Timothy Horrell

Application of ML

Homework 6

**Problem 1:**

1. AirModel consists of two layers, the first being an LSTM layer which feeds a fully connected linear layer. The input of the LSTM is the input sequence given in airline-passengers.csv. The LSTM is one layer with 50 features which represent the hidden state h. The linear layer then has an output of a single value, the estimated airline passengers. This model is trained using the Adam optimizer and MSE Loss function.
2. Lookback parameter tuning:

Lookback = 4



Lookback = 5



Lookback = 6



Lookback = 7



Lookback = 8



Lookback = 10



Lookback = 20



Lookback = 30



Given these results, the best value for the lookback parameter in the data loader is with lookback=6.

**Problem 2:**

1. The transformer consists of encoder layers, decoder layers, and within these is a feed forward neural network and self-attention calculations. These attributes are established when the transformer is initialized. To start, the embedding is size 512, with the feed forward network having dimensionality of 512 also. To start with there are two encoder structures and two decoder structures before assessing output probabilities. Note that the first encoder and decoder also use an embedding layer. The transformer uses cross entropy loss and the Adam optimizer for training.
2. Training completed, sentences below:

A screen shot of a computer

Description automatically generated

1. Layers = 2

A black screen with white text

Description automatically generated

Layers = 3

A black screen with white text

Description automatically generated