https://github.com/cwinsor/uml_comp5300/blob/main/hw7_long_transformers/hw_bert_classification_final.ipynb