GEO-SEMANTICS ANALYSIS OF ENVIRONMENTAL DISASTERS IN NIGERIA USING NATIONAL PRINT MEDIA DATA FOR DISASTER MANAGEMENT

Benedict Ajanaku, Rashidat Sikiru, Anthony Soronnadi, Ife Adebara, Olubayo Adekanmbi

INTRODUCTION

Disasters often prompt immediate information sharing, with location being a critical aspect. Timely identification of affected areas is essential for assessing impact and coordinating response efforts. However, in Nigeria, disaster management remains a challenge due to limited resources and response capacity. While the country experiences fewer severe natural disasters compared to other regions, improving preparedness and response mechanisms is crucial for mitigating risks. Artificial Intelligence (AI) has demonstrated significant potential in disaster management, particularly through natural language processing (NLP), machine learning, and predictive analytics. These technologies enhance real-time decision-making, risk assessment, and resource allocation. This paper focuses on the use of Geo-NLP (using NLP to extract geographical information), with a focus on using information from news articles to help disaster responses and humanitarian efforts. The aim is to demonstrate the usefulness of NLP to contribute to the ongoing discussion to improve emergency response capabilities.

METHODOLOGY

The proposed framework comprises five (5) stages as shown below: Data Collection, Data Preprocessing, Custom NER Model Training, Geo-parsing, and Geo-Visualization. This methodology was adopted from Idakwo et al. (2023) paper.

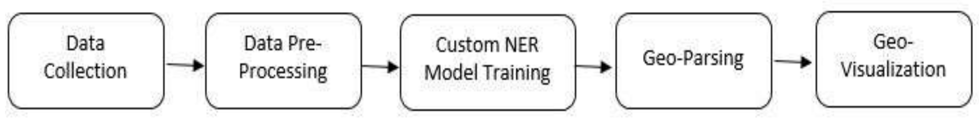

Figure 1: Methodology flow

A total of 3,722 disaster-related articles (fire outbreaks, floods, oil spills) were collected from Tribune Nigeria using a systematic query method. Data was scraped with Beautiful Soup and organized into three columns: headline, article link, and article text, then saved as a CSV file. To retain only relevant content, a RoBERTa-based text classification model—chosen for its superior performance over BERT—was used to filter the corpus. Named Entity Recognition (NER) annotation was conducted with the SpaCy NER Annotator as shown below, focusing on geographical entities critical to disaster analysis: Village, Market, Town/City, Centre, State, LGA, River, Road, Community, and Landmark. This manual process yielded 1,286 annotated records, which were stratified into training (80%), development (10%), and test (10%) sets. The annotated dataset supports the development of a domain-specific NER model for disaster monitoring in Nigeria.

