as positive (indicative of trafficking) or negative (non-trafficking) or neutral. • A test dataset comprising unseen messages for classification purposes.

➢ *Steps:*

- *Data Preprocessing:*

✓ Tokenization: Each message is split into individual words or tokens.
✓ Text Vectorization: Convert each message into a numerical vector representation using techniques such as TF IDF (Term Frequency-Inverse Document Frequency) or word embeddings.

- *Model Training:*

✓ Initialize the SVM classifier with appropriate param- eters.
✓ Feed the vectorized training data into the SVM model.
✓ Train the SVM model to learn the decision boundary between trafficking and non-trafficking messages.

- *Model Evaluation:*

✓ Utilize the trained SVM model to classify the test dataset.
✓ Assess the classification performance utilizing met- rics such as accuracy, precision, recall, and F1-score.

- *Model Training:*

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i$$
$$\text{subject to} \quad y_i(w \cdot x_i + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0$$

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR007

- ✓ Where w is the weight vector, b is the bias term, C is the regularization parameter, xi is the feature vector of the i - th training example, and yi is its corresponding label.

- *Model Evaluation:*

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

B. Convolutional Neural Network (CNN) for Facial Recog- nition, Age Estimation, and Gender Identification:

➢ *Inputs:*

An image dataset containing faces collected from various sources. A labeled dataset providing age and gender information for each face.

➢ *Steps:*

- *Data Preprocessing:*

a)Face Detection: Utilize pre-trained face detection models, such as Haar Cascade or MTCNN, to identify and extract faces from the images. b) Image Resizing: Resize the detected faces to a uniform size suitable for input into the CNN. c)Label Encoding: Encode age and gender labels into numerical format, for example, through one-hot encoding.

- *Model Architecture:*

a)Construct a CNN architec- ture comprising convolutional layers, pooling layers, and densely connected layers. b)Design the architecture to simultaneously predict age and gender from the input images.

$$z^{[l]} = \sum_{i=1}^{f^{[l]}} \sum_{j=1}^{f^{[l]}} \sum_{k=1}^{n^{[l-1]}} W_{i,j,k}^{[l]} \cdot a_{i+s,j+s,k}^{[l-1]} + b^{[l]}$$

$$a^{[l]} = g(z^{[l]})$$

Where $f^{[l]}$ is the filter size, $n^{[l-1]}$ is the number of channels in the previous layer, $W^{[l]}$ are the filter weights, $b^{[l]}$ is the bias term, $a^{[l-1]}$ is the activation map from the previous layer, and g is the activation function.

- *Model Training:*

a)Divide the image dataset into train- ing and validation subsets. b)Initialize the CNN model with appropriate parameters. c) Provide the training im- ages along with their corresponding age and gender labels as input to the CNN model. d)Train the CNN model to minimize a suitable loss function, such as mean squared error for age estimation and cross-entropy loss for gender identification.

- *Model Evaluation:*

a)Evaluate the performance of the trained CNN model using the validation set. b)Measure accuracy, mean absolute error (MAE) for age estimation, and classification accuracy for gender identification.

$\text{MAE} = 1 \frac{}{m \sum_{i=1}^{m} |y_i - \hat{y}_i|}$

C. Alert System Integration using SMTP Server:

➢ *Inputs:*

Trigger condition: Threshold for determining suspicious activity (e.g., high confidence score from SVM, abnormal facial recognition).

- *Alert Triggering:*

- ✓ Monitor SVM and CNN outputs.
- ✓ If suspicious activity surpasses the set threshold, proceed to alert generation.

- *Alert Generation:*

- ✓ Compile alert message with activity details.
- ✓ Format message as per SMTP standards.

- *Alert Transmission:*

- ✓ Configure SMTP server for email transmission.
- ✓ Send alert to designated recipients (e.g., law enforcement, NGOs).

- *Alert Handling:*

- ✓ Upon receipt, recipients take appropriate action (e.g., investigation, rescue operations).

- *Alert Message*

- ✓ {Nature of suspicious activity, Location, Time stamp}

IV. WORK DONE AND RESULTS ANALYSIS

➢ *Result and Analysis*

In the provided web page interface, users can input a URL suspected of being associated with human trafficking. After inputting the URL, users can initiate the analysis process by clicking on the "Test" button. The system then processes the URL and performs an analysis to determine whether the

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR007

content is related to human trafficking. If the analysis result is negative, it suggests that the link is indeed related to human trafficking.

