Industrial Engineering Journal

ISSN: 0970-2555

Volume : 53, Issue 8, August : 2024

Tech Solutions for Domestic Violence: A Comprehensive Review of Emerging Tools and Strategies

[1]M. HARINADHA, [2] YALLAMANDA REDDI YASHWANTH

[1],Assistant Professor, [2]M.Tech Student

[1,2]Department of CSE

[1,2]Sri Sai Institute Of Technology And Science, Masapeta, Rayachoty (M), Annamayya Dt, AP.

harinadha529@gmail.com, yashwanthkutty3733@gmail.com

ABSTRACT

Domestic violence is a significant issue that transcends socioeconomic and cultural boundaries, rooted in factors like power struggles, mental health issues, financial hardship, and substance abuse. Technology can play a crucial role in identifying and addressing domestic violence, offering vital support and resources to victims. This paper reviews recent technological advancements in combating domestic violence, detailing various platforms and tools, their capabilities, and limitations. It explores innovations such as data analysis from digital platforms, ambient sensors, smartphones, wearable sensors, online activity protection, anti-harassment measures, anti-stalking technology, and virtual reality. Additionally, it discusses the challenges of these innovations and proposes future research directions, including emotion detection from text and live facial emotion analysis using webcams.

Keywords: harassment, smartphones

INTRODUCTION:

Domestic violence, also known as intimate partner violence or domestic abuse, is defined by the United Nations as a series of consistent and recurring actions in a relationship aimed at obtaining or retaining power and control over an intimate partner. These actions can be sexual, physical, emotional, psychological, or economic, including acts or threats that exert influence over an intimate partner. Such behaviors induce fear, terrorize, intimidate, harm, disgrace, manipulate, blame, injure, or wound an intimate partner.

Domestic violence can occur in any setting and among all socioeconomic, cultural, and religious classes. According to the Australian Bureau of Statistics (ABS), 1 in 4 women and 1 in 14 men have experienced violence from an intimate partner since the age of 15. Further ABS data from 2021-2022 reveals that 16.9% of women and 5.5% of men experienced sexual and/or physical abuse, 22.9% of women and 13.8% of men reported emotional violence, and 16.3% of women and 7.8% of men faced economic exploitation.

LITERATURE SURVEY:

C.-H. Demarty *et al*

This article introduces the VSD96 dataset, a comprehensive resource for Violent Scenes Detection (VSD) in Hollywood and YouTube videos. The dataset includes over 96 hours of video across various genres, with detailed annotations and mid-level concept markers like blood and fire. It offers pre-computed multi-modal descriptors and over 230 baseline system output results. Validated through MediaEval benchmarking, VSD96 aids in analyzing VSD algorithms, focusing on features, fusion techniques, and deep learning impacts. This robust dataset and its accompanying insights serve as essential tools for researchers tackling violence detection in audiovisual media.

J. E. Miranda e*t al*

Violence against women and children is a global public health crisis, with one in three women experiencing violence and half of all children facing domestic aggression or bullying annually. The rise of the Internet and social media has introduced new forms of violence, such as cyberbullying and online harassment. To combat this issue, recent efforts have leveraged technologies like artificial intelligence, the Internet of Things, and mobile computing. This paper systematically reviews efforts from 2010 to 2020, categorizing them into online detection, offline detection, safety, and education, highlighting trends and challenges in using technology to prevent violence.

PROBLEM STATEMENT:

