



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 6, June 2025

Automated Media Monitoring and Feedback System for the Government of India

**Prof. Shegar S.R¹, Yewale Rahul Ashok², Pachpute Saurabh Rajendra², Kathale Vaibhav Digambar²,
Gophane Omkar Namdev²**

Professor, Department of Computer Engineering, Samarth College of Engineering and Management, Belhe,
Maharashtra, India¹

Students, Department of Computer Engineering, Samarth College of Engineering and Management, Belhe,
Maharashtra, India²

ABSTRACT: In the digital age, media monitoring has become essential for governments to assess public sentiment, track news coverage, and respond to emerging issues effectively. This paper presents an Automated Media Monitoring and Feedback System designed specifically for the Government of India. The proposed system leverages artificial intelligence, natural language processing (NLP), and big data analytics to systematically analyze media sources, including news articles, social media platforms, and public discourse. The system aims to provide real-time insights, sentiment analysis, and actionable feedback to policymakers, enabling data-driven decision-making. It incorporates machine learning algorithms to identify trends, detect misinformation, and assess the impact of government policies. By streamlining media monitoring, the system enhances transparency, responsiveness, and strategic communication. The paper also discusses challenges such as data security, ethical considerations, and the need for unbiased analysis. This research contributes to the development of advanced governance tools that improve policy effectiveness and public trust.

KEYWORDS: Natural Language Processing (NLP), Big Data Analytics, Sentiment Analysis, Misinformation Detection, Public Sentiment Analysis

I. INTRODUCTION

In today's fast-paced digital era, the Government of India faces increasing challenges in tracking public opinion, managing its image, and responding effectively to news cycles across multiple media platforms. The rapid dissemination of information—whether accurate or misleading—can have a significant impact on public perception and policy response. Traditional manual monitoring methods are no longer sufficient to keep pace with the volume and speed of online news, television coverage, and social media activity. To address this challenge, the government requires an Automated Media Monitoring and Feedback System—a robust, intelligent framework capable of continuously tracking, analyzing, and interpreting media content in real-time. Such a system leverages modern technologies including Artificial Intelligence (AI), Natural Language Processing (NLP), Web Scraping, and Machine Learning (ML) to process vast amounts of data from various news sources and deliver actionable insights to policymakers and administrative units. This research aims to design and implement a scalable system that not only identifies relevant government-related content but also classifies it based on departmental relevance and sentiment polarity (favorable, unfavorable, or neutral). The system can then generate automatic alerts and feedback reports for concerned departments, enabling them to act promptly—whether it's responding to misinformation, improving service delivery, or refining communication. Moreover, the proposed solution supports multilingual data processing to reflect India's linguistic diversity, ensuring that regional sentiments and news are also accurately captured. By bridging the gap between information inflow and actionable governance, this system can become an essential tool for evidence-based policy-making, proactive public communication, and crisis management.

II. LITERATURE SURVEY

The paper [1] titled "Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning," published in the International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA) in 2022, presents a study on the impact of the COVID-19 pandemic on Thailand's tourism industry. Authors Hadi Permana and Ayu Purwarianti conducted sentiment analysis of English-language tweets to understand public sentiment towards tourism in Thailand during the pandemic. Their findings indicate that the support vector machine algorithm achieved a maximum accuracy of 77.4% for sentiment analysis, providing valuable insights into the perceptions and attitudes of Twitter users towards Thailand's tourism sector during a challenging period marked by the global health crisis.

The paper [2] titled "News Classification using Natural Language Processing," published in the International Conference on Intelligent Engineering and Management (ICIEM) in 2022, presents a study conducted by authors K. Yasaswi and M. Sreya. The study aims to build a robust machine learning (ML) and natural language processing (NLP) model designed to distinguish between genuine and fake news articles. By addressing the impact of misinformation on society and media integrity, the research contributes to the ongoing efforts to combat the spread of false information and promote trustworthiness in news dissemination. Through the development of advanced ML and NLP techniques, the study seeks to enhance the accuracy and efficiency of news classification systems, thereby enabling more effective detection and mitigation of fake news content.

