Part 3 tagger2

If a word in the training data is not found exactly in the pre trained:

1. If the lower case version is found in the pretrained we will copy the embedding and use it only for this word in original form.
2. If the lower cases version is not found in pretrained in no capitalization, we will create vectors with glorot

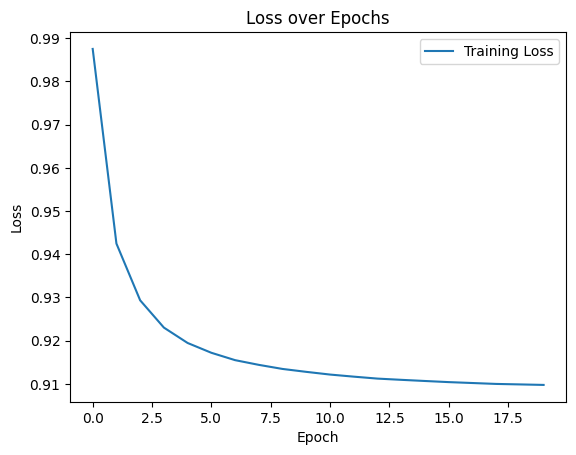
We are finetuning the word vectors

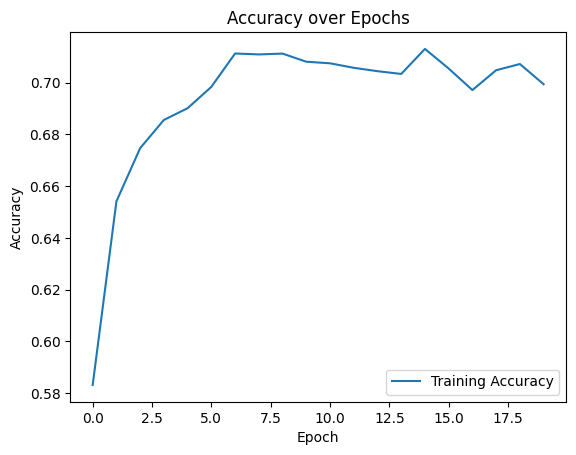
There is a risk with using in test vectors that are only pre trained and not fine tuned. Need to think this through.

NER

Hidden dim 20

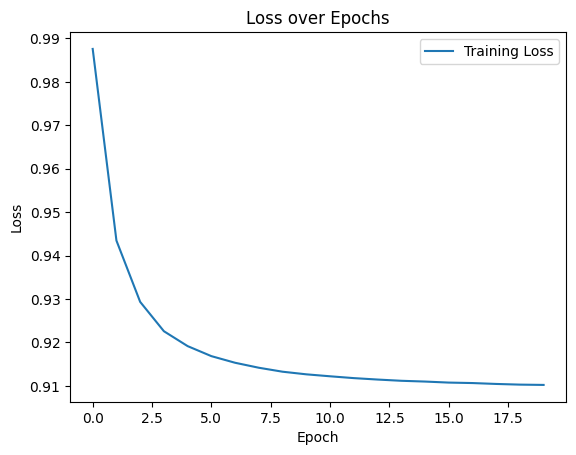
Accuracy peaked in epoch 6 with 0.7113201055166877

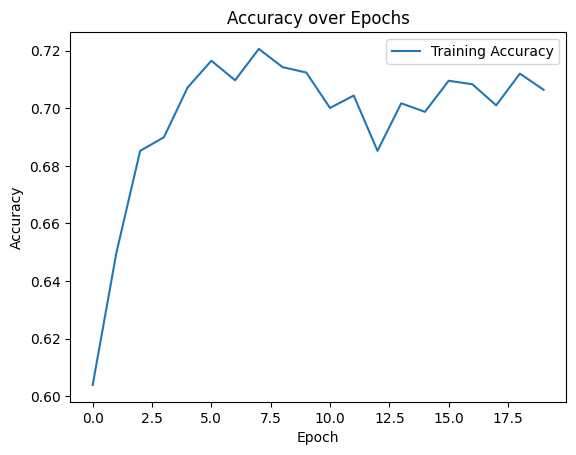




Hidden dim 40

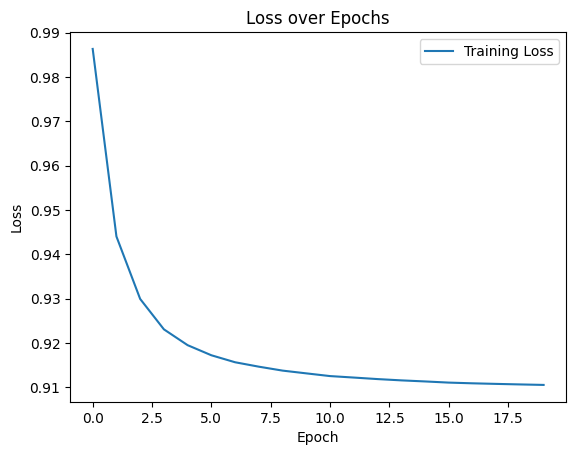
Peak in epoch 7 Accuracy: 0.7206000229016375

