* Train bert model on glue datasets (https://huggingface.co/datasets/glue):
  + glue mrpc data
  + Glue stsb data