Different Running configurations for RNN:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hidden | Vocabulary | Epoch | Learning rate | Vector size | Accuracy | Note |
| 70 | 2000 | 10 | 0.0025 | 300 | 20% |  |
| 70 | 5000 | 10 | 0.0025 | 300 | 50% |  |
| 70 | 5000 | 100 | 0.0025 | 300 | 64% | Learning rate decreases to 0 |
| 70 | 9000 | 10 | 0.0025 | 300 | 49% |  |
| 300 | 5000 | 10 | 0.0025 | 300 | 83.4% | The program runs very slowly |
| 300 | 9000 | 10 | 0.0025 | 300 | 82.4% |  |
| 500 | 5000 | 10 | 0.0025 | 300 | 85% |  |
| 700 | 5000 | 10 | 0.0025 | 300 | 87% |  |
| 700 | 5000 | 20 | 0.0025 | 300 | 78.2% |  |
| 700 | 5000 | 10 | 0.00125 | 300 | 76% |  |
| 700 | 6000 | 20 | 0.00125 | 300 | 86.6% |  |
| 300 | 6000 | 100 | 0.00125 | 300 | 85.6% | Cost decreases down to 0.27835 |
| 300 | 5000 | 10 | 0.0050 | 300 | 79.4% |  |
| 300 | 5000 | 10 | 0.00125 | 300 | 86.2% |  |
| 200 | 3000 | 10 | 0.0025 | 300 | 73% |  |
| 200 | 5000 | 10 | 0.0025 | 300 | 62.4% |  |
| 300 | 6000 | 10 | 0.00125 | 300 | 82.2% |  |
| 200 | 5000 | 10 | 0.00125 | 300 | 65.8% |  |
| 200 | 6000 | 10 | 0.00125 | 300 | 57.80% |  |
| 300 | 7000 | 10 | 0.00125 | 300 | 83.2% |  |
| 500 | 7000 | 10 | 0.00125 | 300 | 85.6% |  |
| 200 | 6000 | 20 | 0.00125 | 300 | 74.40% |  |
| 300 | 5000 | 20 | 0.00125 | 300 | 76.6% |  |
| 300 | 6000 | 20 | 0.00125 | 300 | 77.80% |  |
| 300 | 7000 | 20 | 0.00125 | 300 | 77.4% |  |
| 500 | 5000 | 20 | 0.00125 | 300 | 85.8% |  |
| 500 | 6000 | 20 | 0.00125 | 300 | 87.2% |  |
| 500 | 7000 | 20 | 0.00125 | 300 | 86.4% | Cost gets down to 0.49755 |
| 700 | 6000 | 100 | 0.0025 | 300 | % |  |
| 500 | 6000 | 100 | 0.0025 | 300 | 86% | Cost gets down to 0.1612 |
| 700 | 5000 | 100 | 0.0025 | 300 | 86% |  |
| 500 | 5000 | 100 | 0.0025 | 300 | % |  |
| 300 | 5000 | 100 | 0.0025 | 300 | % |  |
| 700 | 6000 | 100 | 0.00125 | 300 | % |  |
| 500 | 6000 | 100 | 0.00125 | 300 | % |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Different Running Configurations for LSTM

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hidden | Vocabulary | Epoch | Learning rate | Vector size | Accuracy | Accuracy\_train |
| 100 | 5000 | 10 | 0.1 | 300 | 84.55 | 86.94 |
| 100 | 5000 | 10 | 0.2 | 300 | 84.03 | 90.94 |
| 100 | 5000 | 10 | 0.3 | 300 | 83.19 | 89.8 |
| 100 | 5000 | 10 | 0.08 | 300 | 84.45 | 85.84 |
| 100 | 5000 | 20 | 0.1 | 300 | 85.92 | 93.64 |
| 100 | 6000 | 20 | 0.1 | 300 | 84.55 | 95.06 |
| 100 | 5000 | 50 | 0.1 | 300 |  |  |
| 300 | 5000 | 20 | 0.1 | 300 | 85.39 | 93.42 |
| 300 | 5000 | 10 | 0.1 | 300 | 85.08 | 87.06 |
| 500 | 5000 | 10 | 0.1 | 300 | 84.55 | 87.5 |
| 700 | 5000 | 10 | 0.1 | 300 | 85.29 | 86.14 |
| 700 | 5000 | 20 | 0.1 | 300 |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

The difference between the results of LSTM and RNN is that LSTM seems more robust and reliable.