CrisisAnalyzer: A Multimodal Approach to Real-Time Disaster Summarization

Samarth Khandelwal, Yessenia Chapeta, Bruno Sanchez, Griffin Kanzeg, Moussou Kanadjigui, Dataar Ghuman, Mentor: Nick Jarvis
ExLAI Scholars Program, University of Cincinnati

Introduction

- **Background:** Social media platforms such as X (formerly Twitter), provide real-time information from people affected by a disaster. However, the overwhelming amount of data combined with the **potential misinformation** poses a significant challenge for extraction of useful insights to support emergency response (Imran et al. 12).
- **Literature Review:** Previous studies (e.g., Ofli et al., 2020; Aronson, 2018) have demonstrated the utility of Natural Language Processing (NLP) for classifying disaster-related tweets and computer vision for extracting information from images to improve situational awareness.

The above approaches has been effective in processing the data, but they come up short when it comes to providing actionable information and misinformation handling

- **Proposed Solution:** The CrisisAnalyzer system addresses these issues by integrating multimodal inputs (tweet text and images) with advanced NLP techniques to generate concise, **semantically accurate** summaries while identifying misinformation through contextual knowledge.

This ensures that emergency responders receive reliable, useful updates.

- **Objective:** The primary task of the proposed model is to process and summarize real-time social media data to provide accurate information to assist in decision-making and coordination during crises.

Methodology

- **Dataset:** The **CrisisMMD** dataset includes six subsets containing disaster-related tweets, corresponding metadata, and associated images.
- **Processing:** The data is grouped into sets of 100 tweets, each containing a URL to an associated image. These groups are processed by the GPT-4o mini model via the OpenAI API, where both textual and visual information are analyzed.
- **Model:** The GPT-4o mini model processes each set and generates concise summaries, ensuring that essential situational details and incident-specific context are preserved.
- **Evaluation:** Generated summaries are evaluated against three reference **news articles** reporting on the same disaster event.

Validation Metrics:

- **ROUGE-1**: Measures the overlap of individual words between the generated summary and the reference text, providing a basic evaluation of lexical similarity.
- **ROUGE-L**: Assesses the longest common subsequence between the summary and the reference text, capturing structural similarity and coherence.
- **BERTScore**: Evaluates the semantic similarity by comparing contextual word embeddings, ensuring the summary retains the underlying meaning of the original content.

Results

The model was validated using six subsets from the CrisisMMD dataset.

- **Median ROUGE-1: 0.22**, indicating limited lexical overlap
- **Median ROUGE-L: 0.10**, reflecting differences in structure and phrasing
- **Median BERTScore: 0.78**, shows the model's preservation of semantic content

Inferences:

- Preliminary results indicate the model generates summaries that accurately capture meaning, although phrasing and structure are significantly altered.
- Despite these limitations, the system remains effective for real-time disaster monitoring, prioritizing the delivery of meaningful content over exact phrasing.

Figure 1:

