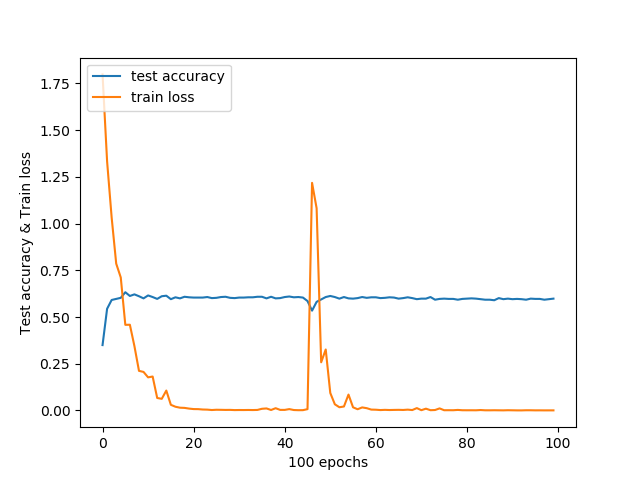
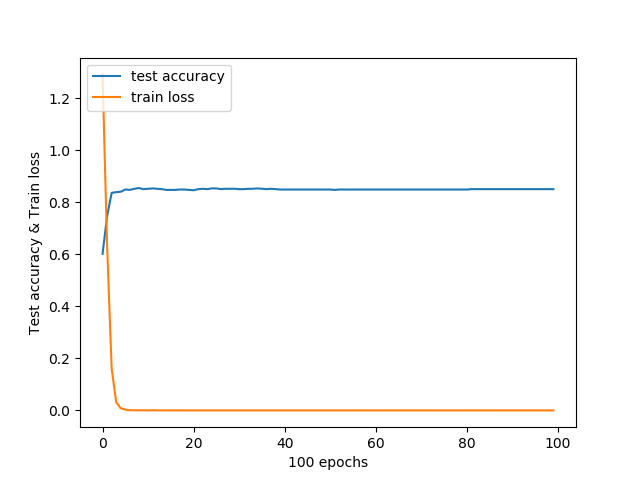
**Part B: Text classification**

**1. Plot the entropy cost on the training data and the accuracy on the testing data against training**

**epochs.**

****

* 1. **2. Plot the entropy cost on training data and the accuracy on testing data against training epochs.**
  2. ****

**3. Design a Character RNN Classifier that receives character ids and classify the input. The RNN is GRU layer and has a hidden-layer size of 20.**

**Plot the entropy cost on training data and the accuracy on testing data against training epochs.**

**4. Design a word RNN classifier that receives word ids and classify the input. The RNN is GRU layer and has a hidden-layer size of 20. Pass the inputs through an embedding layer of size 20 before feeding to the RNN.**

**Plot the entropy on training data and the accuracy on testing data versus training epochs.**

**5. Compare the test accuracies and the running times of the networks implemented in parts (1) – (4).**

**Experiment with adding dropout to the layers of networks in parts (1) – (4), and report the test accuracies. Compare and comment on the accuracies of the networks with/without dropout.**

|  |  |  |
| --- | --- | --- |
| **Model** | **Test Accuracy** | **Running Times (seconds)** |
| Q1 without dropouts | 0.5985714 | 1106 (JC’s CPU) |
| Q2 without dropouts | 0.85 | 183 (JC’s CPU) |
| Q3 without dropouts |  |  |
| Q4 without dropouts |  |  |
| Q1 with dropouts | 0.6414286 | 1030 (JC’s CPU) |
| Q2 with dropouts | 0.9028571 | 186 (JC’s CPU) |
| Q3 with dropouts |  |  |
| Q4 with dropouts |  |  |

**6. For RNN networks implemented in (3) and (4), perform the following experiments with the aim of improving performances, compare the accuracies and report your findings:**

**a. Replace the GRU layer with (i) a vanilla RNN layer and (ii) a LSTM layer**

**b. Increase the number of RNN layers to 2 layers**

**c. Add gradient clipping to RNN training with clipping threshold = 2.**