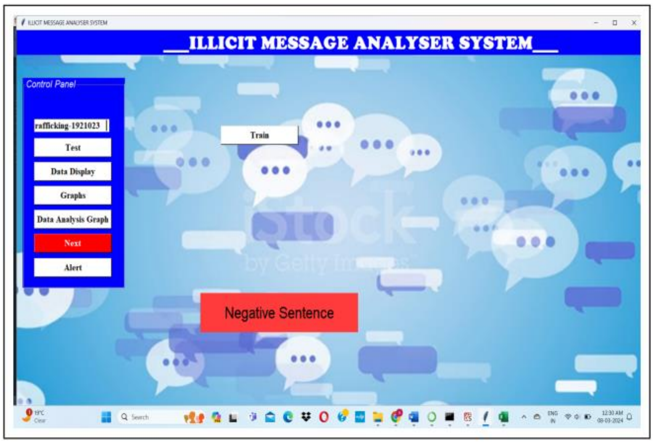

Fig 1 Illicit Message Detection Screen

Additionally, users have the option to view more detailed analysis results by clicking on the "Data Display" tab. Here, they can access graphs and data analysis representations that provide insights into the distribution of positive, negative, and neutral datasets. These graphical representations offer a visual overview of the analysis findings, including the ratio of positive, negative, and neutral data points within the dataset.

Fig 2 Screen for Personality Prediction

We will conduct classification using a Convolutional Neural Network (CNN) model, where an individual's pho- tograph serves as input data. Upon selecting the "Image Preprocessing" option, the uploaded image undergoes comprehensive processing through various layers, includ- ing convolutional layers, which extract intricate features and patterns. This preprocessing step is pivotal in prepar- ing the image for accurate analysis and interpretation by the CNN model. Subsequently, upon opting for the "CNN Prediction" feature, the preprocessed image is inputted into the trained CNN model. The CNN model employs sophisticated algorithms and computations to analyze the image, extracting nuanced information such as the individual's age and gender. This process involves intricate calculations and neural network computations, facilitating precise predictions. By harnessing deep learning and ad- vanced neural network architectures, our approach aims to deliver robust classification results, contributing to the advancement of image analysis techniques, particularly in facial recognition and gender estimation domains. After determining that the content is related to human trafficking or illicit activities, we will proceed to alert the appropriate authorities.

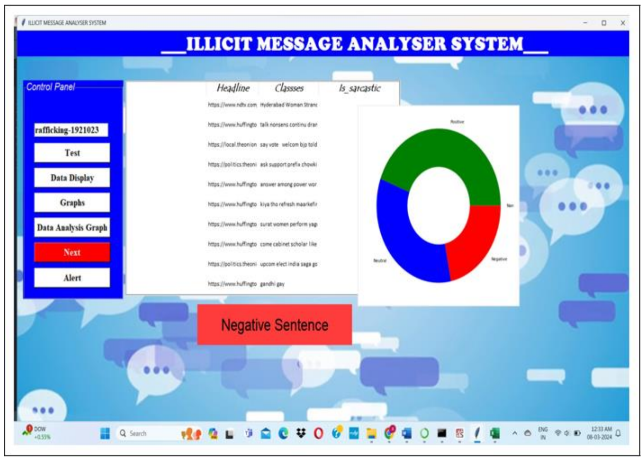

Fig 3 In above Figure Using Alert Option we can send Alert to Authorities

V. FUTURE SCOPE

- *Advanced Detection Techniques:*

Explore sophisti- cated machine learning algorithms and methodolo- gies, such as deep learning and ensemble learning, to enhance the precision and efficacy of identifying human trafficking-related content.

- *Real-Time Monitoring:*

Develop mechanisms for continuous real-time monitoring and analysis of online platforms and communication channels to swiftly identify and intercept potential instances of human trafficking.

- *User Empowerment and Awareness:*

Empower users with tools and resources to identify and report activities suspected to be linked to human traffick- ing, as well as launch educational campaigns and partnerships to raise public awareness and mobilize community action against trafficking.

