1 RNN

The hyper parameters that I used were the same as the recommended ones in the assignment file, except the drop-out rate which I did not include in my model. The table above shows

	RNN			LSTM	
Dataset	train	test		train	test
Accuracy	99.12%	60.12%		99.5%	70.1%

my results, its clear that the 20 epochs might have over fitted the model a little too much resulting in relatively low accuracies on the test sets.

2 Proof

a) $v_j(y_j) = \max_{y_{j-1}} \sum_{i=1}^{j} S(x, i, y_{i-1}, y_i) = \max_{y_{j-1}} \left[S(x, j, y_{j-1}, y_j) + V_{j-1}(y_{j-1}) \right]$

Lets first take the LHS of the equation and expand

$$\max_{y_{j-1}} \sum_{i=1}^{j} S(x, i, y_{i-1}, y_i) = \max_{y_{j-1}} S(x, i, y_{i-1}, y_i) + \max_{y_{m-2}} \sum_{i=1}^{j-1} S(x, i, y_{i-1}, y_i)$$

Then based on the formula above we have

$$\max_{y_{m-2}} \sum_{i=1}^{j-1} S(x, i, y_{i-1}, y_i) = V_{j-1}(y_{j-1})$$

So we have it such that

$$\max_{y_{j-1}} S(x, i, y_{i-1}, y_i) + V_{j-1}(y_{j-1})$$

b) if k are the length of a layers and n is the number of layers

So the runtime would be $O(k^2 n)$

3 Viterbi

I wasn't sure which set of numbers we had to report, but these were the highest accuracies that I could attain. The top half of the numbers are the results from the **ner.test** and the bottom half are the numbers from the**ner.dev**

```
processed 51578 tokens with 5917 phrases; found: 7496 phrases; correct: 2383.
accuracy:  40.83%; (non-O)
accuracy:  83.08%; precision:  31.79%; recall:  40.27%; FB1:  35.53
              LOC: precision:  87.18%; recall:  57.60%; FB1:  69.36  1209
             MISC: precision:  67.94%; recall:  62.14%; FB1:  64.91  836
              ORG: precision:  36.34%; recall:  43.55%; FB1:  39.62  1607
              PER: precision:   4.60%; recall:   9.66%; FB1:   6.24  3844
processed 46666 tokens with 5616 phrases; found: 7573 phrases; correct: 2064.
accuracy:  37.19%; (non-O)
accuracy:  80.49%; precision:  27.25%; recall:  36.75%; FB1:  31.30
              LOC: precision:  85.78%; recall:  54.32%; FB1:  66.52  1055
             MISC: precision:  54.32%; recall:  49.36%; FB1:  51.72  637
              ORG: precision:  37.33%; recall:  45.36%; FB1:  40.95  2001
              PER: precision:   1.70%; recall:   4.12%; FB1:   2.41  3880
```