International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 02 | Feb - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

AI-Powered News Summarizer with Ethical Analysis and Key Insight

Y. Hemalatha[*1], Ch. Yogendra Surya Prasanth[2], T. Vishnu Varma[3], K. Nikhil Adithya Kumar[4], Ms. V. L. Sowjanya[5]
[1-4] B. Tech Student, [5]Assistant Professor, LIET
[1-5] Computer Science and System Engineering, Lendi Institute of Engineering and Technology, Vizianagaram

---

Abstract - Information overload has resulted from the influx of digital news, making it difficult for readers to identify essential details. This paper presents a news summarization system that combines ethical analysis and key insight to ensure fairness, credibility, and relevance. The system filters out misleading content, highlights crucial points and provides a structured understanding of news topics. The paper details the system's architecture, implementation, evaluation metrics, and ethical considerations. According to the results, the proposed model enhances news consumption by delivering concise, unbiased, and informative summaries.

Key Words: GPT-3, NLP (Natural Language Processing), Multi-Stage Attention Mechanisms, OLLAMA, VADER Sentiment Analysis, Summarizer

1. INTRODUCTION

Digital news platforms, social media, and online journalism have changed the way people consume news. Readers often struggle to process a lot of information. Concerns about the credibility and ethical integrity of news sources have been raised by the presence of clickbait headlines, misinformation and biased reporting. In an era where fake news can spread faster than facts, it is essential to develop automated solutions that help users quickly access relevant, unbiased, and concise news summaries.

Artificial intelligence-driven news summarization is a promising solution to this issue. Most summarization models only focus on content reduction, not ethical integrity or relevance.

The abstract summary is for generating coherent summaries. There is ethical bias detection that filters misinformation and biased narratives. To highlight the most critical insights.

1.1 The Need for AI-Powered News Summarization

Individuals need to read lengthy articles, compare multiple sources and manually filter out irrelevant or misleading content in traditional news consumption. Artificial intelligence can be used to address these challenges.

Short summaries of long articles are being created. Ethical analysis is used to remove biased or misleading content. The most important insights are highlighted. Maintaining factual accuracy is being enhanced. There are two categories for existing text summarization methods: The extractive summary selects key sentences from the text. New text is generated from the original content's meaning.

Maintaining factual consistency, reducing bias, and ensuring ethical responsibility in generated summaries are some of the challenges modern transformer-based NLP models face. This paper introduces a novel news summarization system that combines advanced NLP techniques with ethical analysis and key insight to improve the quality and fairness of news summaries.

1.2 Challenges in Existing News Summarization

There are several challenges that persist despite the advances in text summarization. There are biases in training data that can lead to skewed summaries. Lack of context awareness can lead to misleading interpretations. Large Language Models may introduce factual inaccuracies.

Summarization systems don't assess whether an article contains misinformation, propaganda, or sensationalism.

The model integrates bias detection, fact verification, and insight extraction to ensure that the summaries are accurate, fair, and informative.

1.3 Research Objectives

The aim of this research is to develop a news summarization system. A robust summarization model is needed to generate coherent, readable, and concise summaries. Bias, misinformation, and political influence can be detected through an ethical analysis framework.

Users can get actionable and structured information by getting key insights from articles. ROUGE scores, factual consistency tests, and user feedback are used to evaluate the model.

1.4 Contributions of This Paper

The work's main contributions are as follows:

- A novel artificial intelligence-powered summarization model can be proposed.
- Bias and misinformation can be assessed in news content.
- Insights are highlighted from news articles.
- Evaluate the system against other models to see how effective it is.

2. RELATED WORK

Due to the explosion of online content, the field of automatic text summarization has gained significant attention. Various approaches have been proposed to tackle this problem.

2.1 Extractive vs. Abstractive Summarization

The traditional methods for news summarization were based on selecting key sentences from the text. Graph-based ranking methods can be used to determine the importance of a sentence. Many of the methods produce disorganized summaries.