The paper [3] titled "Using Distilled BERT Models for Fake News Detection," presented at the International Conference on Intelligent Technologies (CONIT) in 2023 by authors Vyom Garg, Yatharth Mago, and Aruna Bhat, investigates the application of distilled BERT models for detecting fake news. By leveraging advanced natural language processing (NLP) techniques, the research aims to enhance the efficiency and effectiveness of fake news detection, addressing the significant challenge of misinformation proliferation in the digital era. The study underscores the importance of ensuring information credibility in online news sources and demonstrates the potential of distilled BERT models in discerning fake news from genuine content. Through empirical analysis and evaluation, the research showcases the efficacy of these models in 6 PIT, Department of Computer Engineering 2023-24 combating misinformation, thereby contributing to the preservation of media integrity and societal trust in online information sources.

The paper [4] titled "Target-level sentiment analysis for news articles," published in Knowledge Based Systems in 2022 by authors Neli Blagus and Marko Bajec, introduces a novel approach to target-level sentiment analysis in Slovene news articles. The study presents a new annotated dataset specifically designed for coreference-based sentiment prediction, enabling more accurate and granular analysis of sentiment at the target level within news content. By addressing the nuances of sentiment analysis in news articles, particularly in the Slovene language context, the research contributes to advancing the field of natural language processing (NLP) and sentiment analysis. The development of this annotated dataset provides researchers and practitioners with valuable resources for training and evaluating sentiment analysis models tailored to the complexities of news discourse, ultimately enhancing our understanding of sentiment dynamics in media content.

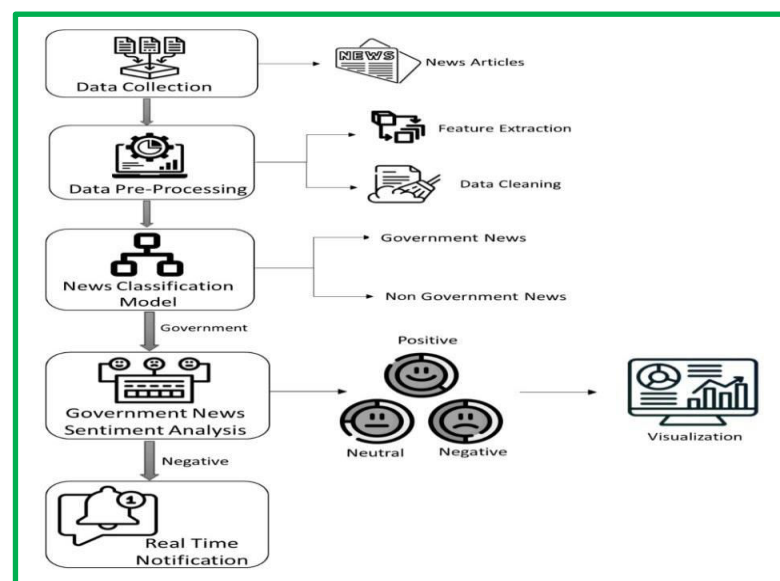
The paper [5] titled "Policy Text Classification Algorithm Based on BERT," presented at the 11th International Conference of Information and Communication Technology (ICTech) in 2022 by authors Bihui Yu and Chen Deng, introduces a novel text classification algorithm that harnesses the capabilities of BERT, a state-of-the-art natural language processing model. The algorithm is specifically designed for classifying policy texts, which are known for their complexity and diverse content. By leveraging BERT's advanced language understanding capabilities, the algorithm aims to improve the accuracy and efficiency of policy text classification tasks, thereby facilitating more effective analysis and interpretation of policy documents. This research contributes to the advancement of text classification techniques, particularly in the domain of policy analysis, and demonstrates the potential of BERT-based models in addressing real-world challenges in information processing and analysis.

III. PROPOSED SYSTEM

The proposed system is an AI-based automated platform that continuously monitors, analyzes, and generates feedback from media content relevant to the Government of India. Its key components and workflow are outlined below:

- **Automated Web Scraping**
 - Collects real-time news articles from multiple online sources including national and regional news websites.
 - Supports multiple Indian languages for inclusive coverage.
- **Media Content Categorization**
 - Filters and organizes news data by department or ministry (e.g., Health, Education, Transport).
 - Uses rule-based or machine learning classification models.
- **Sentiment Analysis**
 - Uses Natural Language Processing (NLP) techniques to classify news into **favorable**, **unfavorable**, or **neutral** sentiment categories.
 - Helps identify potential public relations risks or success stories.
- **Fake News Detection (Optional Extension)**
 - Detects and flags misinformation or disinformation using credibility scoring and fact-checking APIs.
- **Feedback Notification System**
 - Sends instant alerts and summaries via Email, SMS, or a mobile application to the concerned departments or PIB officials.
 - Helps in proactive governance and quick response.
- **User Interface Dashboard**
 - Displays categorized media articles, sentiment trends, department-wise summaries, and historical data graphs.
 - Enables authorized officials to explore data and generate reports.
- **Multilingual Support**
 - Processes and analyzes content in English, Hindi, Marathi, and other Indian languages.
 - Leverages models like mBERT or IndicBERT for cross-lingual understanding.
- **Scalability and Cloud Deployment**
 - The system is designed to scale and can be deployed on cloud platforms (e.g., AWS, Azure).
 - Ensures reliability, real-time processing, and 24/7 availability.