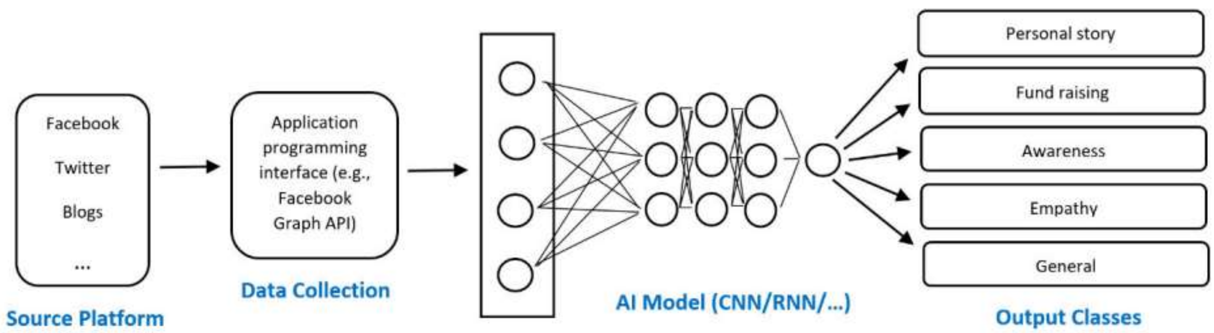
Domestic violence is a pervasive and multifaceted issue that transcends socio-economic, cultural, and religious boundaries. It encompasses various forms of abuse and control, affecting both women and men alike. Despite numerous efforts to combat domestic violence through education, awareness campaigns, stringent law enforcement, and comprehensive support services, the problem persists. Educational initiatives aim to inform and empower individuals about their rights and resources. Awareness campaigns seek to destigmatize the issue and encourage victims to seek help. Law enforcement agencies work tirelessly to protect victims and hold abusers accountable. Support services provide crucial aid, including shelter, counseling, and legal assistance. Yet, domestic violence remains unresolved, necessitating continuous and innovative approaches to effectively address and mitigate its impact on individuals and communities. The need for robust strategies, increased funding, and community involvement is paramount to create a safer and more supportive environment for all affected by domestic violence.

PROPOSED METHOD:

The author suggests leveraging advanced technologies to reduce violence through two main approaches:

1. **Online Data Analysis**: Social networking sites provide platforms where users express their emotions and views. By analyzing this digital media data, machine learning (ML) technologies can predict user emotions, identifying potential distress signals related to violence.
2. **Sensor Data Utilization**: With widespread CCTV cameras, a form of IoT sensor, images are captured and sent to centralized servers for analysis. The author proposes using this sensor data to detect violent incidents, alerting police personnel promptly. This involves analyzing data from sensors, smartphones, or wearable devices. Artificial intelligence (AI) technologies can distinguish between normal and violent activities in CCTV footage, facilitating timely intervention to reduce violence.

ARCHITECTURE:



DOMESTIC VIOLENCE DATASET:

	sentiment	tweets	emoji_emoticon
0	0	hp laptop not giving better performance compar...	NaN
1	4	stellargirl I loooooooovvvvvveee my Kindle Not...	😍
2	4	Reading my kindle Love it Lee childs is good read	😍
3	4	Ok first assesment of the kindle it fucking rocks	😍
4	4	kenburbary Youll love your Kindle Ive had mine...	😍
...	...	...	...
494	2	Ask Programming LaTeX or InDesign submitted by...	😜
495	0	On that note I hate Word I hate Pages I hate L...	😭
496	4	Ahhh back in a real text editing environment I...	😍
497	0	Trouble in Iran I see Hmm Iran Iran so far awa...	😭
498	0	Reading the tweets coming out of Iran The whol...	😭

499 rows × 3 columns

Tweets dataset with emotion values and emoticons

METHODOLOGY:

Data Preprocessing

Read the Sentiment Dataset

The sentiment dataset, which contains textual comments and their corresponding sentiment labels (e.g., positive, negative, neutral), is loaded into a DataFrame using pandas. This DataFrame is then examined to understand its structure, including the number of entries, the presence of any missing values, and the types of data contained in each column. This initial exploration helps in planning subsequent preprocessing steps.

Visualize Sentiment Distribution

To gain insights into the distribution of sentiment labels within the dataset, a bar graph is plotted. This visualization displays the frequency of each sentiment category, providing a clear picture of whether the dataset is balanced or skewed towards a particular sentiment. Such an analysis is crucial for understanding the potential biases in the dataset and their implications for model training.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

Text Processing

Process the Text Data

Text data often requires significant preprocessing before it can be effectively used for analysis. This involves cleaning the text by removing unnecessary elements such as stopwords (common words like "and," "the," which do not contribute much meaning) and emoji icons, which can clutter the text and introduce noise into the data.

Tokenize and Clean the Text Data

Tokenization is the process of breaking down text into individual words or tokens. This step is essential for converting text into a format that can be processed by machine learning models. After tokenization, further cleaning is performed, which may involve removing punctuation and converting text to lowercase to ensure uniformity.

