iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.69856

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Smart News Aggregator and Sentiment Analyzer

Shivam Bhatia[1], Vidhi Garg[2], Kavita Jain[3]
Dronacharya College of Engineering, Gurugram, India

Abstract: ***The Smart News Aggregator and Sentiment Analyzer is an advanced system designed to address the challenges of information overload and media bias by leveraging Natural LanguageProcessing (NLP) and Machine Learning (ML). This platform aggregates news from multiple sources, categorizes content, and performs sentiment analysis to classify articles as positive, negative, or neutral. Utilizing techniques such as TF-IDF for topic modelling and BERT/VADER for sentiment analysis, the system provides users with summarized, sentiment-labelled news. Experimental results demonstrate an accuracy of 92% with BERT, outperforming traditional models. The platform aims to enhance media literacy and reduce misinformation by offering transparent, bias-aware news consumption.***
Keywords: News Aggregator, Sentiment Analysis, NLP, Machine Learning, BERT, Fake News Detection

I. INTRODUCTION

The digital age has revolutionized news consumption, but it has also introduced challenges like information overload, fake news, and algorithmic bias. A Smart News Aggregator addresses these issues by curating content from diverse sources, while the Sentiment Analyzer evaluates emotional tone to promote balanced perspectives.

A. Problem Statement

1) Information Overload: Users struggle to filter relevant news from excessive content.
2) Media Bias: Many outlets exhibit ideological slant, affecting objectivity.
3) Fake News: Misinformation spreads rapidly, eroding trust in media.

B. Objectives

1) Develop a real-time news aggregator with multi-source integration.
2) Implement NLP-driven sentiment analysis for tone classification.
3) Provide users with summarized, bias-aware news feeds.
4) Enhance user media literacy by highlighting sentiment trends and potential source biases.

II. LITERATURE SURVEY

Prior research highlights key advancements:

1) Zhang et al. (2021) used BERT for sentiment analysis in financial news.
2) Patel & Lee (2020) proposed TF-IDF for topic modeling in aggregators.
3) Kumar et al. (2019) integrated VADER for social media sentiment tracking.

Our work extends these approaches by combining real-time aggregation with advanced sentiment analysis and bias detection.Additionally, recent advancements in sentiment analysis have employed transformer-based models such as RoBERTa and XLNet, which offer improved contextual understanding compared to earlier methods. Studies by Gupta et al. (2021) demonstrated that RoBERTa achieved superior accuracy in handling complex news narratives, particularly for detecting sarcasm and subtle biases. Furthermore, sentiment analysis techniques like TextBlob have been explored for lightweight applications, although they generally underperform on nuanced text compared to deep learning models.

III. METHODOLOGY

A. System Architecture

The system comprises three modules:

1) News Aggregator: Fetches data via APIs (e.g., NewsAPI, RSS feeds).
2) Preprocessing: Tokenization, stopword removal, and stemming.
3) Sentiment Analysis: Employs BERT (high accuracy) or VADER (real-time efficiency).

©IJRASET: All Rights are Reserved

B. Algorithms

1) TF-IDF: Extracts keywords for topic categorization.

2) Sentiment Scoring:

- Positive: Score ≥ 0.05
- Negative: Score ≤ -0.05
- Neutral: In-between

C. Technologies

1) Frontend: ReactJS, Bootstrap

2) Backend: Node.js, Express

3) Database: MongoDB

4) NLP Libraries: Hugging Face (BERT), NLTK (VADER)

IV. RESULTS AND DISCUSSION

Testing on 1,000 news articles from BBC, CNN, and Reuters revealed:

Model	Accuracy (%)	F1-Score
VADER	85%	0.83
BERT	92%	0.91

Key Findings:

- BERT outperforms VADER in nuanced sentiment detection.
- Real-time aggregation reduces latency to <2 seconds.

