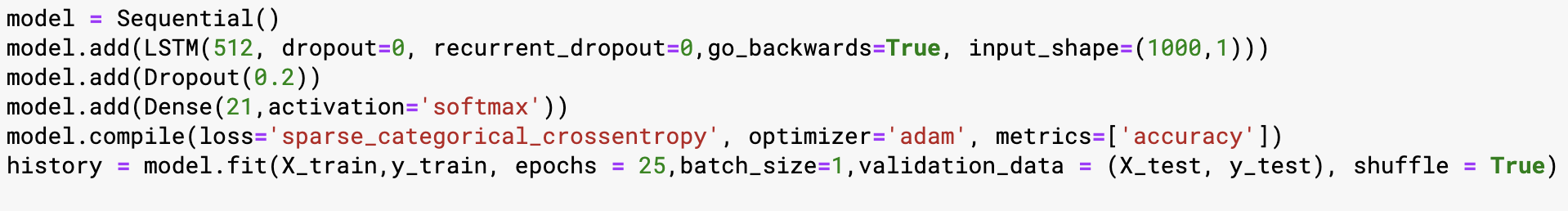
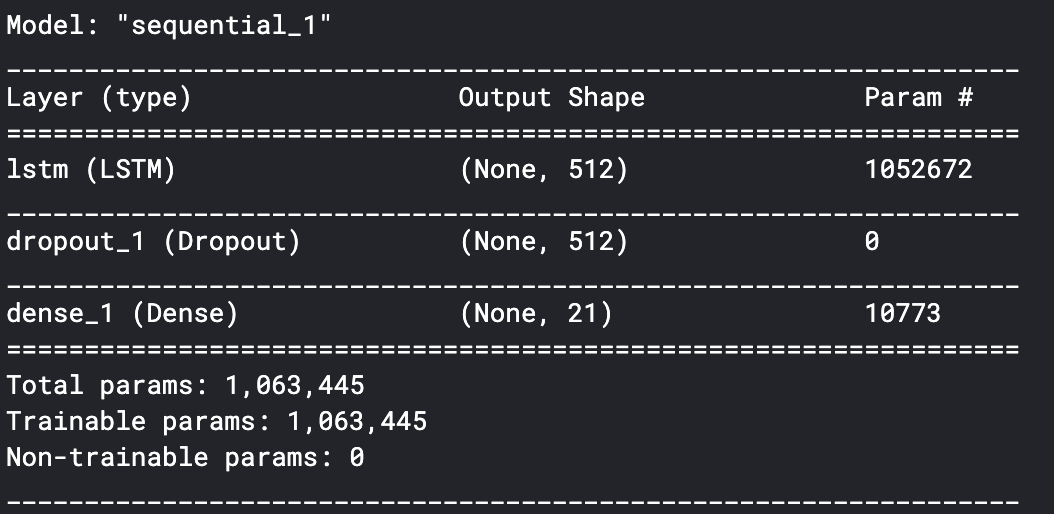
Model Architecture:





Metrics:

* Accuracy: 0.841002583503723
* Balanced accuracy: 0.71849001379358
* Precision: 0.7542789869222116
* Recall: 0.7184900137935801
* F1: 0.8144635071319509

Long-short term memory networks are a subclass of recurrent neural networks. Unlike CNN’s and ANN’s, RNN’s are not feedforwarded neural networks so they have memory of previous data and incorporate that to make their predictions. RNN’s are trained using backpropagation techniques and are specifically trained on sequential data. LSTM’s vary from RNN’s in the sense that they use special units in addition to standard units to learn longer-term dependencies Their extra memory cell allows them to more effectively keep track of patterns and sequences from previous inputs, and use this information to more accurately predict outputs. For our model, we used a simple LSTM structure of an LSTM layer, followed by a Dropout layer, and a fully connected Dense layer for classification. In our LSTM layer, we experimented with multiple neuron values and found that 512 worked the best. In our model, we set go-backwards to true, meaning we created a bi-directional LSTM. The difference between a regular uni-directional LSTM and bi-directional LSTM is that the bi-directional LSTM not only stores memory of past inputs, but also of future inputs and uses them to make predictions. Our dropout layer of 20% was added to help reduce overfitting and our dense layer consisted of 21 neurons, 21 being the number of total malware families we were to classify. We trained this model of 25 epochs with a batch size of 1. This batch size allows us to perform stochastic gradient descent while training on the samples.