**Hands-on Assignment 5**

**Due Date: See web**

In Tutorial 5, you have seen the use of a BERT model for a spam classification task through fine-tuning. In this hands-on assignment, you are asked to fine-tune the BERT model for another classification task --- sentiment analysis with the IMDB Dataset (which can be found in the same folder). Note that labels in IMDB Dataset.csv are “positive” and “negative”, while the labels in the spam dataset in Tutorial 5 are “0” and “1”.

You need to randomly shuffle the dataset and split the dataset into a training set and a testing set. The whole dataset consists of 50k samples. You should randomly select 5k samples for the training set and 500 samples for the testing set.

In your experiments, you should

* Vary the learning rate (at least three options)
* Vary the batch size (at least two options)
* Vary the dropout rate (at least two options)
* Vary the maximum number of epochs (at least three options)

In total, you will need to run at least 10 experiments. Write a simple report to summarize the results, including the test result of each run. A simple report should not only discuss the changes made in the model hyper-parameters and the corresponding result, but also analyze the results. A report should be organized and must be in docx or pdf format.

Please submit the report on Canvas. There is no need to submit your code. Similarity scores will be computed for this assignment.

*Copyright Statement: The materials provided by the instructor in this course are for the use of the students enrolled in the course. Copyrighted course materials may not be further disseminated.*