

1. Using the corpus used in Assignment 1, implement Byte Pair Encoding and WordPiece algorithms in the following settings.
 - a. Set the total number of merge steps to 32000.
 - b. Set the vocabulary size to 32000.

Implement all these from scratch without using any predefined Python packages.