iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

SENTIMENTAL ANALYSICS ON SOCIAL MEDIA

Dr. Shobha Hugar,
Associate Professor, Department of electronics and communication Engineering, Sapthagiri college of engineering, Bangalore-57, Karnataka, India Affiliated to Visvesvaraya Technological University, Belagavi
Sinchana C, **Sreevatsa S**, **Tejaswini S S**, **V Vishnu**,
UG Students, Department of electronics and communication Engineering, Sapthagiri college of engineering, Bangalore-57, Karnataka, India Affiliated to Visvesvaraya Technological University, Belagavi

ABSTRACT
This project aims to combat misinformation on social media by developing a system that combines sentiment analysis, fake news detection, and deep fake recognition. Using VADER lexicon, the sentiment analysis module will categorize social media posts into positive, negative, or neutral sentiment. A fake news detection module will identify misleading content using advanced techniques, while a deep fake recognition system, powered by PyTorch and LSTM networks, will analyze media for anomalies indicative of synthetic manipulation. This integrated approach seeks to enhance trust and accountability in online information, contributing to a more informed and safer digital environment.
KEYWORD: VADER, LSTM, Sentiment Analysis, Fake News Detection, Deep fake Recognition.

INTRODUCTION
The rise of social media has created both opportunities and challenges. While platforms like Twitter and Facebook offer valuable insights into public sentiment, the unstructured data and proliferation of misinformation, including fake news and deepfakes, pose significant threats. This project addresses these challenges by developing a unified framework integrating sentiment analysis, fake news detection, and deepfake identification. Using the VADER lexicon, the project analyzes social media text for sentiment, while a PyTorch-based LSTM model detects deepfakes by analyzing temporal patterns in media. A dedicated fake news module further enhances the system's ability to identify and mitigate the spread of false information. This integrated approach aims to foster transparency and trust in online communication, benefiting industries, governments, and academia by enabling more informed decision–making.

OBJECTIVES
The key objectives of this research are:

1. Develop a Sentiment Analysis Module:
    a. Implement an efficient and accurate sentiment analysis system using the VADER Lexicon, specifically optimized for processing informal and context rich social media text.
    b. Address challenges such as informal language, sarcasm, slang, and evolving trends to classify user-generated content into categories like positive, negative, or neutral.
    c. Provide actionable insights for organizations to understand public opinion, consumer behavior, and market trends effectively.

2. Implement a Fake News Detection Module:
    a. Design a robust system capable of identifying and mitigating the spread of fake news across multiple online platforms.
    b. Use state-of-the-art NLP techniques such as word embeddings, LSTM networks, or transformer models to detect and analyze the linguistic and contextual cues associated with fake news.
    c. Enhance the detection system with data augmentation methods to improve accuracy across diverse datasets, including noisy or incomplete inputs.

3. Design a Deepfake Detection System:
    a. Develop a PyTorch-based LSTM model to analyze temporal patterns in media data, focusing on detecting manipulated audio, video, and image content.
    b. Integrate feature extraction methods that identify anomalies in deepfake media, such as

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

inconsistent lip movements, unnatural blinking patterns, or irregular audio-visual synchronization.
    c. Ensure adaptability to counter evolving deepfake technologies and emerging manipulation techniques.

4. Integrate All Modules into a Unified Framework:
    a. Combine sentiment analysis, fake news detection, and deepfake identification into a cohesive and user-friendly platform.
    b. Ensure that the framework is interoperable, modular, and scalable, capable of processing large volumes of data in real- time across diverse social media platforms.
    c. Provide APIs or interfaces for seamless integration into existing systems or applications in industry and governance.

5. Enhance Trust and Transparency in Digital Communication:
    a. Develop tools that empower industries to protect their reputation by identifying misinformation and manipulated media in real time.
    b. Assist governments in combating misinformation campaigns and ensuring public safety and informed policy-making.
    c. Foster ethical online behavior by providing academia and researchers with a platform for studying and mitigating digital threats.

METHODOLOGY

The project follows a structured approach to achieving sentiment analysis, fake news detection, and deepfake recognition. The methodology is divided into five primary stages:
1. Data Collection & Pre-processing
2. Sentiment Analysis using VADER
3. Fake News Detection using Machine Learning
4. Deepfake Recognition using LSTM Networks
5. Integration & Real-time Monitoring

1. Data Collection & Preprocessing: This stage involves collecting and preprocessing textual and media data required for sentiment analysis, fake news detection, and deepfake recognition.
    a. Data Collection:Gather textual data from social media (tweets, posts, comments).Collect news articles and headlines from verified and unverified sources.Obtain deepfake and real video datasets from publicly available sources (FaceForensics++, Celeb-DF, etc.).
    b. Data Cleaning & Preprocessing:Remove stopwords, special characters, emojis, and URLs from text.Convert text to lowercase and tokenize sentences.Convert videos into individual frames and extract key features.
    c. Feature Extraction: Use TF-IDF vectorization for text-based feature extraction. Apply OpenCV and deep learning techniques to extract image/video features.
    d. Data Storage: Store pre-processed textual data in a structured database (MongoDB, Firebase, or PostgreSQL).Store media files in a cloud storage service for accessibility.