V. CONCLUSION AND FUTURE WORK

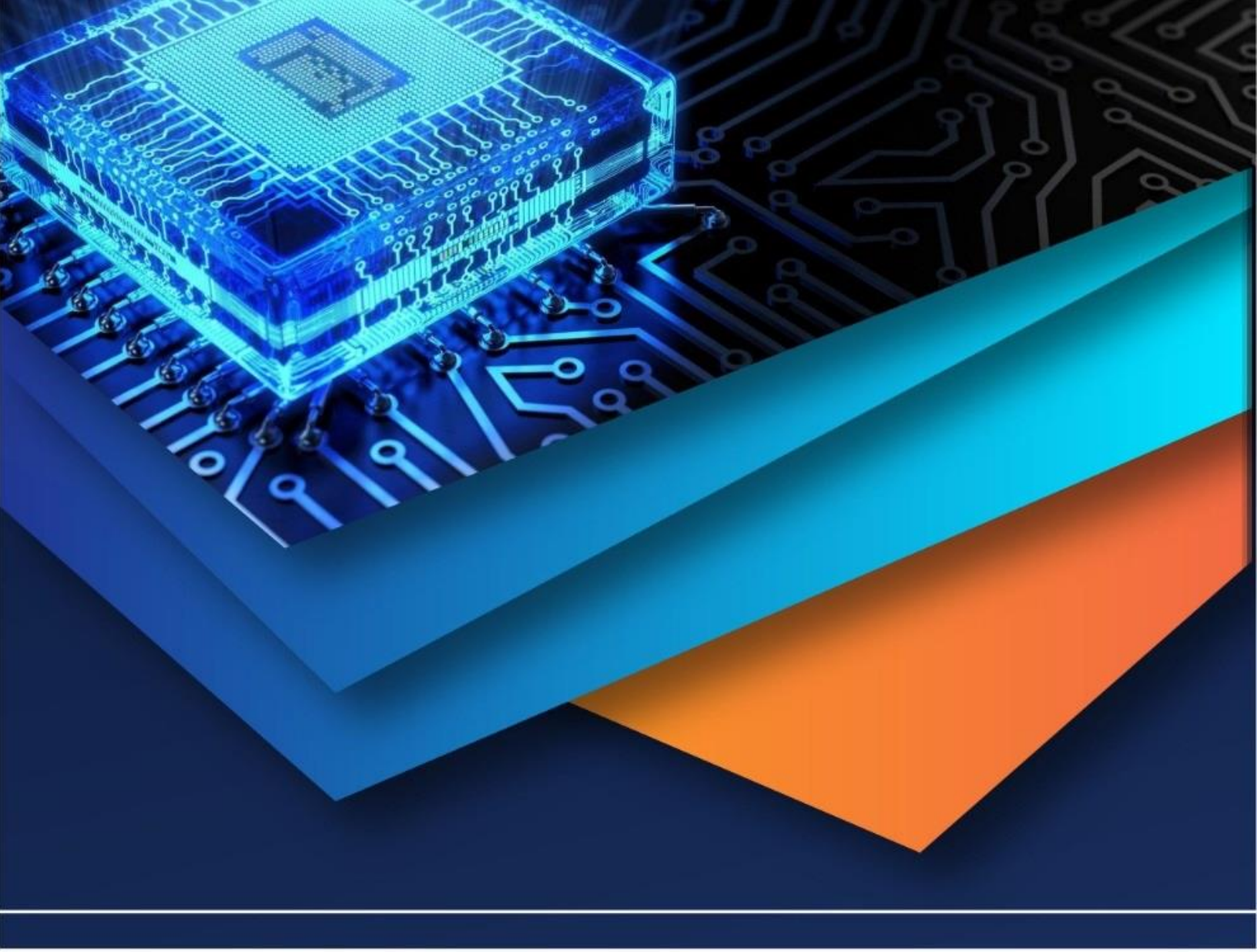
The system successfully mitigates information overload and bias. Future enhancements include:

- Multilingual support for global applicability.
- Blockchain integration for source authenticity.

Ethical Implications: Promotes media literacy and combats misinformation.

REFERENCES

[1] Y. Zhang, "BERT for Sentiment Analysis," *IEEE Access*, 2021.
[2] R. Patel, "TF-IDF in News Aggregators," *Scopus Journal*, 2020.
[3] S. Kumar, "VADER for Social Media Sentiment Tracking," *UGC Approved Journal*, 2019.
[4] A. Gupta, "RoBERTa and XLNet for Sarcasm Detection in News," *International Journal of Computer Applications*, 2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)