• Summary of the report.

This project works on Natural Language Understanding, as its task is to predict whether a tweet indicates a real disaster, or just metaphorism. The dataset consists of text of tweets, keywords indicating possible disaster, and location where the tweet is sent. Some common text processing methods are applied on tweets' text, along with the observation on the usefulness of location and keywords. After that, Group 1 adopt two attention-based models RoBERTa and XLNet for the task, yielding result ~0.8 Kaggle score. Further improvement may lay on how to utilize keyword and location of the tweets, or using a larger model for the task.

• Describe the strengths of the report.

It clearly shows us the advantages of XLNet in comparing to BERT, and point out some directions for further improvement

• Describe the weaknesses of the report.

Maybe more details are better, as everything in this report is just on the surface. I know that even our group report, we do not detailly describe the dataset, but I think it's better if you include the structure of dataset to the poster, rather than let us to go the Kaggle site to see it. And, it seems that Prof. did not show the architecture of RoBERTa in class, so you should point out the difference between it and BERT.

• Evaluation on quality of writing (1-5): Is the report clearly written? Is there a good use of examples and figures? Is it well organized? Are there problems with style and grammar? Are there issues with typos, formatting, references, etc.? Please make suggestions to improve the clarity of the paper, and provide details of typos. (4)

• Evaluation on presentation (1-5): Is the presentation clear and well organized? Are the language flow fluent and persuasive? Are the slides clear and well elaborated? Please make suggestions to improve the presentation. (3)

The presentation surprisingly has much more content than the poster, and in each slide of Section 3, there is MANY text, so it's hard to follow. It affects the organization of slides, as they are so dense. Also, the overall content seems to be discussing a model, instead of experimenting with it.

• Evaluation on creativity (1-5): Does the work propose any genuinely new ideas? Is this a work that you are eager to read and cite? Does it contain some state-of-the-art results? As a reviewer you should try to assess whether the ideas are truly new and creative. Novel combinations, adaptations or extensions of existing ideas are also valuable. (4)

• Confidence on your assessment (1-3) (3- I have carefully read the paper and checked the results, 2- I just browse the paper without checking the details, 1- My assessment can be wrong) (2)