**Project 1: Fact-checking Disaster Information**

The prevalence of false information, or fake news, is a significant obstacle to effective communication and public safety, particularly during disasters. The spread of misinformation can cause confusion and fear and even prompt dangerous actions among the public. For example, fake news can undermine official warnings and emergency responses issued by agencies such as FEMA, endangering lives. Disseminating false information during critical times disrupts the coordinated efforts needed for community safety and hinders the efficient allocation of emergency services and resources.

To address this issue, our project will develop a comprehensive fact-checking methodology that aims to accurately identify and classify misinformation within the context of disaster scenarios. This methodology will involve two crucial steps: evidence retrieval and claim classification. Initially, relevant evidence will be gathered from credible sources to ensure the accuracy of any given claim. Subsequently, the claim will then be classified as either supported, refuted, or inconclusive based on the evidence gathered.

Furthermore, we plan to enhance disaster tweet datasets with detailed annotations that include geographical location, time and date, severity level, and other pertinent information to further refine our approach. This multi-dimensional analysis, which will incorporate sentiment analysis, source verification, and proposed response actions, will significantly improve our ability to assess the reliability and context of information, thereby minimizing the impact of fake news during disaster events.

**Project 2: Multilingual Event Detection**

Recent contributions have highlighted the challenges inherent in multilingual event detection, a complex area of natural language processing (NLP).
- One of the critical challenges is the difficulty of detecting multiple languages, which significantly complicates language detection and analysis.
- Furthermore, NLP tools that are not universally compatible can be an additional hindrance to comprehensive analysis across various languages.
- In addition, semantic ambiguity, a reflection of the complex semantic structures present within languages, adds an extra layer of complexity to textual interpretation.
- Translation and transliteration issues pose a risk of losing vital information, which is exacerbated by dialect variations that introduce unfamiliar vocabulary and grammar into the analysis.
- Misinterpretation and translation errors due to diacritical marks pose yet another obstacle to clear communication.
- Additionally, cultural nuances underscore the importance of understanding cultural idioms and conventions for accurate translation and interpretation of text.

To address these challenges, our proposed solution will utilize a systematic approach that begins with language detection and progresses to translation, preprocessing, and crowdsourcing, ultimately leading to advanced NLP models for final classification. This methodological pipeline will include translating text into English, grammatical correction, and refinement of translations. It then moves on to the automation of information extraction, ensuring that classification is not only precise but also culturally and contextually aware. The process will also incorporate strategies to address dialectal intricacies, diacritical marks, and cultural nuances, ultimately aiming to

prevent the loss of critical information and minimize the risk of misinterpretation. This comprehensive approach to NLP will navigate linguistic diversity and enrich the analysis and application of language technology in global communication and event detection.