"SentiNEWS: Multilingual Summarization and Sentiment Analysis with SerpAPI for Efficient and Accessible News Consumption"

Prolay Shankar Mazumder
prolay23065@iiitd.ac.in

Shaksham Singhal
shaksham23087@iiitd.ac.in

Shivani Sharma
shivani23142@iiitd.ac.in

Argharupa Adhikary
argharupa23020@iiitd.ac.in

Arunoday Ghorai
arunoday23023@iiitd.ac.in

Sandip Pal
sandip23080@iiitd.ac.in

1 Introduction

In an era dominated by information, staying informed is crucial, but navigating through vast amounts of news can be overwhelming. Introducing SentiNEWS, a revolutionary application designed to transform the way we consume news – making it efficient, accessible, and insightful.

2 What is the problem identified in the project?

- News is cluttered and not organised
- News pages seems too much work to read
- People usually dont have time to read through whole article
- Need for single word sentiment or tone analysis
- News is not easily accessible by people with disability(visual disability).

3 Why is this problem important?

- **Accessibility and Readability:**
  - News content should be easily accessible and readable.
- **Cultivating News Reading Habits:**
  - Raise awareness about the habit of news reading in everyday life.
- **Interest over Mundanity:**
  - Make news interesting rather than mundane and boring.
- **Efficient Information Retrieval:**
  - Allow users to input keywords for obtaining relevant news links through the SERP API.
- **Multilingual Accessibility:**
  - Overcome language barriers and foster inclusivity.
- **Content Summarization for Quick Understanding:**
  - Automate the summarization of the top 5 news articles for quick comprehension.

Preprint. Under review.

4 Related works:

- **Google News:** Personalized news summaries based on user preferences and trending topics.
- **inShorts:** News aggregation app providing concise summaries.
- **Dailyhunt:** News app and aggregator offering summaries across categories.

5 Novelty of Work:

- **W**e are using SerpApi so we get the best news pages available on the internet based on google news search engine.
- **W**e are generating important images that can quickly summarize and generate attraction of the users..
- **W**e give user news related to the user query which is limited instead of a vast sea of information on the internet

6 Techniques/Algorithms to be Used:

- Hugging face models Text-to-image and text-summarization and Sentiment Analysis
- Streamlit in Python and Google Colab for notebook
- SerpApi for searching across the web
- Langchain Interface for smooth Interfacing for Hugging face models

7 How will we evaluate our work?

Precision of SerpAPI Links, Summary Conciseness and Clarity, Polarity Score Accuracy, User Satisfaction and Engagement, Multilingual Support Effectiveness, Real-time Updates and Freshness of Content, Racall. Any one or a combination of the mentioned metrics will be used for evaluation in our project.

8 Potential Contributions:

- **Serp API Implementation:** - Arunoday Ghorai, Sandip Pal
- **UI:** - Shaksham Singhal, Prolay Shankar Mazumder
- **Summary and Polarization:** - Argharupa Adhikary, Shivani Sharma
- **Dataset, Report, Documentation, Multilingual Support:** - All
- **Logic Implementation:** - All

9 References:

1. Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., van der Goot, E., Halkia, M., Pouliquen, B., and Belyaeva, J. (2010): Sentiment analysis in the news. In: Proceedings of LREC'10, ELRA.
2. Kabadjov, M., Steinberger, J., Pouliquen, B., Steinberger, R., Poesio, M. (2009): Multilingual statistical news summarisation: Preliminary experiments with English. In: Proceedings of the Workshop on Intelligent Analysis and Processing of Web News Content at the IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WIIAT), ACM.
3. Wenxin Jiang, Nicholas Synovic, Matt Hyatt, Taylor R. Schorlemmer, Rohan Sethi, Yung-Hsiang Lu, George K. Thiruvathukal, James C. Davis, "An Empirical Study of Pre-Trained Model Reuse in the Hugging Face Deep Learning Model Registry", ACM/IEEE 45th International Conference on Software Engineering (ICSE) 2023, Cornell University, 2023.

4. Maximilian Speicher, Andreas Both, Martin Gaedke, "S.O.S.: Does Your Search Engine Results Page (SERP) Need Help?", 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, pp: 1005-1014.
5. Alexandra Balahur, Ralf Steinberger, Mijail Kabadjov, Vanni Zavarella, Erik van der Goot, Matina Halkia, Bruno Pouliquen, Jenya Belyaeva, "Sentiment Analysis in the News", Computation and Language (cs.CL), 2013.