Part 4

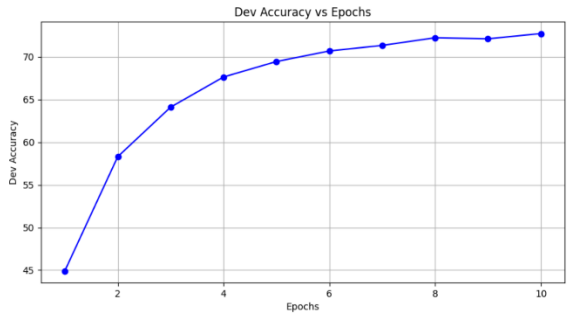
For part 4 we built a network using an embedding of the prefix and suffix of 3 words, for each of the 5 words that represents the input for the model.

There are some words which has only 1-2 characters so in order to relate to these words also, we defined an embedding not for 3 characters prefix/suffix but for prefix/suffix for 1-3 characters long.

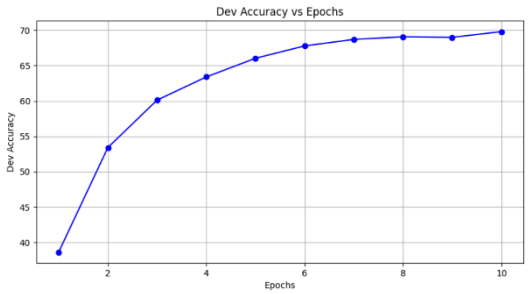
Every suffix/prefix then embedded to a vector of size 50. We observed a difference in performance between the presence/unpresence of the pre-trained embedding.

These are the DEV accuracies for NER, with and without the pre-trained embeddings,

with embedding – 72.74% accuracy

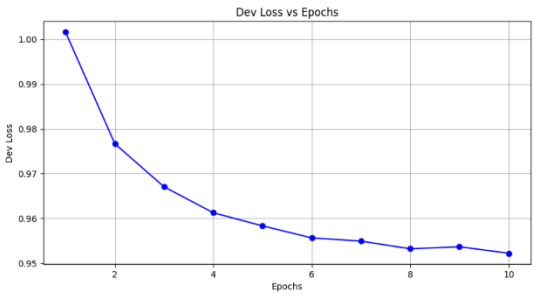


without embedding – 69.79% accuracy

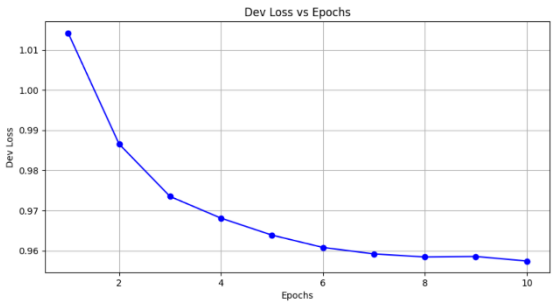


And these are the losses on the DEV set:

with embedding – 0.9522 loss

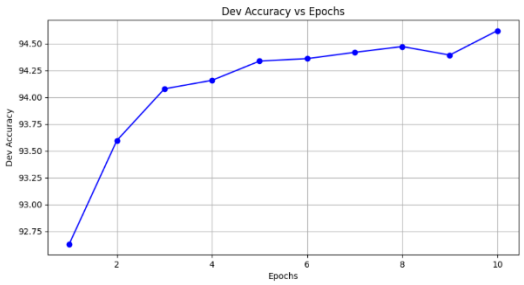


without embedding – 0.9574 loss

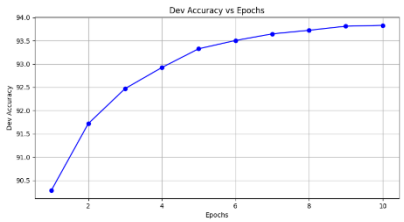


These are the DEV accuracies for POS, with and without the pre-trained embeddings,

with embedding – 94.62% accuracy

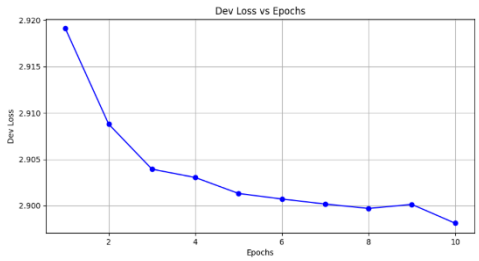


without embedding – 93.82 % accuracy



And these are the losses on the DEV set:

with embedding – 2.8981 loss



without embedding – 2.906 loss