2. Sentiment Analysis using VADER: This stage involves analyzing text data (e.g., social media posts, news headlines) to determine sentiment polarity: positive, negative, or neutral. The VADER (Valence Aware Dictionary and Sentiment Reasoner) lexicon is used because it is optimized for short, informal text, making it ideal for social media content. Input Preprocessed Text Data: Load the cleaned and tokenized text data from the database. Apply VADER Sentiment Analysis: Calculate sentiment scores using the VADER lexicon. Classify text into Positive, Negative, or Neutral categories based on compound sentiment scores. Store Sentiment Results: Save the sentiment classification results in the database for further analysis. Generate Insights & Visualizations: Analyze trends and generate graphs or statistical summaries to track sentiment shifts.
3. Fake News Detection using Machine Learning: Fake news detection is crucial for identifying misleading, biased, or fabricated information. This methodology employs machine learning techniques to classify news articles or headlines as real or fake based on textual patterns and linguistic features.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

a. Input Preprocessed News Data: Load the cleaned and tokenized news articles or headlines.
b. Feature Extraction: Convert text into numerical features using TF-IDF (Term Frequency - Inverse Document Frequency).Train Machine Learning Model
c. Test and Evaluate Model: Split the dataset into training (80%) and testing (20%).Evaluate model performance using accuracy, precision, recall, and F1-score.Classify News Articles using the trained model to classify new articles as Fake or Real.
d. Store & Display Results: Store results in a database for real-time monitoring. Provide a dashboard for visualization.

4. Deepfake Recognition using LSTM Networks: Deepfake recognition is a critical component of this project, designed to detect synthetic media, including manipulated videos, images, and audio. This stage employs Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), to identify anomalies in deepfake content. The methodology consists of the following key steps:
    a. Data Collection: Gathering real and synthetic media datasets.
    b. Preprocessing: Transforming media into structured data (frames, spectrograms).
    c. Feature Extraction: Identifying deepfake-specific patterns.
    d. LSTM Model Training: Training the neural network to detect anomalies.
    e. Model Evaluation: Testing model accuracy and fine- tuning.
    f. Deployment: Integrating the system for real-time deepfake detection.
5. Deepfake Recognition using LSTM Networks: Deepfake recognition is a critical component of this project, designed to detect synthetic media, including manipulated videos, images, and audio. This stage employs Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), to identify anomalies in deepfake content. The methodology consists of the following key steps:
    a. Data Collection: Gathering real and synthetic media datasets.
    b. Preprocessing: Transforming media into structured data (frames, spectrograms).
    c. Feature Extraction: Identifying deepfake-specific patterns.
    d. LSTM Model Training: Training the neural network to detect anomalies.
    e. Model Evaluation: Testing model accuracy and fine- tuning.
    f. Deployment: Integrating the system for real-time deepfake detection.

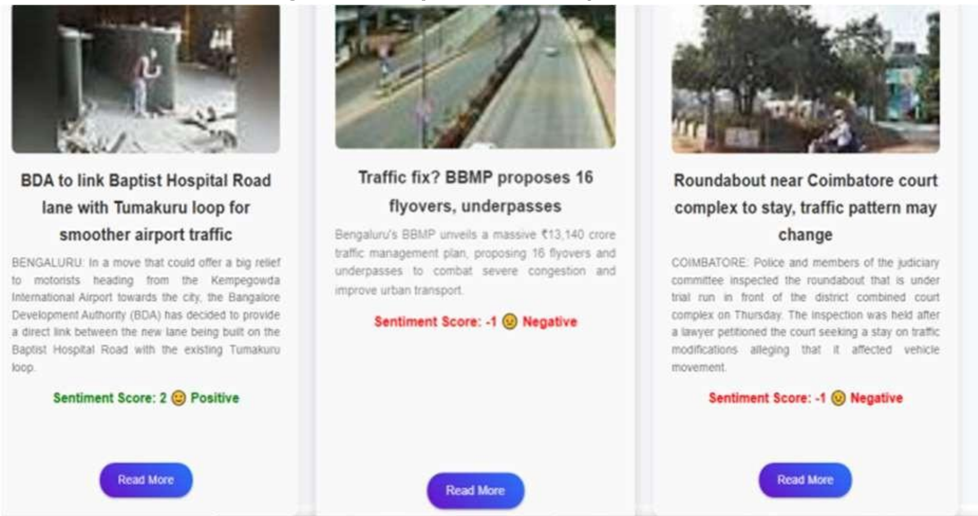
6. Integration & Real-Time Monitoring: The final stage of the project focuses on integrating all components—sentiment analysis, fake news detection, and deepfake recognition—into a unified system with real-time monitoring capabilities. This ensures seamless operation, efficient data processing, and continuous tracking of digital content authenticity.
    a. System Architecture & API Development: Backend API using FastAPI and REST & WebSocket communication
    b. Data Stream Processing: Streaming from social media/news APIs and Using Kafka/RabbitMQ for data flow.
    c. 3. Model Deployment & Inference: Docker Kubernetes for scalable deployment TensorFlow Serving /TorchServe for inference.
    d. Real-Time Dashboard:Built using React.js and Visualizes sentiment trends, fake news alerts, deepfake detections.
    e. Continuous Model Improvement:User feedback for misclassifications and MLOps tools (MLflow) for retraining.
    f. Final System Outcome: A real-time AI-powered digital content monitoring system.

