**Assignment 3**

Rohan Charudatt Salvi (rcsalvi2)

**Method**

1. I took the problem as a multilabel classification problem
2. Based on the word in the training dataset, I extracted the sentence it is in.
3. Then used the BERT tokenizer to tokenize the sentence
4. This was followed by generating the sentences embeddings as well as mention masks
5. Finally, I used BERT with the MultiLabelSoftMarginLoss function to train it on the training dataset

For prediction, the initial steps were the same (1-4 steps above).

1. Then on the model, I predicted and extracted only the highest probability label.
2. If the label had multiple labels, I transformed it so it has all the labels for the given label.
3. Then generated a new tab file for the test set

**Results**

Results can be obtained by evaluating the ‘rcsalvi2\_predictions.tab’ file through the scorer

**Findings**

I found that the model reasonably learns the basic NER task i.e for Level 1 over the epochs. And does get it quite right for Level 1. When we come to Level 2 the performance falls a lot and it mixes up between classes e.g I noticed it predicts FAC.structure while the label is FAC.building. Lastly, coming down to Level 3 the model isn’t performing too well and it tends to pick the same label for all truth labels belonging to Level 3 meaning that it predicts ‘FAC.Structure.Bridge’ mostly for all the labels which have level 1 and 2 as ‘FAC.Structure’.

I believe the reason could be less training data, especially considering the number of output classes we have thus model needs more data or information. Additionally, training it for several epochs might also improve the score by a bit since I trained it for 5 epochs only.