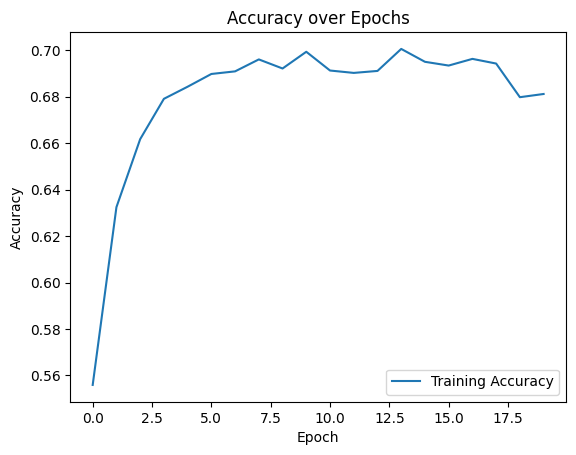




Hidden dim 60

Peak in epoch 9 with Accuracy: 0.6993471538197228





Part 4

Tagger 3

For words that are Smaller then 3, this means they don’t have prefix and suffix of 3.

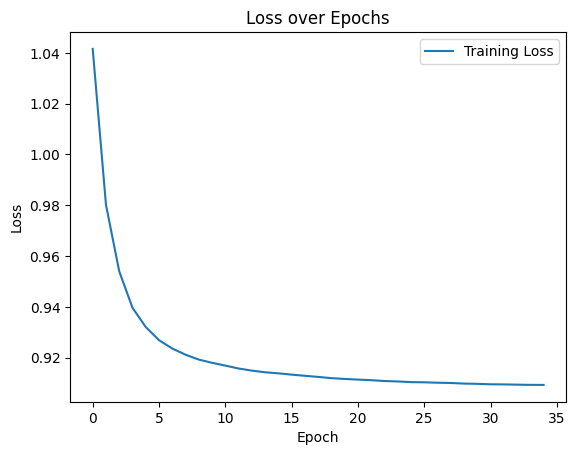
So for each word of under length 2 we will create a suffix and prefix (identical)

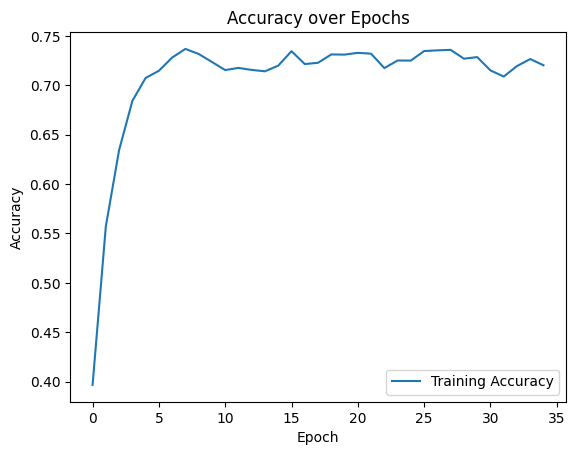
Also for the padding words there is need for some vector

Tagger 3 without pre trained vectors

Hidden dim 20

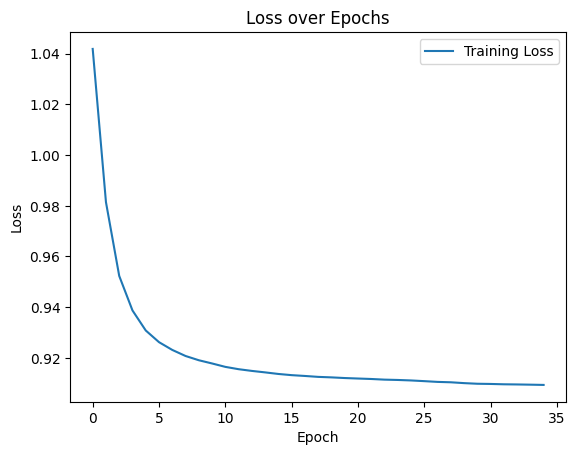
Peak in epoch 27 Accuracy: 0.7358390119250426

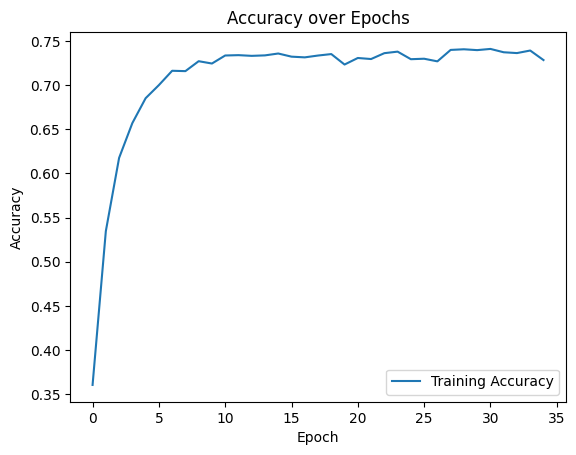




Hidden dim 40

Peak in epoch 30 Accuracy: 0.7409065472859541





Hidden dim 60

Peak in epoch 30 Accuracy: 0.7281620439636322