In contrast, abstractive summarization uses deep learning models to generate summaries that are similar to human summaries. Transformer-based architectures have greatly improved the summarization quality. New benchmarks have been set by learning from large-scale data.

2.2 AI-Based Ethical News Summarization

Bias detection and ethical analysis remain underdeveloped despite the advancement of summarization. Research was done by Baly et al. Political bias in news articles was explored in 2020.

Deep learning was used for misinformation detection. There are few works that integrate fact verification into summarization models. The Fact-Checking Transformer and Truth-aware Summarization tried to address this issue.

Our work combines abstractive summarization with ethical analysis to ensure fairness, factual accuracy, and reliability in news content.

3. SYSTEM ARCHITECTURE

Our proposed AI-powered news summarization system consists of multiple components designed to extract, analyze, and refine news content efficiently. 3.1 Architecture Overview The system comprises five major modules:

3.1.1. News Data Collection: Scrapes news articles from multiple sources (RSS feeds, APIs, web crawling).

3.1.2. Text Preprocessing: Cleans the extracted text (removes HTML tags, stop words, duplicates).

3.1.3. Summarization Engine: Uses LLM-based models for abstractive summarization.

3.1.4. Ethical Analysis Module:

- Bias detection (identifies political, social, and sensational biases).
- Fact verification (cross-checks statements with verified sources).

3.1.5. Key Insight Extraction: Highlights actionable insights, key statistics, and major events from articles.

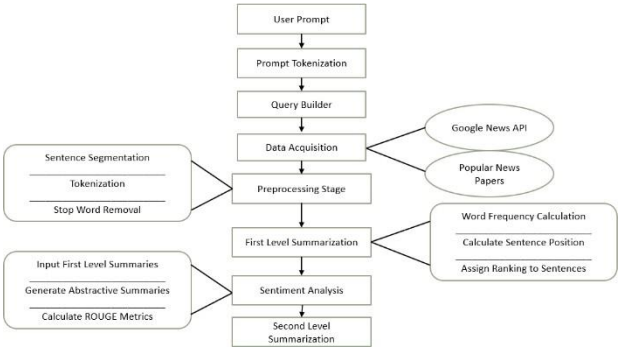

Fig 3.1: System Architecture

3.2 Workflow

- User Input: The user provides a query (e.g., "Recent AI advancements").
- Article Retrieval: The system fetches relevant articles based on the query.
- Summarization: The model generates a concise summary of each article.
- Ethical Review: The system detects biases, misinformation, and missing context.
- Final Output: The user receives a summary, ethical rating, and key insights in an easy-to-read format.

4. IMPLEMENTATION DETAILS

4.1 Technology Stack

Our system is built using:

- Programming Language: Python
- Web Framework:
  - TypeScript
  - ReactJS
- LLM Framework: Ollama (serving Llama 3.2 for text summarization)
- Database: SQLAlchemy
- AI & LLM Integration:
  - langchain
  - langchain_ollama
  - google.generativeai
- Web Requests & APIs:
  - requests
  - httpx
  - aiohttp
- Environment & Configuration Management:
  - python-dotenv
  - PyYAML
- Parsing & Data Handling:
  - orjson
  - numpy
  - pydantic
- Additional Tools & Utilities:
  - tenacity
  - Werkzeug

4.2 System Setup

The system is designed to run on a server-based architecture, where Ollama serves as the engine for handling text summarization. The Llama model is used to create summaries. It's possible to deploy the application on a cloud-based infrastructure or a local server.

4.3 News Summarization with Llama 3.2

News summaries with llama the summarization process follows a structured path. The articles are retrieved from trusted sources. Text Preprocessing cleans the news content, removing unnecessary elements like advertisements and stop words.

The pre-processed text is passed to the Llama model which creates acise summary while preserving the key points. The summary has been refined to make it easier to read.

