

COMS 4705 - Natural Language Processing

Homework 3 (Programming)

Aayush Maini - am4810

Part 1

```
>>> python src/eval.py trees/dev.conll outputs/dev_part1.conll
```

Unlabeled attachment score	82.47
Labeled attachment score	79.34

Part 2

```
>>> python src/eval.py trees/dev.conll outputs/dev_part2.conll
```

Unlabeled attachment score	82.89
Labeled attachment score	79.71

The accuracy increased (slightly) because the 2nd model uses 400 units in the hidden layers. Thus, this network can capture more complex features thus enabling better classification.

Part 3

Variation	Unlabeled attachment Score on (dev)	Labeled attachment Score on (dev)
#Hidden Units = 200 + LeakyReLU + #Epochs = 7	81.95	78.41
#Hidden Units = 200 + Dropout(0.2)	83.72	80.63
#Hidden Units = (500, 500) + Dropout(0.2) + #Epochs = 15	84.33	81.17
#Hidden Units = (500, 500) + Dropout(0.3) + #Epochs = 15	84.45	81.35

- LeakyReLU performs worse than ReLU. So, I drop the idea of using this activation.
- I use Dropout layer because it is a regularization criteria and will prevent overfitting. I validate this choice by calculating score on dev.conll. Since, the score increases, I choose to use Dropout in the subsequent models.

- iii. Now I increase the number of units within hidden layers to 500 each and increase the #Epochs to 15 The score on dev.data increases because more the number of neurons, more complex the model. But it is possible that the model is overfitting, so I try another model with increased regularization. I use Dropout(0.3) in the following model.
- iv. The score on dev.conll increases while the train score is the same which implies that the model has better generalization now. I propose this as my best model.