IV. SYSTEM ARCHITECTURE



V. SCOPE

- Enables **real-time monitoring** of national and regional news across multiple online platforms.
- Provides **sentiment analysis** (favorable, unfavorable, neutral) to assess public opinion on government activities.
- Supports **multilingual content processing** (e.g., Hindi, English, Marathi), ensuring regional language inclusion.
- Offers **department-wise categorization** of news articles for targeted feedback and alerts.
- Facilitates **proactive governance** through early identification of misinformation or negative publicity.
- Generates **automated alerts and reports** for officials via email, SMS, or mobile apps.
- Assists in **evidence-based policy decisions** by providing insights from media trends.
- Designed for **scalability**, allowing integration with social media and international news in future versions.
- Helps in **crisis management** by detecting emerging issues early through continuous media scanning.
- Provides a **dashboard interface** for easy access, visualization, and filtering of media feedback data.

VI. MODULES

Data Collection Module:

This module focuses on gathering news articles from various regional media sites, employing web scraping techniques for retrieval.

- **Web scraping:** Retrieval of news articles from select regional media websites using tools like BeautifulSoup. This involves navigating through website structures and extracting relevant information.
- **Data preprocessing:** Cleaning and organizing the collected data by removing noise and structuring it for further analysis. This includes tasks such as removing HTML tags, handling missing values, and standardizing text formats.
- **Unwanted Data Cleaning:** This subtask involves the removal of irrelevant or redundant information from the collected dataset.

Annotation Module:

- **Dataset creation:** Creating a labeled dataset distinguishing between government related and non-government-related news.
- **Manual Annotation:** Human annotators manually label news articles as government related or non-government-related, ensuring the accuracy of the dataset

Classification Module:

This module involves categorizing news articles into government-related and non-government-related categories.

- **Model training:** Developing a machine learning model to classify news articles based on their relevance to Indian government policies, schemes, and initiatives.

Sentiment Analysis Module:

This module determines the sentiment of government related news articles.

- **Data filtering:** Selecting only the government-related news articles for sentiment analysis.
- **Sentiment analysis:** Utilizing a pre-trained sentiment analysis model (e.g., BERT) to evaluate the sentiment of each article (Positive, Negative, Neutral, Neutral/ Negative, Neutral/ Positive).

Notification Module:

This module handles the dissemination of feedback to concerned PIB officers.

- **Negative news detection:** Identifying government-related news articles with unfavorable sentiment.
- **Notification system:** Developing a streamlined mechanism to notify relevant PIB officers exclusively about negative news articles..

