Part1:

Yuval Ramot - 208115840 Guy Ben Razon - 209207364

Parameters:

    In both tasks the model looked the same:
- Embedding layer (randomly initialized)
- Fully connected (linear) layer
- Activation layer, tanh (as requested)
- Dropout (p=0.2)
- Fully connected layer
- Softmax Layer

    We used Adam optimizer with lr=0.001 (default)
    Hidden layer size 150

We ran 10 Epochs for each task,
Each epoch ran on 32 batch size

Considerations:

    For words not in the training set but are in the dev set we used a special token, <unknown>. When testing results on validation (dev) whenever we encountered a word that was not seen in training we transferred it to the unknown token. We tried to randomly assign words to the unknown token when training believing it would help the model learn how to treat it but gained worse results so We removed it.

    For the last and first words of the data we added 2 special tokens, <start>, <end>, we added 2 at the beginning of the data and 2 at the end respectively.

Model's best Results:
Pos: acc: 90.8% loss: 2.937
Ner: acc: 63.2% loss: 0.968

Dev Accuracies pos vs Epochs

Dev Losses pos vs Epochs

Dev Accuracies ner vs Epochs

Dev Losses ner vs Epochs