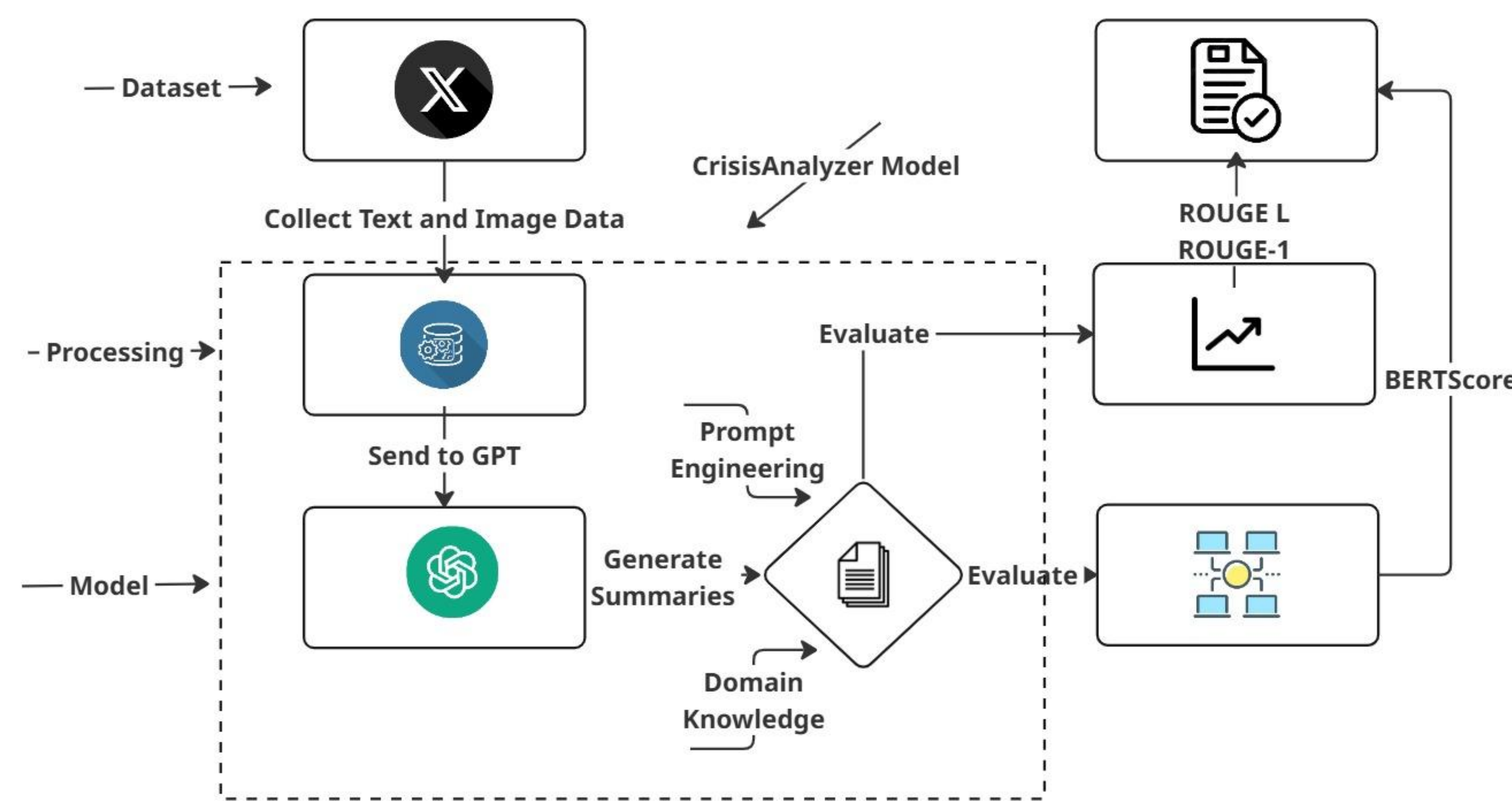


Figure 2:

