National and Kapodistrian University of Athens

Department of Informatics and Telecommunications

software development for algorithmic problems

Assignment 3 - B - LSTM Autoencoder Finetuning

Project implemented by team № **59**:

CHARALAMPOS MARAZIARIS - 1115201800105
SPYRIDON CHALKIAS - 1115201800209

Contents

1. Fine-tuning

1.1 B

1.1.1 Batch Size

We experimented with batch sizes of 4, 16, 32, 64, 128, 256, 512, 1024, 4096 samples. The results showcased next, indicate that a batch size of **64** or **128** is the most preferable.

Figure 1: Batch size = 16

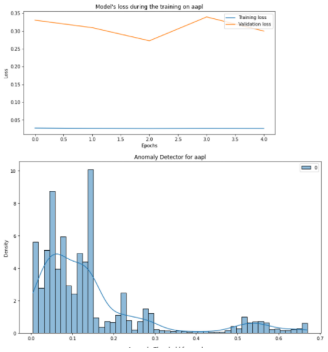


Figure 2: Batch size = 32

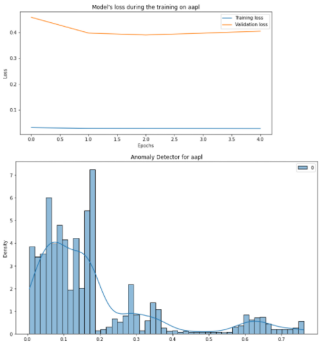
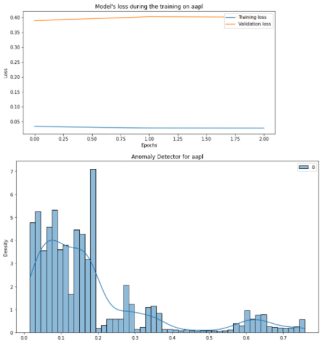


Figure 3: Batch size = 64



1.1.2 Dropout probability

We experimented with dropout probabilities of $0.0, 0.1, 0.2, 0.3, 0.4, 0.5$. Dropout layers are used to reduce the chances that our model will overfit the validation dataset. The results below indicate that a dropout probability between **20% and 30%** is more favorable,

Figure 4: Batch size = 128

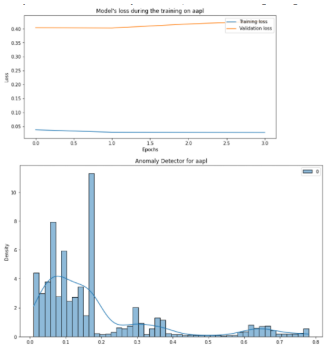
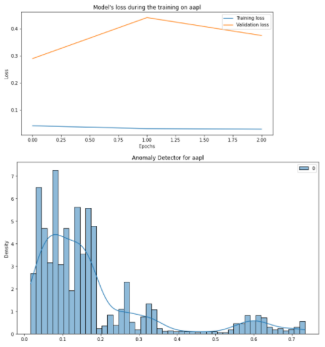


Figure 5: Batch size = 256



obtaining very low Train and Validation losses while preventing overfitting, when compared to a model with no dropout layers.

1.1.3 Filter layers and Filter sizes

We experimented with **2 to 6** layers, and more specifically:

- 2 layers: [50, 10], [100, 50], [200, 100]
- 3 layers: [50, 10, 5], [100, 50, 25], [200, 100, 50]
- 4 layers: [100, 50, 25, 10], [200, 100, 50, 25], [300, 200, 100, 50]
- 5 layers: [400, 300, 200, 100, 50], [300, 200, 100, 50, 10]
- 6 layers: [500, 400, 300, 200, 100, 50]

We present the results below.

