Intro Deep Learning Homework 1

Jaskin Kabir
Student Id: 801186717
GitHub:
https://github.com/jaskinkabir/Intro_Deep_Learning/tree/master/HM1

January 2025

1 Problem 1: Character Prediction Small Dataset

1a. *Training Curves*

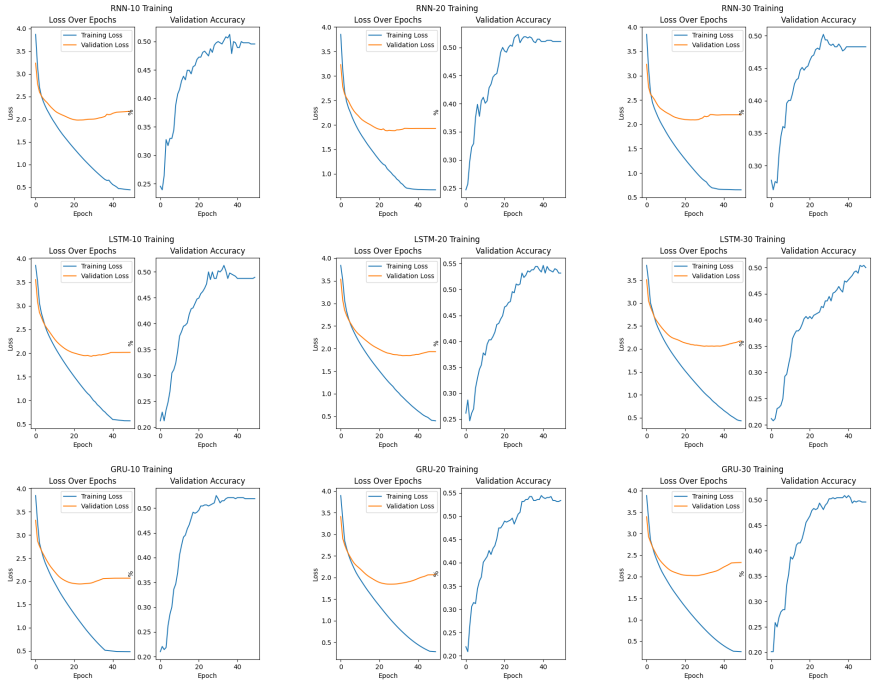

Figure 1: Problem 1 Training Curves

1b. *Results Comparison*

1c. *Discussion*

The parameter count of the LSTM network was more than triple that of the basic RNN network, and the GRU was almost halfway between the two. In terms of training time, the RNN was the quickest to train, followed by the GRU and then the LSTM. This is to be expected, as the LSTM adds significant complexity to the base RNN architecture, and the GRU reduces that complexity. This complexity comes from the number of gates in each architecture. The RNN-20 model had the lowest overfit, while the LSTM-30 model had the highest. The GRU-20 model had the highest accuracy, while the RNN-30 model had the lowest.

Model	Parameter Count	Training Time (s)	Overfit (%)	Accuracy (%)
RNN-10	44846	1.93	389.68	49.58
RNN-20	44846	3.65	185.97	51.05
RNN-30	44846	5.24	235.41	48.31
LSTM-10	143918	5.79	253.80	48.95
LSTM-20	143918	11.21	381.00	53.16
LSTM-30	143918	18.23	399.18	50.00
GRU-10	110894	5.30	329.01	51.89
GRU-20	110894	9.88	633.11	53.38
GRU-30	110894	15.41	460.00	50.85

Table 1: Problem 1 Data Comparison