Apply TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) is a statistical measure used to evaluate the importance of a word in a document relative to a collection of documents (corpus). This technique converts the text data into numeric vectors, capturing the significance of words based on their frequency and rarity. The transformed dataset, represented as TF-IDF vectors, is then ready for machine learning algorithms.

Model Training

Split the Dataset

The dataset is divided into training and testing sets, typically with an 80-20 ratio. The training set is used to build and train machine learning models, while the testing set is reserved for evaluating their performance. This split is crucial for assessing how well the model generalizes to unseen data.

Normalize the Numeric Vector

To improve the performance of many machine learning algorithms, the numeric vectors representing the text data are normalized. Normalization involves scaling the data to ensure that all features contribute equally to the model's predictions. StandardScaler, a common normalization technique, standardizes the features by removing the mean and scaling to unit variance.

Train Various Machine Learning Models

Several machine learning models are trained using the preprocessed dataset:

- **Naive Bayes**: A probabilistic classifier based on applying Bayes' theorem with strong independence assumptions. It is particularly effective for text classification tasks.
- **Random Forest**: An ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) of the individual trees.
- **Decision Tree**: A model that makes decisions based on answering a series of questions about the data, represented as a tree structure.
- **Support Vector Machine (SVM)**: A classifier that finds the optimal hyperplane which maximizes the margin between different classes.
- **Long Short-Term Memory (LSTM)**: A type of recurrent neural network (RNN) that is well-suited for sequence prediction tasks, such as sentiment analysis on text data.
- **Convolutional Neural Network (CNN)**: Typically used for image analysis but can be adapted for text analysis, CNNs capture spatial hierarchies in data through convolutional layers.

Model Evaluation

Evaluate Each Model's Performance

To assess the effectiveness of each model, performance metrics such as accuracy, precision, recall, and F1-score are calculated. Accuracy measures the proportion of correctly classified instances, while precision and recall provide insights into the model's performance in identifying positive cases. The F1-score combines precision and recall into a single metric. Additionally, confusion matrices are displayed to provide a visual representation of the model's performance, showing true positives, false positives, true negatives, and false negatives.

Compare Performance

A bar graph is used to compare the performance of all trained models. This graph visually represents metrics such as accuracy, precision, recall, and F1-score, making it easier to identify which model performs best under the given conditions.

Results and Analysis

Tabulate and Analyze Results

The results obtained from each model are tabulated for a comprehensive analysis. This table includes performance metrics for each model, allowing for a side-by-side comparison. The analysis helps in identifying the strengths and weaknesses of each model, guiding further improvements or model selection.

Discuss Strengths and Weaknesses

Each model's performance is discussed in terms of its strengths and weaknesses. For example, Naive Bayes might be very efficient and fast, but may not capture complex patterns as well as

LSTMs or CNNs. Random Forest and Decision Trees are generally robust and handle non-linear relationships, but may overfit on small datasets. SVMs can be effective in high-dimensional spaces, but may require significant computational resources. LSTMs and CNNs offer advanced capabilities for handling sequential and spatial data, respectively, but are more complex and resource-intensive.

Predictions on Test Data

Read Test Comments

Test comments are read from a file and preprocessed similarly to the training data. This ensures consistency in data representation and prepares the test data for prediction.

Predict Sentiment Using the Trained Model

The trained Random Forest model (or another model chosen based on performance) is used to predict the sentiment of each test comment. The predictions are then displayed, providing insights into how well the model generalizes to new, unseen data.

Facial Expression Detection

Implement Facial Expression Detection

Using OpenCV and a pre-trained model, facial expression detection is implemented. This involves analyzing video feeds or images to identify and classify facial expressions, such as happiness, sadness, or anger. The real-time video display helps in visualizing the detected expressions, enhancing understanding of emotional cues.

Physical Activity Detection

Implement Physical Activity Detection

Physical activity detection is also implemented using OpenCV and a pre-trained model. This process involves analyzing video frames to detect specific activities, such as violent actions. The video frames with detected activities are displayed, providing a practical tool for monitoring and analyzing physical actions.