1.2 Experiments with various Threshold values

Figure 6: Batch size = 512

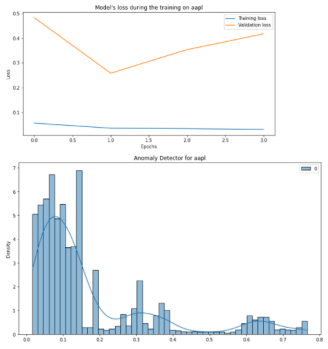


Figure 7: Batch size = 1024

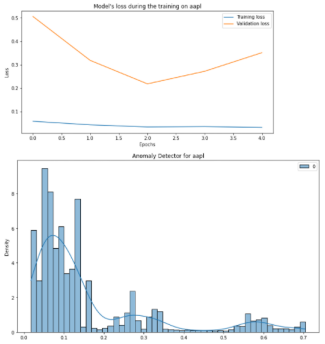


Figure 8: Batch size = 4096, Many epochs to train

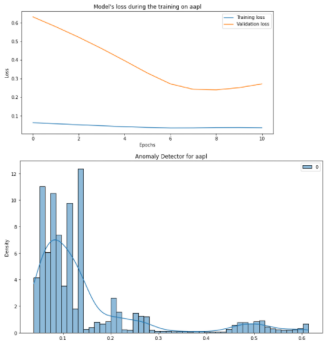


Figure 9: Dropout probability = 0%, No dropout

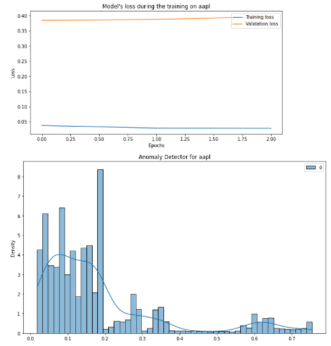


Figure 10: Dropout probability = 10%

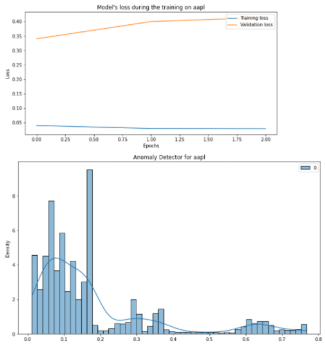


Figure 11: Dropout probability = 20%

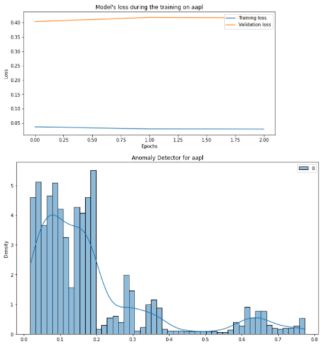


Figure 12: Dropout probability = 30%

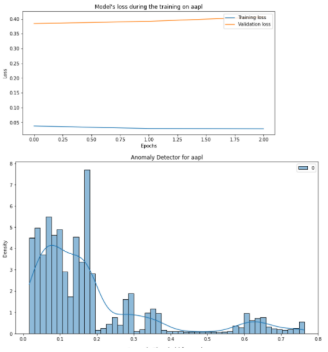


Figure 13: Dropout probability = 40%

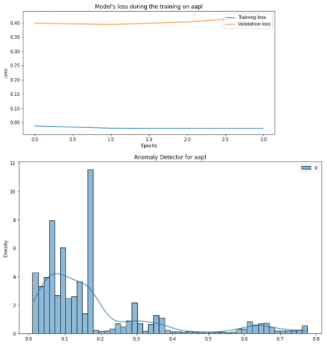


Figure 14: Dropout probability = 50%

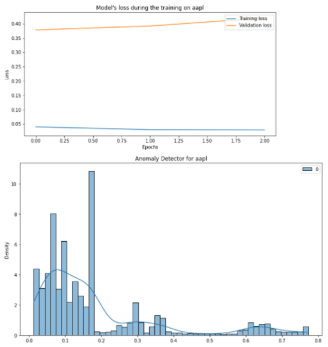


Figure 15: NN Layers = [50, 10]

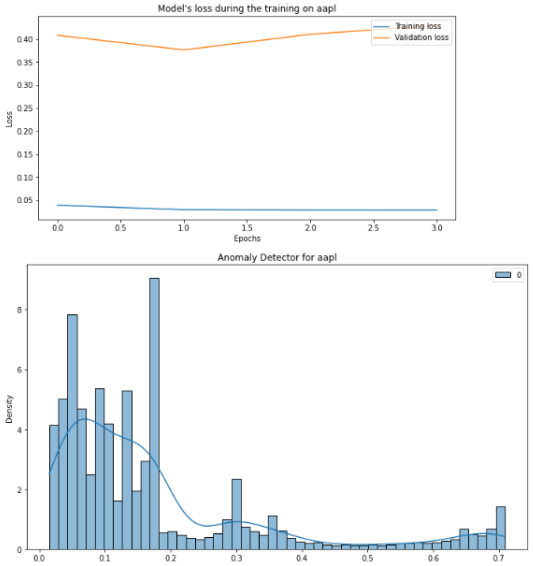


Figure 16: NN Layers = [100, 50]

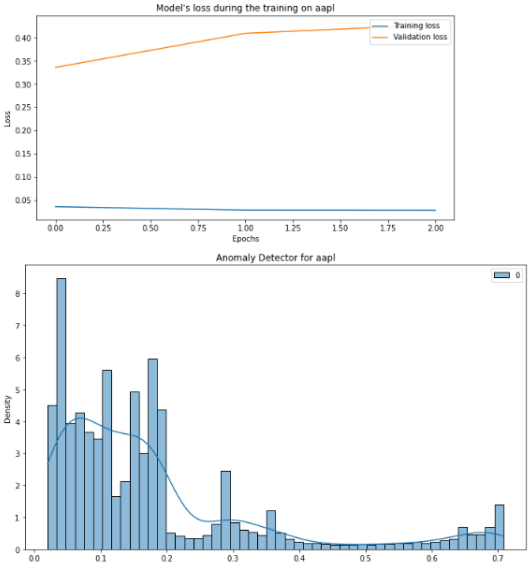


Figure 17: NN Layers = [200, 100]