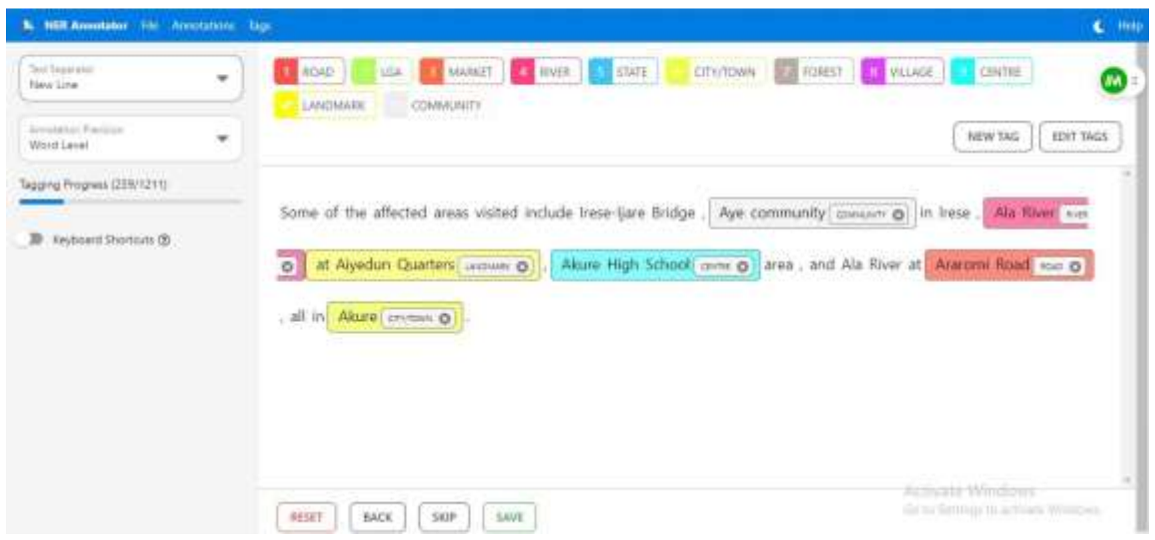

Figure 2: NER Spacy annotator tool

RESULT

Models	Precision	Recall	F1-score
spaCy	0.95	0.77	0.85
BERT	0.99331	0.99349	0.99297
DistilBERT	0.99236	0.99297	0.99240

Table 1: Result of model comparison

The fine-tuned models exhibit variations in performance, highlighting the impact of domain-specific adaptation. BERT achieved the highest performance with a precision of 0.99331, recall of 0.99349, and f1-score of 0.99297, closely followed by DistilBERT with a precision of 0.99236, recall of 0.99297, and f1-score of 0.99240. In contrast, SpaCy exhibited lower performance with a precision of 0.95, recall of 0.77, and f1-score of 0.85. These results confirm the effectiveness of transformer-based models for NER tasks. The superior precision and recall of BERT and DistilBERT highlight their robustness in identi- fying disaster-related entities. Their f1-scores indicate suitability for real-world applications, ensuring accurate extraction of critical disaster information for emergency response teams.

Locations	Entity	Latitude	Longitude
Awe LGA	LGA	8.111812	9.146371
Nasarawa	STATE	8.438787	8.238285
Lokoja	CITY/TOWN	7.802355	6.743033

Table 2: Sample geocode location from corpus

In this phase, toponyms were extracted, and their corresponding latitude and longitude coordinates are presented above. These locations were then visualized on an interactive map, providing a clear representation of the spatial distribution of the identified toponyms.

Figure 3: Map showing disaster distribution

CONCLUSION

The ability to accurately extract and locate disaster-affected areas is vital for enhancing situational awareness and improving the efficiency of emergency response operations. By precisely pinpointing affected regions, our method facilitates the more effective allocation of resources, enabling timely assistance to impacted populations. The findings highlight the critical role of advanced NER models in disaster management, illustrating their capacity to transform textual data into actionable geographic insights. This technological advancement not only aids immediate response efforts but also contributes to long-term disaster planning and preparedness strategies. In addition, we recognize that the current framework primarily utilizes news articles, which, while valuable, may not capture the full immediacy of rapidly unfolding events. Datasets sourced from platforms like Twitter or Bluesky, which offer robust and timely data streams, could further enhance the framework's effectiveness by reducing disaster response times.

REFERENCES

- Patricia Ojonoka Idakwo, Olubayo Adekanmbi, Anthony Soronnadi, and Amos David. Geo-parsing and geo-visualization of road traffic crash incident locations from print media for emergency response and planning. In 5th Workshop on African Natural Language Processing, 2023.
- Ni Putu Ayu Sherly Anggita Sugiartaa and Ngurah Agus Sanjaya ERa. Location named-entity recognition using rule-based approach for balinese texts. Jurnal Elektronik Ilmu Komputer Udayana p-ISSN, 2301:5373, 2021.
- Lishuang Li, Degen Huang, and Chunrong Chen. Identification of location names from chinese texts based on support vector machine. JOURNAL-DALIAN UNIVERSITY OF TECHNOLOGY, 47(3):433, 2007.
- Anthony Ivan. Spatial visualizations and analysis in python with folium. Towards Data Science, 2018.