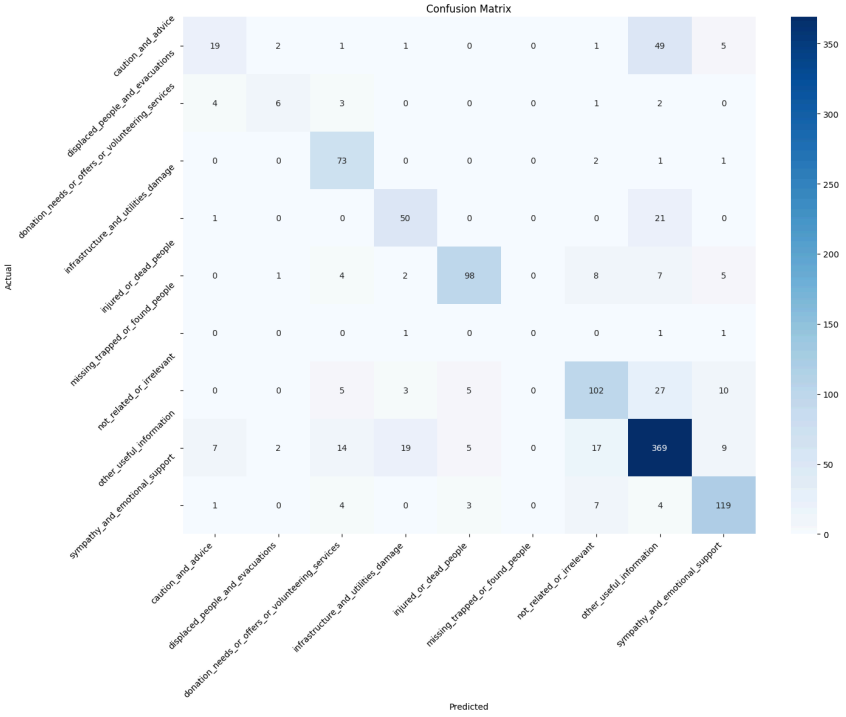
MAIS 202
Deliverable 3
Team members: Ruolin Hu, Sihui Wei, Shengyi Zhong

Social Media-Enabled Crisis Informatics

1. Final Training Results

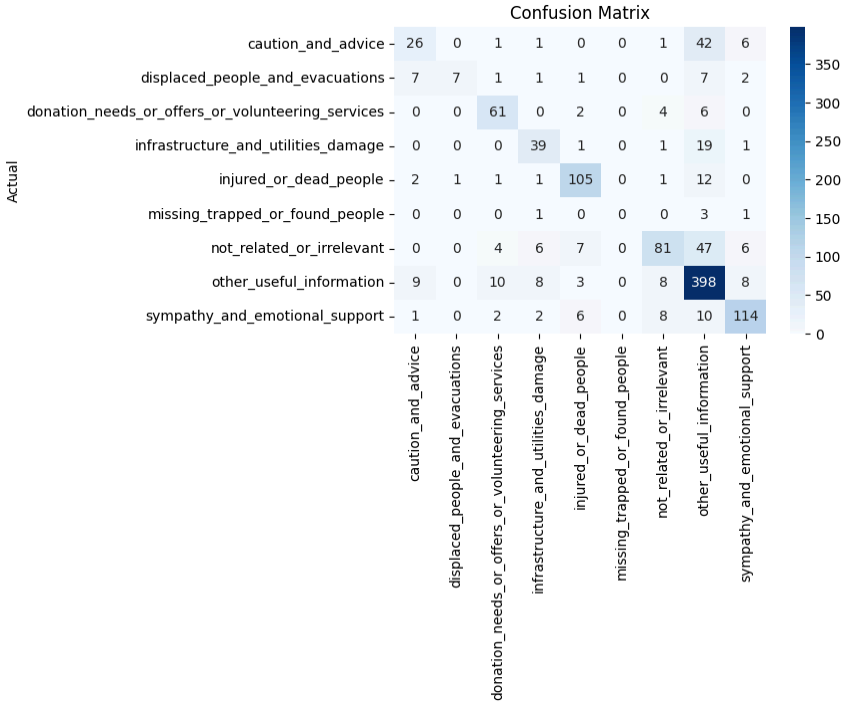
We used two approaches to improve the performance of our model:

1. Pre-given word embeddings [1] on CNN
2. DistilBERT



1.1 Result of Pre-given word embeddings on CNN

1.2 Result of DistilBERT



1.3 Preliminary result of Random Forest Classifier with simple text preprocessing

DistilBERT shows good performance on classes with more distinct patterns, such as "other useful information" and "injured or dead people." This suggests that the model captures semantic context effectively. However, misclassifications are observed, particularly between "caution and advice" and "sympathy and emotional support," likely due to overlapping vocabulary. The model struggles with rare classes, as seen with "missing, trapped, or found people."

The CNN model appears to perform well in some categories, but the performance is generally lower than DistilBERT in terms of capturing textual patterns. Misclassifications are more common, especially with classes such as "not related or irrelevant" and "displaced people and evacuations." This indicates that CNN's performance might be limited due to its lack of contextual understanding compared to transformer-based models like DistilBERT.

The Random Forest model provides reasonable baseline performance on categories like "other useful information" and "sympathy and emotional support." It shows that even with simple text preprocessing, the model can differentiate broad classes. The model's performance is generally less robust, with significant misclassifications across categories. This indicates that basic vectorized representations do not capture the semantic depth required for distinguishing closely related categories.

Overall, DistilBERT outperforms CNN & Word Embeddings and Random Forest Classifier in terms of both overall accuracy and category-specific performance.

2. Final demonstration proposal

In the final demonstration, the user can either upload a CSV file containing tweets or enter text directly into a search bar on the website. Once the user submits the input, the backend processes this data by first applying text preprocessing.

The backend uses the machine learning model with the best performance (CNN, DistilBERT, or Random Forest) to classify the processed tweets into predefined categories, such as "infrastructure damage", "missing persons", or "donation needs".

Once the classification is complete, the backend sends the results back to the frontend, where they are displayed to the user. The frontend interface will present a categorized list of the submitted tweets, showing the predicted category for each tweet.

Additionally, the application will generate visual representations, such as pie charts, to illustrate the distribution of classifications.

References

[1] Dat Tien Nguyen, Kamela Ali Al-Mannai, Shafiq Joty, Hassan Sajjad, Muhammad Imran, Prasenjit Mitra. Robust Classification of Crisis-Related Data on Social Networks using Convolutional Neural Networks. In Proceedings of the 11th International AAAI Conference on Web and Social Media (ICWSM), 2017, Montreal, Canada.