Problem Set 4: Recurrent Neural Networks

Martin Tran

`mtran01@g.uafs.edu`

UAFS Computer Science Department

April 9, 2022

Data

The data set consists of 51 columns of decimal values where the target variable(s) is/are either 1 or 5 values. The preprocessing step of the data includes reshaping the data to add an extra dimension. Last but not least, I split the data into 80/20 for training and testing purposes.

Models

The model architecture for all tests generally follows either a gated recurrent unit (GRU), long-short term memory (LSTM), or a recurrent neural network (RNN) with an output dense layer matching the number of target variables.

Evaluation

The model uses mean squared error (MSE) as the loss and Adam as the optimizer. All of the results shown below are conducted by taking the mean MSE score of 5 iterations with 15 epochs and a batch size of 32. Of the results shown below, the SimpleRNN and sequence-to-output model performed the best in my testing. However, the only drawback is that models with SimpleRNN took the longest to train. I am not sure why it performed better than the other state-of-the-art models though.

	Seq-to-Output	Seq-to-Vec	Seq-to-Seq
SimpleRNN	**0.003**	**0.00954**	0.0088
LSTM	0.00378	0.01194	**0.00706**
GRU	0.00408	0.0129	0.0074