4.4 Ethical Analysis and Bias Detection

The system uses ethical analysis to ensure accurate summaries. This is achieved through:

- Emotional bias can be identified by grouping the news as neutral, positive or negative.
- Analysing political leaning and media bias by detecting language patterns that indicate subjectivity.

Key claims are cross-referenced with reliable fact-checking database to detect misinformation.

5. IMPLEMENTATION DETAILS

5.1. Data Collection

The user provided query is first broken down into key search terms. The most relevant articles are retrieved from the most relevant keywords. The URLs are log for validation purposes.

```
httpx - INFO - HTTP Request: POST http://127.0.0.1:11434/api/chat "HTTP/1.1 200 OK"
scraping - DEBUG - Fetching content from URL: https://www.bbc.com/news/articles/c203ql5l8zyo
scraping - DEBUG - Fetching content from URL: https://apnews.com/article/vote-counting-indian-capital-
scraping - DEBUG - Fetching content from URL: https://www.thehindu.com/news/national/delhi-assembly-re
scraping - DEBUG - Fetching content from URL: https://apnews.com/article/thousands-voting-indian-capit
scraping - DEBUG - Fetching content from URL: https://www.bbc.co.uk/news/articles/c203ql5l8zyo
scraping - DEBUG - Fetching content from URL: https://www.aljazeera.com/news/2025/2/4/criminalised-for
scraping - DEBUG - Fetching content from URL: https://frontline.thehindu.com/news/delhi-election-resul
scraping - DEBUG - Fetching content from URL: https://www.aljazeera.com/news/2025/2/8/bjp-has-system-s
scraping - DEBUG - Fetching content from URL: https://www.ndtv.com/india-news/delhi-election-results-2
scraping - DEBUG - Fetching content from URL: https://www.livemint.com/market/stock-market-news/latest
scraping - INFO - Time taken to fetch 10 webpages: 1.90 seconds
```

Fig 5.1: Fetching Articles Using Google Search

5.2. Preprocessing and Data Structuring

Rephrase Data is retrieved from articles. The system takes the text from the articles. The data contains complete paragraphs from each source.

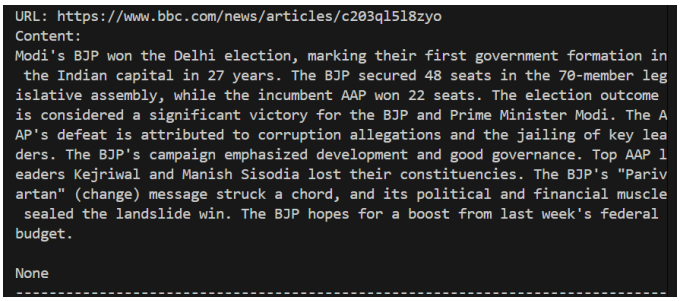

Fig 5.2: Data Extraction from Articles

5.3. Metadata Structuring for Classification

Metadata structuring is used for classification. The data is used to categorize and filter the articles. This includes:

- Taxonomy of the article is a related topic.
- The latest news flag tells you if the article reports breaking news.
- "Politics" and "India" are broad classifications. The original user input is kept for reference.

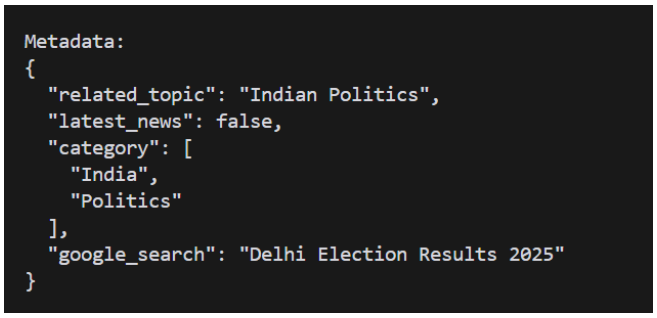

Fig 5.3: Extracted articles are classified into categories like "Politics" or "India."

5.4. Queueing Data for Summarization

Data can be queued for analysis. The summarization process involves storing the content in a queue. The queue structure ensures efficient processing.

