**WEEK 1**

1.1

Sub token vectorization is better because:

* They break down words into smaller subwords, giving us a wide range of vocabulary as there are more combinations that you can make
* Helps reduce size of vocabulary
* Can handle out of vocabulary words more easily because we can combine commonly used words to make unseen/rare words