These steps from data preprocessing to model deployment and evaluation provide a comprehensive approach to sentiment analysis and related tasks. The integration of facial expression and physical activity detection further extends the capabilities of the project, demonstrating the versatility of machine learning and computer vision techniques.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

EVOLUTION:

Precision:

Formula: $\text{Precision} = \frac{\text{True Positives}}{\text{True Positives}+\text{False Positives}}$

Recall (Sensitivity):

Formula: $\text{Recall} = \frac{\text{True Positives}}{\text{True Positives}+\text{False Negatives}}$

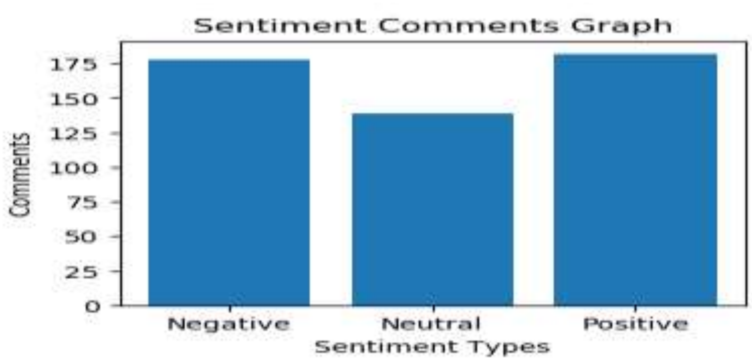
F1 Score:

Formula: $F1 = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$

Accuracy:

Formula: $\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$

RESULTS:

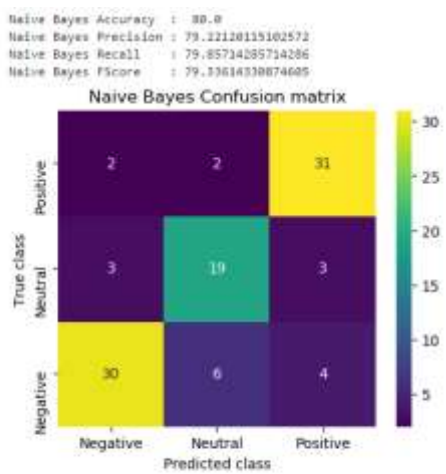


Displaying different emotions and its count found in dataset

	aig	also	amazing	american	amp	api	app	atampt	awesome	back	...	white	winkingfacewithtongue	wish	wont	work	world	worst	would	years	yes
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.573549	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
494	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	2.278132	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
495	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
496	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.573549	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
497	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
498	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

