M7 Project  
Kaggle: Natural Language Processing with Disaster Tweets

Johnathan Wang, Nicholas Charles Vitellaro, Makenna Shae Owens, and Marc Mahanna   
COSC523: Artificial Intelligence, November 15, 2023

***Abstract* - This project focuses on using machine learning and natural language processing approaches to implement models predicting which Tweet validity in disaster situations. The format of this project is that of a Kaggle competition** [1]**. Goals of this project are to 1) implement a baseline score and 2) try to improve on the baseline score.**

**DESIGN CHOICES AND IMPLEMENTATION**

The approach selected was to implement a baseline and an improved model intended to improve score beyond the baseline. … is the default implementation and is provided in the Kaggle competition tutorial [2]. … is improved model providing the best performance after much team experimentation. For each experiment and subsequent baseline and improved implementation, data was loaded, data was cleaned, features were explored and selected, models were trained, and test results were produced.

The baseline model followed the Kaggle tutorial. Data cleaning and preparation was minimal as the selected features …. Resulting performance when submitted to Kaggle was …%.

The improved model was the result of much team experimentation. The improved model uses features that …. Resulting performance when submitted to Kaggle was …%.

**CHALLENGES AND OBSTACLES**

The project group members are familiar with machine learning concepts and have prior experience with …. Use of standard practices of data exploration, cleaning, and normalization were employed making exploration of model performance achievable.

…

**DISCUSSION AND FUTURE WORK**

The team believes that …

Additionally, the team …

**SUMMARY**

This project provided the opportunity to compare the performance of different machine learning and natural language processing models. A consumable data set and a challenge issued in the form of a competition provided an enriching and rewarding experience for the project group.

**REFERENCES**

[1] "Natural Language Processing With Disaster Tweets," [Online]. Available: https://www.kaggle.com/competitions/nlp-getting-started.  
[2] "KerasNLP starter notebook Disaster Tweets," [Online]. Available: https://www.kaggle.com/code/alexia/kerasnlp-starter-notebook-disaster-tweets.