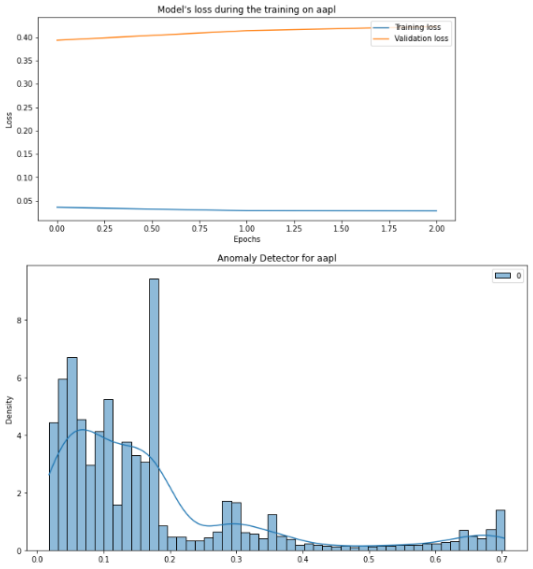


Figure 18: NN Layers = [50, 10, 5]

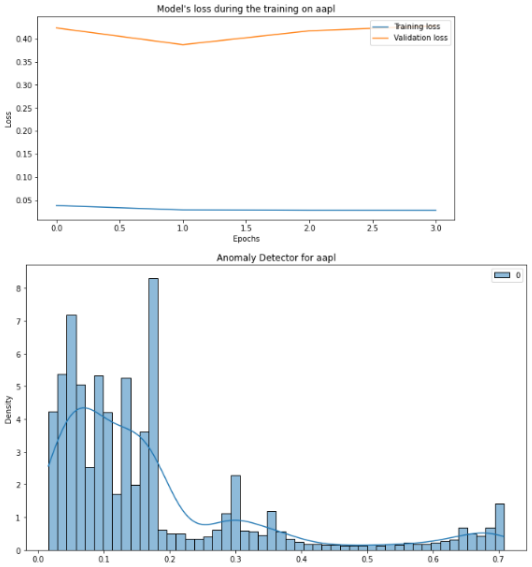


Figure 19: NN Layers = [100, 50, 25]

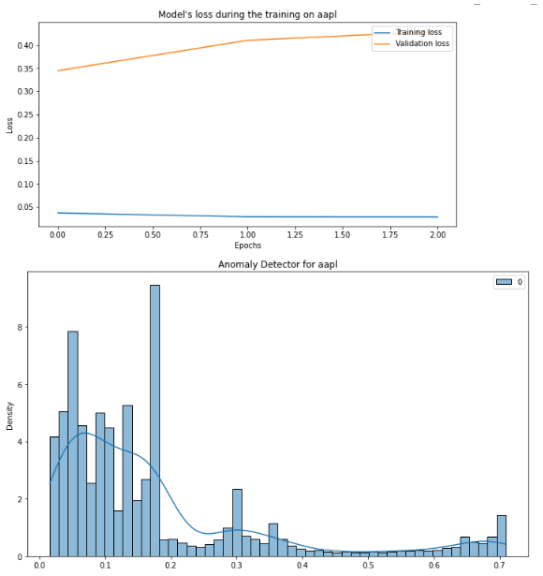


Figure 20: NN Layers = [200, 100, 50]

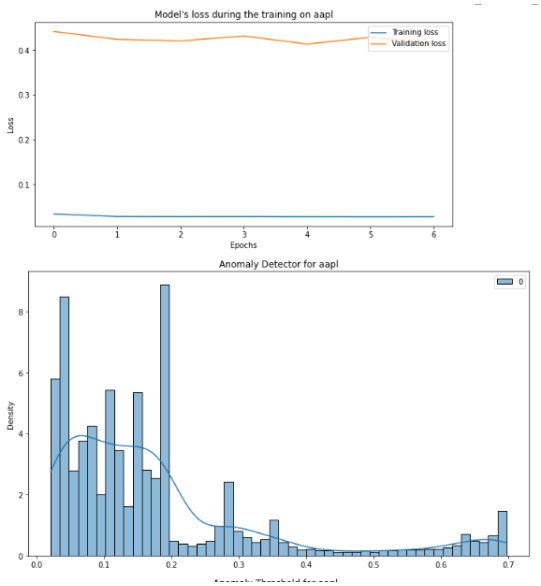


Figure 21: NN Layers = [100, 50, 25, 10]