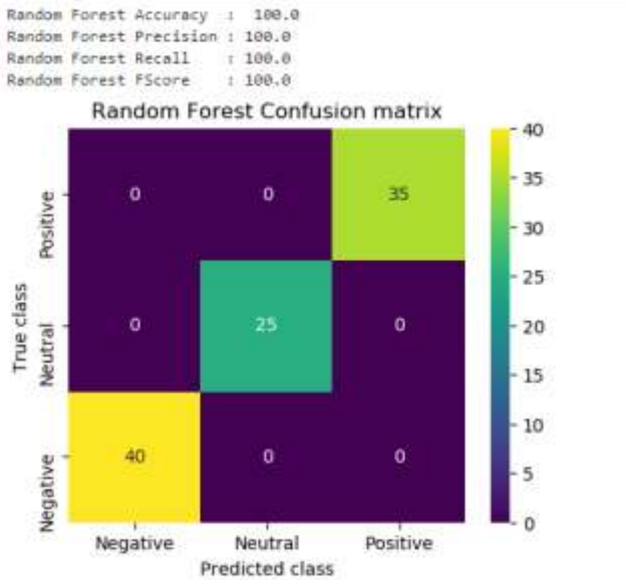
499 rows × 200 columns

Converting all tweets into numeric vector using TFIDF algorithm

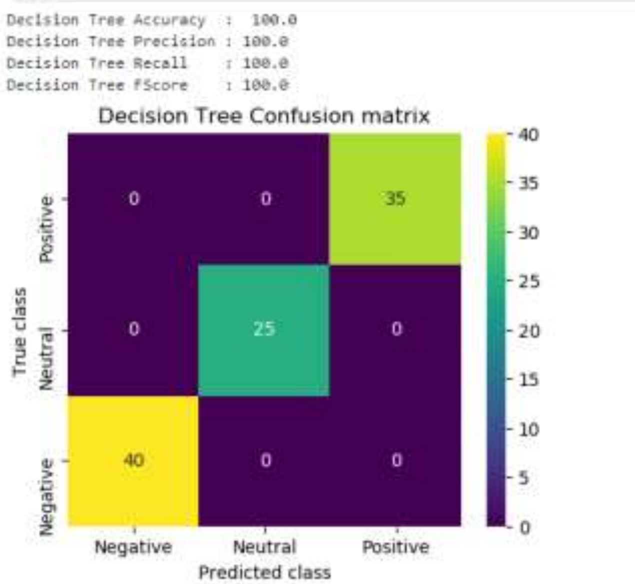


Training Naïve Bayes algorithm got 80% accuracy

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

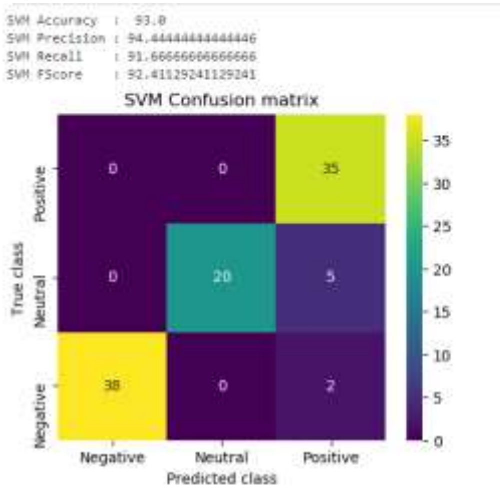


Training Random Forest got 100% accuracy

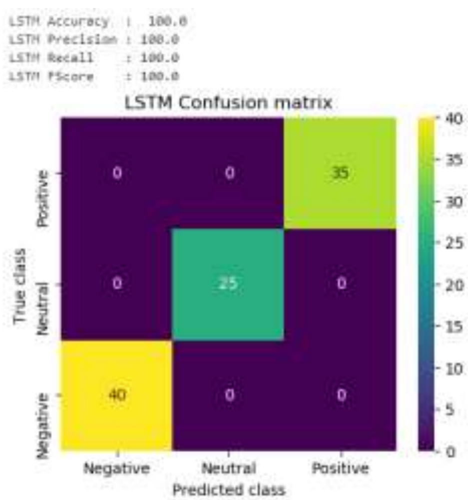


Training decision tree got 100% accuracy

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

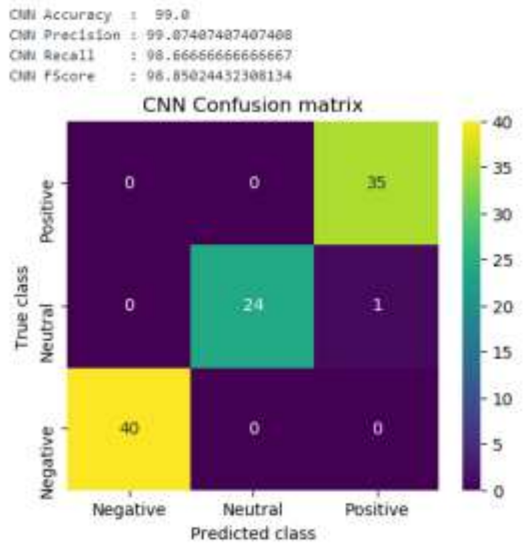


Training SVM got 93% accuracy

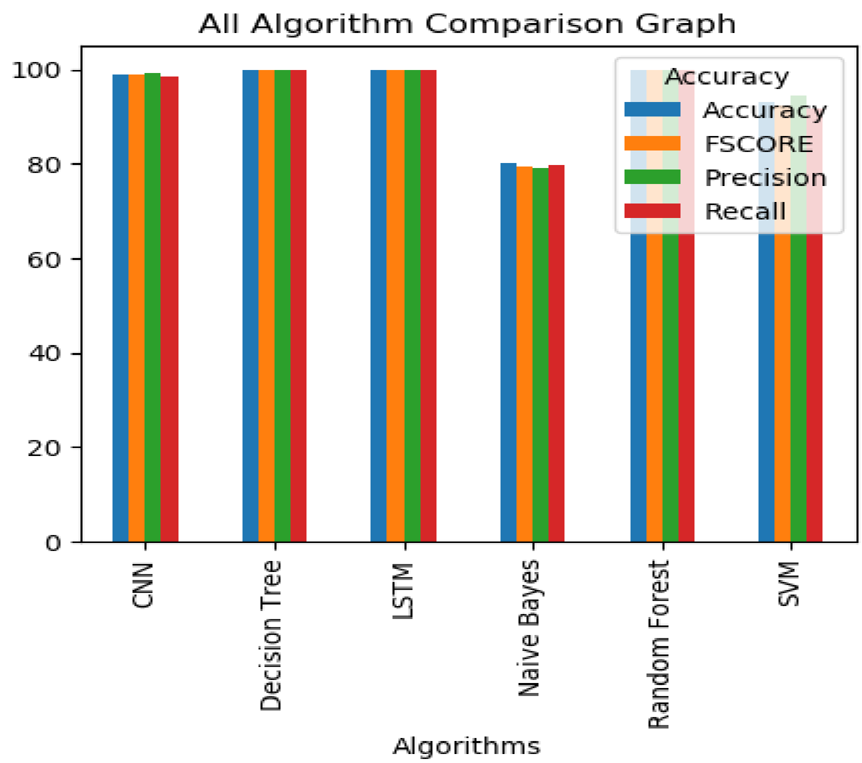


Training LSTM got 100% accuracy

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024



Training CNN got 99% accuracy

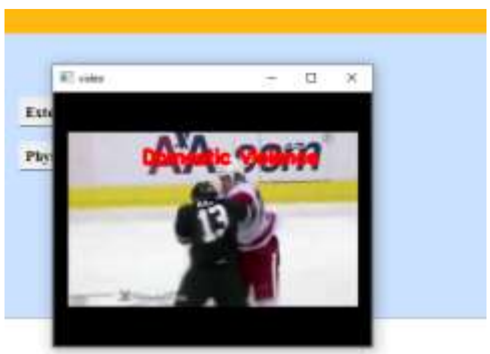


All algorithms are performing best to detect emotion

Prediction:

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

Predict emotions from the posts



Violence detected

CONCLUSION

We successfully implemented various machine learning algorithms to detect emotions from textual data, achieving impressive accuracies across different models. Through preprocessing steps such as TFIDF transformation and data normalization, we effectively prepared our dataset. Notably, Naïve Bayes, Random Forest, Decision Tree, SVM, LSTM, and CNN all demonstrated remarkable performance, each