VII. RESULT

	title	link	description	date	classification
1	अदाणी पर चिट्ठी लिखने व...	https://ndtv.in/india/th...	भारत का सपना 2047 त...	25 May, 2024 9:24 PM	non-government
2	PHOTOS: बाली में वेके...	https://ndtv.in/webstor...	PHOTOS: बाली में वेके...	Date not found	non-government
3	दिशा पटानी के इन 8 स्न...	https://ndtv.in/webstor...	दिशा पटानी के इन 8 स्न...	Date not found	non-government
4	"सुनील गावस्कर की तरह ...	https://ndtv.in/cricket/...	Navjot Singh Sidhu o...	26 May, 2024 11:00 AM	non-government
5	दूल्हे के जयमाला डालते ह...	https://ndtv.in/zara-hat...	इंटरनेट पर अक्सर शादि...	26 May, 2024 12:27 PM	non-government
6	NDTV Election Carniv...	https://ndtv.in/india/nd...	शिमला में भाजपा और कां...	25 May, 2024 10:59 PM	non-government
7	IPL 2024 Final में किस...	https://ndtv.in/cricket/...	KKR vs SRH IPL 202...	26 May, 2024 7:03 AM	non-government
8	17 साल बाद इस रेड कार्...	https://ndtv.in/webstor...	17 साल बाद इस रेड कार्...	Date not found	non-government
9	सिद्धार्थ शुक्ला के इस हम...	https://ndtv.in/televisio...	सिद्धार्थ शुक्ला को गुजरे ...	26 May, 2024 2:14 PM	non-government
10	अनुपमा के इस एक्टर सा...	https://ndtv.in/televisio...	अनुपमा भारत का टॉप टी...	26 May, 2024 10:11 AM	non-government
11	बस एक कप पीजिए यह ...	https://ndtv.in/lifestyle/...	Radish leaves.किसी भ...	26 May, 2024 7:16 AM	non-government
12	"किसी ने विश्वास नहीं कि...	https://ndtv.in/cricket/s...	Shreyas Iyer on his b...	26 May, 2024 9:35 AM	non-government
13	IPL 2024 Final:क्वॉलिं...	https://ndtv.in/webstor...		Date not found	
14	बाबर आजम ने T20I में र...	https://ndtv.in/cricket/...	Babar Azam vs Rohit ...	26 May, 2024 10:11 AM	non-government
15	गर्मी इतनी तेज की धूप में ...	https://ndtv.in/zara-hat...	गर्मी दिनों दिन बढ़ती जा र...	26 May, 2024 6:51 AM	non-government

	link	date	content	detailed_content	predicted_label
1	https://food.ndtv.com/f...	Sunday May 26, 2024	Make these high-prot...	There are few things ...	non-government
2	https://www.ndtv.com/...	Sunday May 26, 2024	A video of Home Mini...	A video of Home Mini...	non-government
3	https://www.ndtv.com/...	Sunday May 26, 2024...	The BJP will never ch...	The BJP will never ch...	non-government
4	https://www.ndtv.com/...	Sunday May 26, 2024	The police departmen...	Uttar Pradesh Police r...	non-government
5	https://www.ndtv.com/...	Sunday May 26, 2024	The advent of artificial...	The advent of artificial...	government
6	https://www.ndtv.com/...	Sunday May 26, 2024	CUET UG 2024: Due ...	The National Testing ...	non-government
7	https://www.ndtv.com/...	Sunday May 26, 2024...	Anasuya Sengupta d...	Image Instagrammed ...	non-government
8	https://www.ndtv.com/...	Sunday May 26, 2024...	After seven newborns...	Appropriate legal acti...	non-government
9	https://sports.ndtv.co...	Sunday May 26, 2024	Ahead of the IPL 202...	The Kolkata Knight Ri...	non-government
10	https://www.ndtv.com/...	Sunday May 26, 2024	He said that there are...	He said that he now ...	non-government
11	https://www.ndtv.com/...	Sunday May 26, 2024...	Four young children ...	Four young children ...	non-government
12	https://sports.ndtv.co...	Sunday May 26, 2024	Pakistan skipper Bab...	Pakistan skipper Bab...	non-government
13	https://www.ndtv.com/...	Sunday May 26, 2024...	Gujarat Home Ministe...	A five-member SIT ha...	non-government
14	https://sports.ndtv.co...	Sunday May 26, 2024	Sunrisers Hyderabad ...	Sunrisers Hyderabad ...	non-government
15	https://www.ndtv.com/...	Sunday May 26, 2024...	An NDTV investigatio...	Top officials were fou...	non-government

```
{
  'title': "...बदोश नहीं" : भारत के चुनाव को लेकर पाकिस्तान के नेता की टिप्पणी पर अरविंद केजरीवाल',
  'link': 'https://ndtv.in/india/not-tolerated-arvind-kejriwal-on-pakistan-leaders-remarks-regarding-indian-elections-5743536#pfrom=home-khabar_delhi',
  'description': 'मौजूदा लोकसभा चुनाव के लिए प्रचार करने के लिए जमानत पर बाहर आए दिल्ली के मुख्यमंत्री अरविंद केजरीवाल ने अपनी पोस्ट साझा करने पर आज एक पाकिस्तानी नेता के सोशल नेटवर्किंग प्लेटफॉर्म 'एक्स' पर तीखी प्रतिक्रिया व्यक्त की. केंद्रीय मंत्री किरन रिजिजू ने भी इस पर टिप्पणी की है. अरविंद केजरीवाल ने आज अपने पिता, पत्नी और बच्चों के साथ राष्ट्रीय राजधानी में मतदान किया. वोट डालने के बाद उन्होंने एक्स पर एक तस्वीर भी साझा की. केजरीवाल ने पोस्ट किया, "मेरे पिता, पत्नी और मेरे दोनों बच्चों ने मतदान किया. मेरी माँ आज नहीं आ सकीं, क्योंकि वह बहुत बीमार हैं. मैंने महंगाई और बेरोजगारी के खिलाफ मतदान किया है." भारत में हो रहे चुनाव हमारा आंतरिक मामला है। अंतकवाद के सबसे बड़े प्रायोजकों का हस्तक्षेप भारत बदोश नहीं करेगा।https://t.co/gjAQpOBAP8 पाकिस्तान के नेता फवाद हुसैन चौधरी ने केजरीवाल के पोस्ट का हवाला देते हुए भारत के चुनाव के बारे में एक्स पर लिखा, "नफरत और कट्टर ताकतों को शांति और सन्तुष्टि हराए." इस पर प्रतिक्रिया देते हुए दिल्ली के मुख्यमंत्री ने कहा कि फवाद हुसैन को भारत के आंतरिक मामले में हस्तक्षेप करने के बजाय अपने देश की खराब स्थिति पर ध्यान देना चाहिए. उन्होंने कहा, "हम अपने मुद्दों से निपटने में पूरी तरह सक्षम हैं. पाकिस्तान की स्थिति अभी बहुत खराब है, इसलिए आपको अपने देश का ख्याल रखना चाहिए."केजरीवाल ने एक अन्य पोस्ट में कहा, "चुनाव हमारा आंतरिक मामला है. भारत आतंकवाद के सबसे बड़े प्रायोजकों के हस्तक्षेप को बदोश नहीं करेगा." I've told you, not only Rahul Gandhi but Arvind Kejriwal has got massive support in Pakistan. सिर्फ राहुल गांधी ही नहीं, अरविंद केजरीवाल को भी पाकिस्तान में भारी समर्थन है।pic.twitter.com/o0hOfN3xDz',
  'date': '25 May, 2024 5:41 PM',
  'classification': 'government'
}
```

