Ass 2 – Part 3

Dorin Keshales – 313298424

Model's Parameters

In both tasks:

- I used dropout, after the activation function, with probability of p=0.5.
- Optimizer: Adam
- Learning rate: 1e-3
- Size of the hidden layer: 150

POS:

- I used batches of size 32 for both training and dev sets.
- Number of epochs: 2

NER:

- I used batches of size 16 for both training and dev sets.
- Number of epochs: 3

Description of the logic of using the pre-trained vectors

- Because of the embedding vocabulary being lower-case, when I searched in it for a match to words, which belong to the training set / dev set / test set , I checked on the word in its lower-case form too. The reason for it is that the training, dev and test sets contains also words written in capital letters.

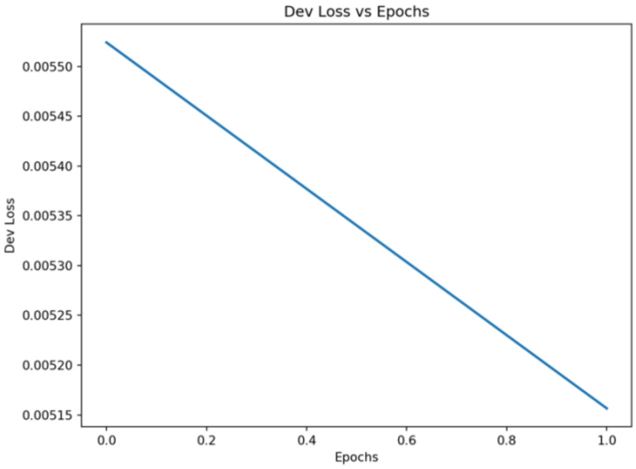
- In order to deal with words that appear in the training set / dev set / test set and not in the pre-trained vocabulary, I assign to them the unknown token - 'UUUNKKK' - which belongs to the pre-trained vocabulary and following logic represents words that are not in the original vocabulary. Therefore, when creating context window to each word in the training \ dev \ test set, the words which does not appear in pre-trained vocabulary get the correspondent embedding vector to the unknown token - 'UUUNKKK'.

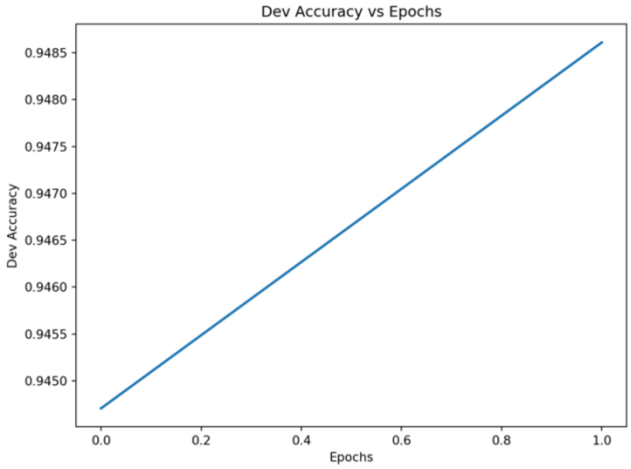
- For words that were recognized as numbers by appropriate regex, I assigned tokens from the pre-trained vocabulary which represent numbers (following logic of course), such as ‘DGDG’, ‘DG.DGDG and ‘NNNUMMM’ (there are 53 of them in the pre-trained vocabulary, when each represents a different pattern of a number).

Compared to the previous model results, accuracy did not improve, but even decreased. I can guess it might happen due to a relatively small overlap between the pre-trained vocabulary to the training and dev sets.

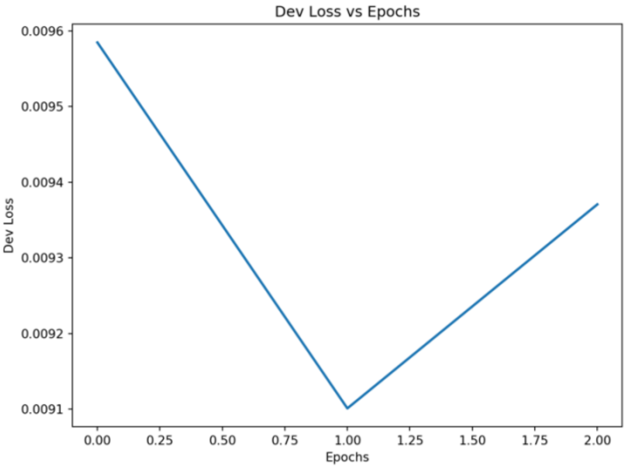
My model’s results:

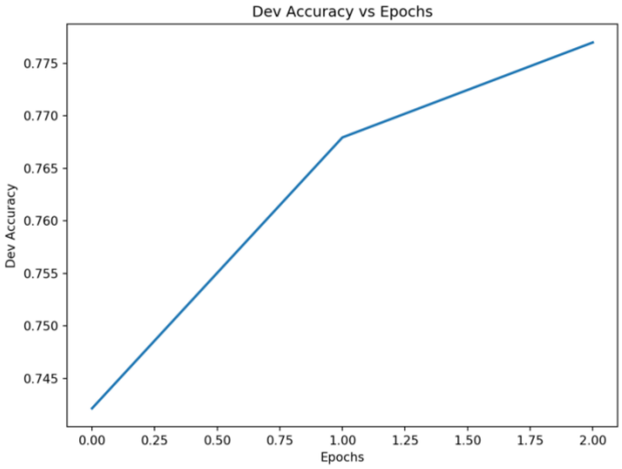
POS: Loss - 0.00515 Accuracy – **94.861%**





NER: Loss - 0.00937 Accuracy – **77.697%**


Dev Loss vs Epochs
Dev Loss
Epochs


Dev Accuracy vs Epochs
Dev Accuracy
Epochs