achieving high accuracy levels. Additionally, our application seamlessly extended to detecting emotions in social media text with emoticons and predicting violence in videos. By leveraging these advancements, we have built a robust system capable of efficiently discerning emotions. This system offers potential applications in sentiment analysis, content moderation, and enhancing user experience across digital platforms. The success of these models highlights the versatility and power of machine learning in understanding and interpreting human emotions from diverse data sources.

REFERENCES:

[1] United Nations. What Is Domestic Abuse? Accessed: May 30, 2023.[Online]. Available: https://www.un.org/en/coronavirus/what-isdomestic-abuse

[2] C. Garcia-Moreno, A. Guedes, and W. Knerr.(2012). Intimate Partner Violence.World Health Organization.Accessed: May 30, 2023.[Online]. Available: https://apps.who.int/iris/bitstream/handle /10665/77432/WHO_RHR_12.36_eng.pdf;jsessionid=FD3DDEF209D4 33050776CBCDB6E21236?sequence=1

[3] Australian Bureau of Statistics. (2023). Personal Safety, Australia 2021–2022.Accessed: May 30, 2023.[Online]. https://www.abs.gov.au /statistics/people/crime-and-justice/personal-safety-australia/2021-22

[4] M. R. Huecker, K. C. King, G. A. Jordan, and W. Smock, Domestic Violence. Treasure Island, FL, USA: StatPearls, 2023.