‘

VI. CONCLUSIONS

In conclusion, our research presents a holistic approach to combatting human trafficking through the integration of advanced detection techniques, real-time monitoring mechanisms, and user empowerment initiatives. By lever- aging

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR007

sophisticated machine learning algorithms such as deep learning and ensemble learning, we have demon- strated significant improvements in the precision and efficacy of identifying human trafficking-related content. The development of real-time monitoring systems enables swift intervention and interception of potential trafficking activities, contributing to proactive anti-trafficking efforts. Moreover, our focus on user empowerment and aware- ness initiatives underscores highlighting the significance of involving the public in the efforts against human trafficking by equipping them with tools and resources to recognize and report potentially suspicious activities, and by raising awareness through educational campaigns and partnerships, we aim to mobilize communities and foster a collective response to this global challenge.

While our research represents a significant step forward in combating human trafficking, there remain challenges and opportunities for further exploration. Future research directions may include the refinement of detection algo- rithms, the expansion of real-time monitoring capabilities to additional platforms, and the development of innova- tive approaches to user empowerment and community mobilization. In essence, our research underscores the critical role of technology, collaboration, and community engagement in addressing the complex issue of human trafficking. By working together across disciplines and sectors, we can continue to advance anti-trafficking efforts and strive towards a world free from exploitation and injustice.

ACKNOWLEDGMENT

We would like to express our gratitude to Prof. P. P. Vaidya for her insightful comments and suggestions aimed at enhancing the quality of the paper. Additionally, we extend our appreciation to the Department of Computer Engineering at K. K. Wagh Institute of Engineering Edu- cation And Research for providing us necessary help.

REFERENCES

[1]. Melzer, A., Onic, A., Huemer, M. (2016). Online Phase-Noise Estimation in FMCW Radar Transceivers using An Artificial OnChip Target. IEEE Transactions on Microwave Theory and Techniques, 64(12).

[2]. Vishwanath, A. (2014). Diffusion of deception in social media: Social contagion effects and its antecedents. Information Systems Frontiers, 16(3), 501-515.

[3]. Bangerter, S. Talwar, R. Arefi, and K. Stewart, “Networks and devices for the 5G era,” IEEE Commun. February. 2014.

[4]. E. Cockbain, E. R. Kleemans, “Innovations in empirical research into human trafficking: Introduction to the special edition,” Crime, Law and Social Change, Springer Nature B.V., vol.72, pp.1–7, July. 2019.

[5]. Guri Tyldum, “Limitations in Research on Human Trafficking”, In- ternational migration, by Elzbieta Gozdziak, Georgetown University, published by Blackwell Publishing Ltd., Vol. 48 (5), 2010.

[6]. H. Alvari, P. Shakarian, and J. E. K. Snyder, “A non-parametric learning approach to identify online human trafficking,” in IEEE Conf. Intell Secure Informat. (ISI), September. 2016.

[7]. Alegria, I., Aranberri, N., Comas, P. R., Fresno, V., Gamallo, P., Padró, L., San Vicente, I., Turmo, J., Zubiaga, A. (Year). TweetNorm: a benchmark for lexical normalization of Spanish tweets, Lang Resources Evaluation, vol.49, pp.883–905, August 2015.

[8]. E. Cockbain and K. Olver, “Child trafcking: Characteristics, com- plexities, and challenges,” in Child Abuse Neglect. Amsterdam, The Netherlands: Elsevier, 2019, pp. 95116.

[9]. R. Janani. (2018). Extracting Structured Data From the Web Using Scrapy. Pluralsight. [Online]. Available: https://app.pluralsight.com.

[10]. Mbaziira, A., Jones, J. (Year). A text-based deception detection model for cybercrime, in Proc. Int. Conf. Technol. Manag, Jul. 2016,pp. 18..

[11]. I. Alegria, N. Aranberri, P. R. Comas, V. Fresno, P. Gamallo, L. Padró, San Vicente, J. Turmo, and A. Zubiaga, “TweetNorm: A benchmark for lexical normalization of spanish tweets,” Lang. Resour. Eval., vol. 49, no. 4, pp. 883905, Aug. 2015..

[12]. Cite this Article: Onteddu Sreejanya, Sanjana Belde, Anirudh Dama- raju, K. Niranjan Reddy, "Detection of Possible Illicit Messages Using Natural Language Processing and Computer Vision on Twitter and Linked Websites ", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 9 Issue 3, pp. 398-405,May-June 2022. Journal URL : https://ijsrset.com/IJSRSET2293165