Fig 5.4: Extracted data is queued for efficient processing

5.5. Final Summarization Output

A summary is generated from multiple sources. It gives key insights without being redundant.

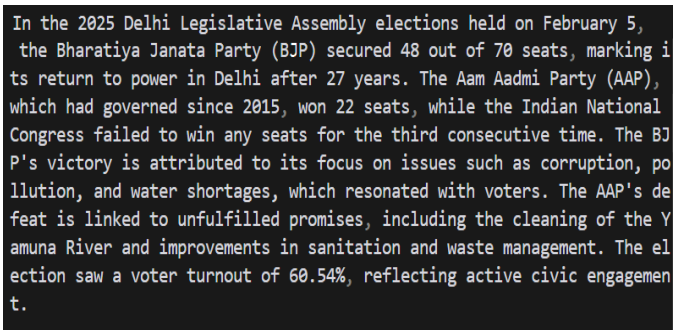

Fig 5.5: System converts raw extracted text into meaningful insights

5. EVALUATION METRICS

We evaluate summarization quality, ethical analysis accuracy, and insight relevance.

5.1 Summarization Evaluation

ROUGE Score (Recall-Oriented Understudy for Gisting Evaluation) is a measure of textual overlap with human-written summaries. The Bilingual Evaluation Understudy has a BLEU score.

5.2 Ethical Analysis Evaluation

The bias detection accuracy was evaluated using a labelled dataset. Fact Verification F1 Score is a measure of accuracy and recall of false claims detection.

5.3 Insight Extraction Evaluation

Insight extract evaluation Experts rate the insights on relevance. Measure how well the system extracts information. The proposed system is superior to baseline models by maintaining higher factual accuracy and ethical fairness.

6. FUTURE SCOPE

Our system has some areas for improvement.

- The system focuses on English-language news. It is possible to enhance accessibility by extending support to multilingual news sources.

- Real-time fact-checking can be used with external databases to improve misinformation detection.
- Users wonder how artificial intelligence makes summaries. Future versions should have explainable techniques.
- A more personalized experience can be offered by giving users the ability to adjust summary length, bias sensitivity and ethical parameters.
- There is integration with news platforms. The system can be used as a mobile app or a browser extension.

REFERENCES

1, Liu, Y., Zheng, Y., Wan, X., & Sun, X. (2022). "News Summarization and Its Ethical Challenges". arXiv preprint arXiv:2209.12356. Retrieved from https://arxiv.org/pdf/2209.12356
2, Narayan, S., Cohen, S. B., & Lapata, M. (2018). "Don't Give Me the Details, Just the Summary! Topic-Aware Convolutional Neural Networks for Extreme Summarization". Proceedings of EMNLP 2018.
3, Zhang, J., Zhao, Y., Saleh, M., & Liu, P. J. (2020). "PEGASUS: Pre-training with Extracted Gap-Sentences for Abstractive Summarization". Proceedings of ICML 2020.
4, Bae, S., & Lee, Y. (2021). "Fact-Check or Fake? Analyzing the Credibility of News Content Using NLP Models". Journal of Computational Journalism, 9(1), 25-40.
5, Brown, T., Mann, B., Ryder, N., et al. (2020). "Language Models Are Few-Shot Learners". NeurIPS 2020.
6, Wang, L., Yang, Y., & Su, Z. (2022). "Bias Mitigation in News Summarization: Challenges and Opportunities". Computational Journalism Conference 2022.
7, Raffel, C., Shazeer, N., Roberts, A., et al. (2020). "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer". Journal of Machine Learning Research.
8, C. Guan, A. Chin, and P. Vahabi, "Enhancing News Summarization with ELearnFit through Efficient In-Context Learning and Efficient Fine-Tuning," 2024.
9, T. Goyal, J. J. Li, and G. Durrett, "News Summarization and Evaluation in the Era of GPT-3," 2023.