Question 3: Find the THREE most frequent bigrams in the following dataset. The symbols <s> and </s> represent the start and end of the sentence respectively. Consider them as valid tokens. [10 Marks]

Dataset

<s> I am Sam </s>

<s> Sam I am </s>

<s> I and Sam </s>

<s> I do not like green eggs and beef </s>

Time: 20 Min