Aliraza Punjani (amp2280)

NLP HW4

Scores:

Part1:

Unlabeled attachment score **83.62**

Labeled attachment score **80.01**

Part2:

Unlabeled attachment score **83.83**

Labeled attachment score **80.4**

Part3:

Unlabeled attachment score **85.2**

Labeled attachment score **82.2**

Explanations:

Part2:

Increasing the number of neurons allows the network to learn more complex functions capable of better fitting the data. However, it could also result in overfitting wherein the model does an excellent job of fitting the training data but doesn't generalize well. In part 2 we observe that the model with 400 neurons in the hidden layers sees a decrease in loss from 0.100 to 0.068 over the training data and an accuracy increase of 0.21% over the dev data, proving our hypothesis above correct. But as we observe in part 3, there seems to be some overfitting and the model can be further improved.

Part3:

We use the same configuration as in Part2 with dropout with probability 0.5 after both the hidden layers.

Trainer: **Adam**
Epochs: **7**
Transfer function: **RELU**
Word embedding dimension: **64**
POS embedding dimension: **32**
Dependency embedding dimension: **32**
Minibatch size: **1000**
First hidden layer dimension: **400**
Second hidden layer dimension: **400**
Dropout probability after first hidden layer: **0.5**
Dropout probability after second hidden layer: **0.5**

Dropout is an approach to regularization in neural networks which helps reducing interdependent learning amongst the neurons. Using dropout we drop or switch off certain neurons in our network. This helps prevent overfitting in the network by making it learn more robust features which generalize better. Here we observe that using dropout with probability 0.5 after both the hidden layers leads to an increase in accuracy from 83.83 to 85.20.