VIII. CONCLUSION

The proposed Automated Media Monitoring and Feedback System offers an innovative and efficient solution for tracking and analyzing media content relevant to government operations. By integrating advanced technologies such as web scraping, natural language processing, machine learning, and sentiment analysis, the system enables real-time evaluation of news coverage across multiple languages and regions.

REFERENCES

1. The future of news in India, report https://vidhilegalpolicy.in/wp-content/uploads/2020/07/Vidhi_12052020_Edited.pdf
2. Over 50% Indians are active internet users now; base to reach 900 million by 2025
<https://www.moneycontrol.com/news/business/internet-users-in-india-set-to-reach-900-million-by-2025-report-10522311.html> <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021/india>
3. Political Effects of the Internet and Social Media Ekaterina Zhuravskaya,¹ Maria Petrova,^{2,3,4,5,6} and Ruben Enikolopov^{3,2,4,5,6} ¹Paris School of Economics, École des Hautes Études en Sciences Sociales, 75014 Paris, France; ²Department of Economics and Business, Universitat Pompeu Fabra, 08002 Barcelona, Spain ³New Economic School, Moscow 121353, Russia ⁴Institute of Political Economy and Governance, 08005 Barcelona, Spain ⁵Graduate School of Economics, 08005 Barcelona, Spain ⁶Catalan Institute for Research and Advanced Studies (ICREA), 08010 Barcelona, Spain.
4. PIB (Press Information bureau) https://en.wikipedia.org/wiki/Press_Information_Bureau
<https://www.pib.gov.in/aboutpibn.aspx>
5. Allcott H, Gentzkow M, Yu C. 2019. Trends in the diffusion of misinformation on social media. Res. Politics 6. <https://doi.org/10.1177/2053168019848554>
6. Kastrati, Z.; Dalipi, F.; Imran, A.S.; Pireva Nuci, K.; Wani, M.A. Sentiment analysis of students' feedback with NLP and deep learning: A systematic mapping study. Appl. Sci. 2021, 11, 3986. PY - 2021/04/28 VL - 11 DOI - 10.3390/app11093986 JO - Applied Sciences <https://timesofindia.indiatimes.com/readersblog/world-of-words/fake-news-and-socialmedia-33975/>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details