ECGR 5106 Homework 3: Language Modeling with RNN, LSTM, and GRU Models

Lauren Bourque (Student ID: 801264851)

March 8, 2024

GitHub Link

Click here to view the code

Problem 1.

RNN for Text Sequence:

Sequence Values	Training Loss	Validation Accuracy	Time
10	0.10003	48.739%	2.41788 seconds
20	0.09559	50.844%	4.48389 seconds
30	0.10339	47.246%	6.62579 seconds

The loss and accuracy was best for sequence values of 20, but the training time was best for sequences of 10.

LSTM for Text Sequence:

Sequence Values	Training Loss	Validation Accuracy	Time
10	0.08919	47.899%	5.77161 seconds
20	0.14930	48.312%	10.82653 seconds
30	0.11420	47.669%	16.70444 seconds

The training loss was best for sequence values of 10, accuracy was best for values of 20, and the training time was best for sequence values of 10.

GRU for Text Sequence:

Sequence Values	Training Loss	Validation Accuracy	Time
10	0.07154	51.471%	7.28464 seconds
20	0.05164	52.531%	14.24456 seconds
30	0.06533	52.331%	21.01641 seconds

The training loss and validation accuracy was best for sequence values of 20, but the training time was lowest for sequence values of 10. Overall, training time increases as the sequence length increases and the validation accuracy was the best for sequence values of 20 for all 3 models.

Problem 2.

LSTM for Shakespeare Text Sequence:

Sequence Values	Training Loss	Validation Accuracy	Time
20	1.23858	58.201%	368.24697 seconds
30	1.23048	58.456%	486.49960 seconds
50	1.22053	59.020%	806.89997 seconds

The training loss was best for sequence values of 20, accuracy was best for values of 50, and the training time was best for sequence values of 20.

GRU for Shakespeare Text Sequence:

Sequence Values	Training Loss	Validation Accuracy	Time
20	1.37995	56.413%	471.56682 seconds
30	1.36881	56.644%	577.76836 seconds
50	1.35796	56.989%	783.59275 seconds

The training loss was best for sequence values of 20, accuracy was best for values of 50, and the training time was best for sequence values of 20. Overall, training time increases as the sequence length increases and the validation accuracy was the best for sequence values of 50 for both models.