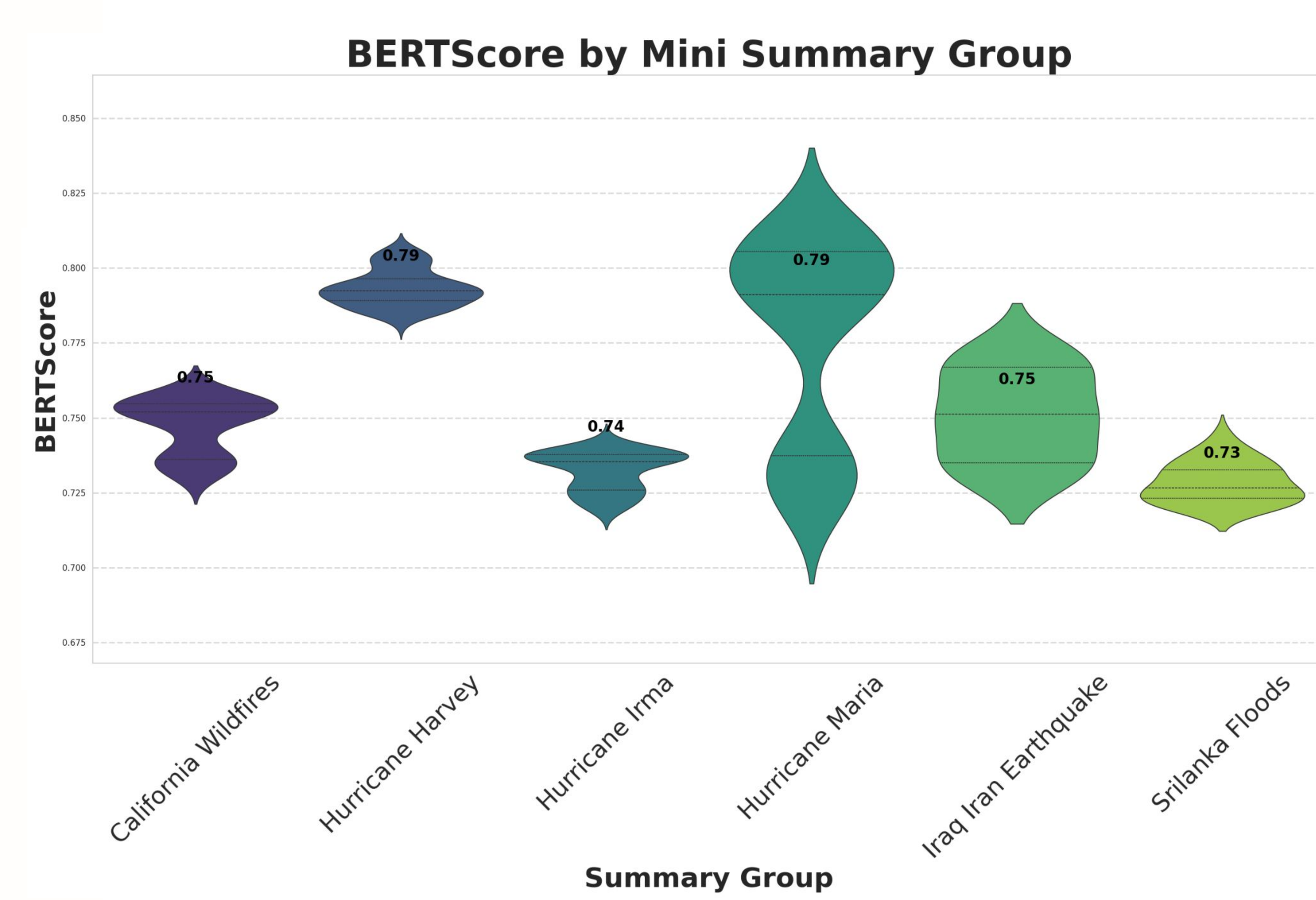
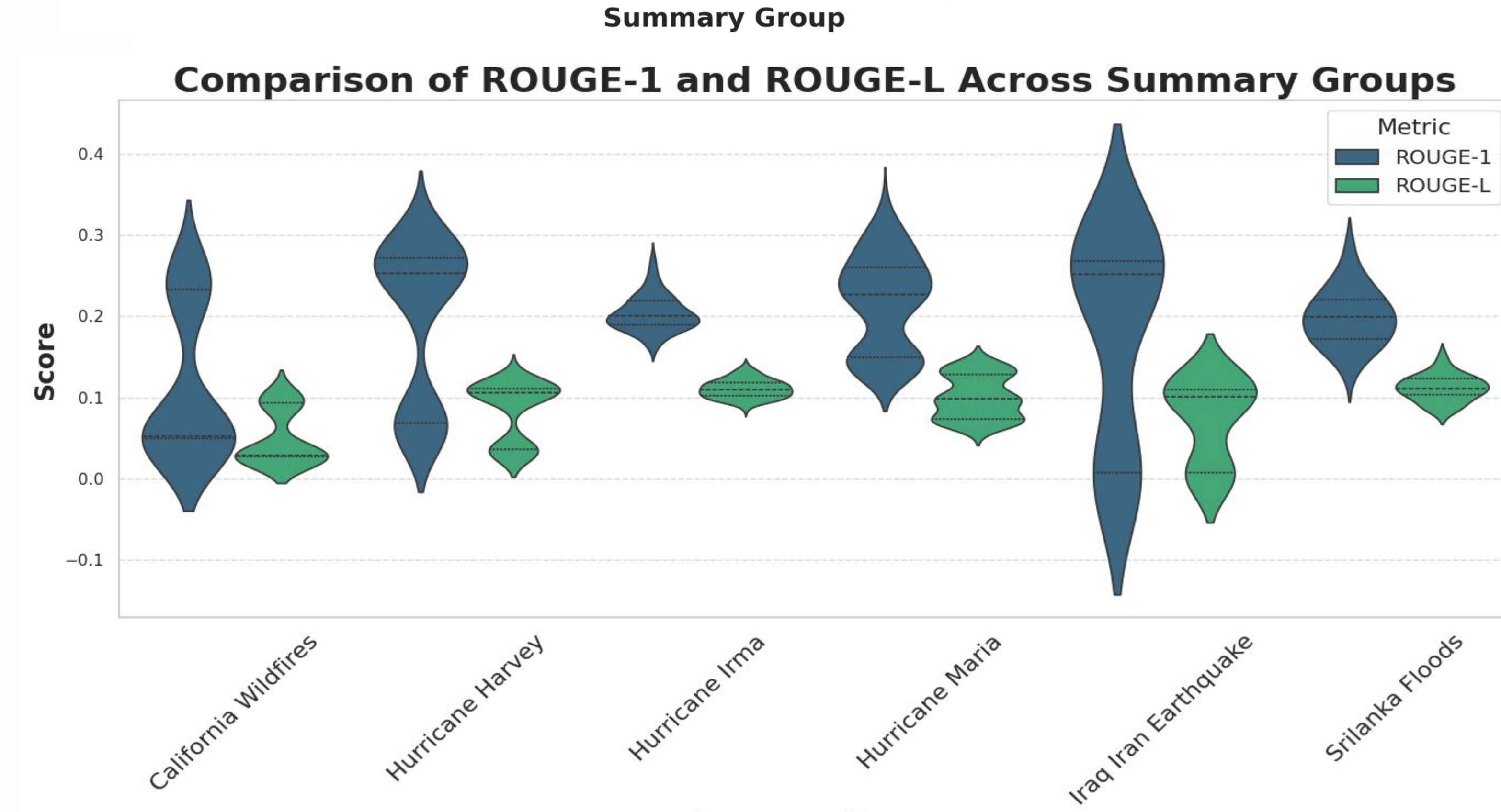


