**Task 4** (*16 points available, 4 for each question*)

Answer the following questions.

**4a.** Describe the performance of each featurization, comparing

* the performance of a same model between the training and test set
* the performance of different models on the training set
* the performance of different models on the test set

(*4 points available, max 150 words*)

*type your answer here*

1. The performance of the training set (real Dutch words) was better than that of the test set (pseudowords) for all models. So, the models learned patterns in real words but couldn't apply the learned to new words. The fasttext model didn't decrease the accuracy when it was used on new words. Therefore, it works similarly to pseudowords and real words. Nonetheless, the RobBERT model managed real words well but pseudowords not that well. Unigram and bigram models are better on real words as well from which bigram was the better model.
2. RobBERT performed the best, second place is fasttext. The bigram performed better than the unigram model but both had less accuracy than RobBERT and fasttext.
3. Fasttext performed best. It dealt better with pseudowords because of character n-grams usage. RobBERT performed worse. Unigram and bigram models performed the worst.