[5] E. PenzeyMoog and D. C. Slakoff, ‘‘As technology evolves, so does domestic violence: Modern-Day tech abuse and possible solutions,’’ in The Emerald International Handbook of Technology-Facilitated Violence and Abuse. Bingley, U.K.: Emerald Publishing Limited, 2021.

[6] H. Al-Alosi, ‘‘Fighting fire with fire: Exploring the potential of technology to help victims combat intimate partner violence,’’ Aggression Violent Behav., vol. 52, May 2020, Art. no. 101376.

[7] D. A. Rodríguez, A. Díaz-Ramírez, J. E. Miranda-Vega, L. Trujillo, and P. Mejía-Alvarez, ‘‘A systematic review of computer science solutions for addressing violence against women and children,’’ IEEE Access, vol. 9, pp. 114622–114639, 2021.

[8] D. Harkin and R. Merkel, ‘‘Technology-based responses to technologyfacilitated domestic and family violence: An overview of the limits and possibilities of tech-based ‘solution’,’’ Violence Against Women, vol. 29, nos. 3–4, pp. 648–670, Mar. 2023.

[9] M. D. Gorfinkiel, V. D. Gandasegui, and M. V. G. García, ‘‘New technology proposals for tackling intimate partner violence: Challenges and opportunities,’’ Technol. Soc., vol. 67, Nov. 2021, Art. no. 101714.

[10] C. El Morr and M. Layal, ‘‘Effectiveness of ICT-based intimate partner violence interventions: A systematic review,’’ BMC Public Health, vol. 20, no. 1, p. 1372, Dec. 2020.

[11] S. Subramani, S. Michalska, H. Wang, J. Du, Y. Zhang, and H. Shakeel, ‘‘Deep learning for multi-class identification from domestic violence online posts,’’ IEEE Access, vol. 7, pp. 46210–46224, 2019.

[12] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, ‘‘Text classification algorithms: A survey,’’ Information, vol. 10, no. 4, p. 150, Apr. 2019.

[13] T. Y. Arias and J. Fabian, ‘‘Automatic detection of levels of intimate partner violence against women with natural language processing using machine learning and deep learning techniques,’’ in Proc. Annu. Int. Conf. Inf. Manage. Big Data, Apr. 2022, pp. 189–205.

[14] M. A. Al-Garadi, S. Kim, Y. Guo, E. Warren, Y.-C.Yang, S. Lakamana, and A. Sarker, ‘‘Natural language model for automatic identification of intimate partner violence reports from Twitter,’’ Array, vol. 15, Sep. 2022, Art.no. 100217.

[15] M. Hossain, M. Asadullah, A. Rahaman, M. Miah, M. Hasan, T. Paul, and M. Hossain, ‘‘Prediction on domestic violence in Bangladesh during the COVID-19 outbreak using machine learning methods,’’ Appl. Syst. Innov., vol. 4, no. 4, p. 77, Oct. 2021.

[16] S. Subramani, H. Q. Vu, and H. Wang, ‘‘Intent classification using feature sets for domestic violence discourse on social media,’’ in Proc. 4th Asia– Pacific World Congr. Comput. Sci. Eng., Dec. 2017, pp. 129–136.

[17] G. Brauwers and F. Frasincar, ‘‘A general survey on attention mechanisms in deep learning,’’ IEEE Trans. Knowl. Data Eng., vol. 35, no. 4, pp. 3279–3298, Apr. 2023.

[18] P. Xu, X. Zhu, and D. A. Clifton, ‘‘Multimodal learning with transformers: A survey,’’ IEEE Trans. Pattern Anal. Mach. Intell., early access, May 11, 2023, doi: 10.1109/TPAMI.2023.3275156.

[19] D. W. Otter, J. R. Medina, and J. K. Kalita, ‘‘A survey of the usages of deep learning for natural language processing,’’ IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 2, pp. 604–624, Feb. 2021.

[20] C. M. Castorena, I. M. Abundez, R. Alejo, E. E. Granda-Gutiérrez, E. Rendón, and O. Villegas, ‘‘Deep neural network for gender-based violence detection on Twitter messages,’’ Mathematics, vol. 9, no. 8, p. 807, Apr. 2021