Figure 3:



Conclusions

- The CrisisAnalyzer system integrates **multimodal** data from social media to produce concise, accurate summaries for **disaster response**.
- Evaluation using ROUGE-1, ROUGE-L, and **BERTScore** validated the models effectiveness, with a BERTScore of **0.78** indicating strong semantic similarity.
- The model supports emergency responders by **filtering noise** and misinformation, providing actionable insights in real time.
- CrisisAnalyzer offers a robust foundation for real-time disaster monitoring and presents opportunities for further development.

Future Recommendations

- **Model Performance**: Using advanced LLMs like OpenAI's o3 and GPT-4.5 for improved summarization.
- **Fine-Tuning**: Training on **domain-specific data** (e.g., dispatcher reports) to enhance quality and relevance.
- **Prompt Engineering**: Employing multi-step, **context-aware** prompts for clearer, higher-quality summaries.

Acknowledgements

We would like to express our gratitude to the ExLAI program and the University of Cincinnati for providing the resources and support that made this project possible. We extend our heartfelt thanks to our mentor, Nick Jarvis, for his guidance, and to Dr. Zhan, the program director, whose leadership was instrumental in the success of this endeavor.

Citations

Alam, Firoj, et al. "CrisisMMD: Multimodal Twitter Datasets from Natural Disasters." *In Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*, 2018, dataverse.mpi-sws.org/dataverse/icwsm18.

Aronson, Jay D. "Computer Vision and Machine Learning for Human Rights Video Analysis: Case Studies, Possibilities, Concerns, and Limitations." *Law & Social Inquiry*, vol. 1188–1209, no. 4, Cambridge UP, season-03 2018, www.jstor.org/stable/26630958.

Imran, Muhammad, et al. "Processing Social Media Messages in Mass Emergency: A Survey." *arXiv (Cornell University)*, Jan. 2014, https://doi.org/10.48550/arxiv.1407.7071.

Ofli, Ferda, et al. "Analysis of Social Media Data Using Multimodal Deep Learning for Disaster Response." *arXiv (Cornell University)*, Jan. 2020, https://doi.org/10.48550/arxiv.2004.11838.