RESULT AND ANALYSIS

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

A. Sentimental analysis score:

Fig.3.1: Result of sentimental analysis.



Feature	Z. Wang[2]	K. Fujihira[4]	Our Output
Sentimental key words	Yes	Yes	Yes
Sentimental score	No	No	Yes

Table 3.1: Comparing the results with existing papers.

The system will search for the top 10 news from the Microsoft binge and analyze the news. Based on the keywords present in the news it is going to predict the sentimental score of that particular news.

B. Fake news result:

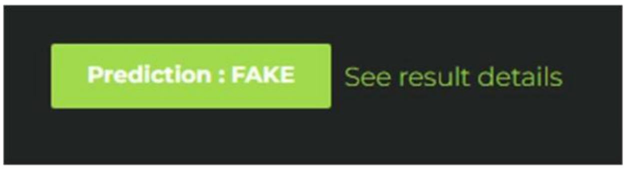
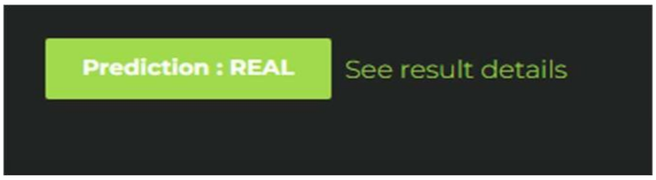


Fig.3.2: Result of Fake news detection



iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Fig.3.3: Result of Fake news detection.

Feature	Realor fakedetection	Accuracy
Jain and A. Kasbe [5]	Yes	97.3%
Our Output	Yes	98.1%

Table 3.2: Comparing the results of fake news detection with existing papers.

C. Deep Fake reorganization:

Fig.3.4: Result of deep fake reorganization.

Feature	Real or fake detection	Confidenceno.
D. Pan[1]	Yes	No
Our Output	Yes	Yes

Table 3.1: Comparing the results with existing papers

The system is feed with an image or a video, the algorithm checks for the facial points and shows the result as real or fake. It also shows the confidence of the result.

CONCLUSION

The Sentiment Analysis, Fake News Detection, and Deep fake Recognition System represents a significant advancement in AI-driven content analysis. It effectively identifies sentiment in text, detects misinformation, and recognizes deep fake media, addressing crucial challenges in digital communication and content authenticity. By leveraging Natural Language Processing (NLP) for sentiment analysis, machine learning models for fake news classification, and deep learning techniques (LSTMs, CNNs, GAN detection algorithms) for deep fake recognition, the system provides a comprehensive solution for individuals, businesses, and policymakers. The system's strengths include its ability to process large-scale data, provide real-time monitoring, and deliver high accuracy in classification tasks. However, challenges such as adversarial AI techniques, evolving deep fake methods, and misinformation tactics necessitate continuous improvements and refinements. This project demonstrates high scalability and adaptability, making it a valuable tool for media

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

agencies, cybersecurity firms, government organizations, and social media platforms in their efforts to combat digital misinformation and synthetic media manipulation.

REFERENCE

[1] *Pan, L. Sun, R. Wang, X. Zhang and R. O. Sinnott*, "Deepfake Detection through Deep Learning," 2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT), Leicester, UK, 2020, pp. 134-143, doi: 10.1109/BDCAT50828.2020.00001.

[2] *Z. Wang, C. S. Chong, L. Lan, Y. Yang, S. Beng Ho and J. C. Tong*, "Fine-grained sentiment analysis of social media with emotion sensing," 2016 Future Technologies Conference (FTC), San Francisco, CA, USA, 2016, pp. 1361-1364, doi:10.1109/FTC.2016.7821783.

[3] *Marutho, Muljono, S. Rustad and Purwanto*, "Sentiment Analysis Optimization Using Vader Lexicon on Machine Learning Approach," 2022 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, Indonesia, 2022,pp.98103,doi:10.1109/ISITIA56226.2022.98553

[4] *K. Fujihira and N. Horibe*, "Multilingual Sentiment Analysis for Web Text Based on Word to Word Translation," 2020 9th International Congress on Advanced Applied Informatics (IIAI-AAI), Kitakyushu, Japan, 2020, pp. 74-79, doi:10.1109/IIAIAAI50415.2020.00025.

[5] *Jain and A. Kasbe*, "Fake News Detection," 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2018, pp. 1-5, doi: 10.1109/SCEECS.2018.8546944.