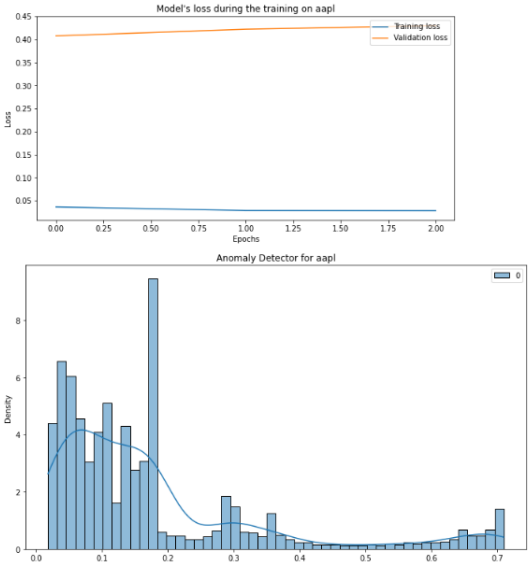


Figure 22: NN Layers = [200, 100, 50, 25]

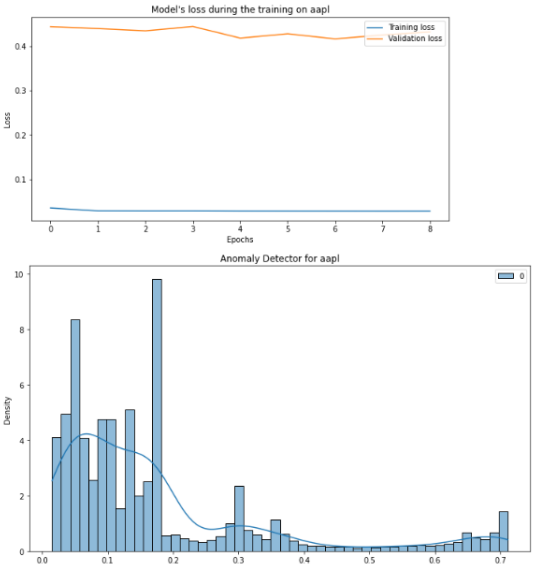


Figure 23: NN Layers = [300, 200, 100, 50]

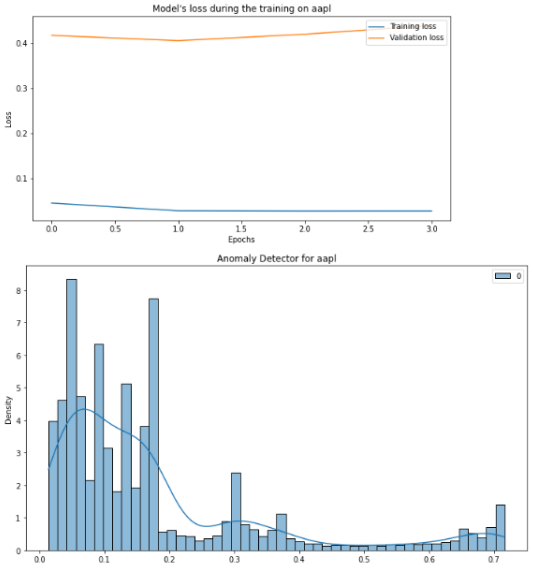


Figure 24: NN Layers = [400, 300, 200, 100, 50]

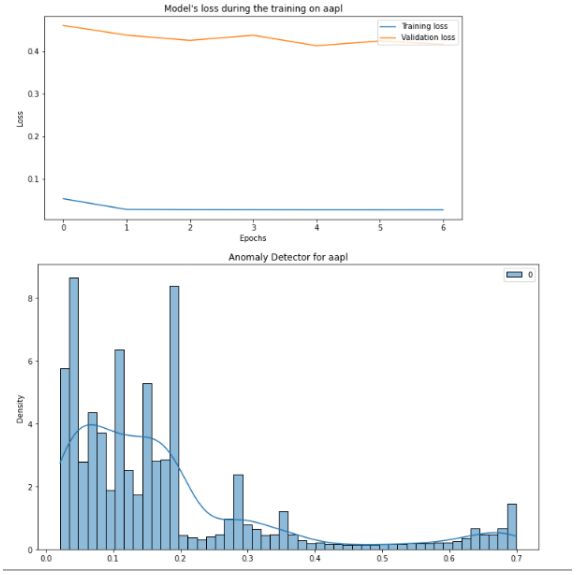


Figure 25: NN Layers = [300, 200, 100, 50, 10]

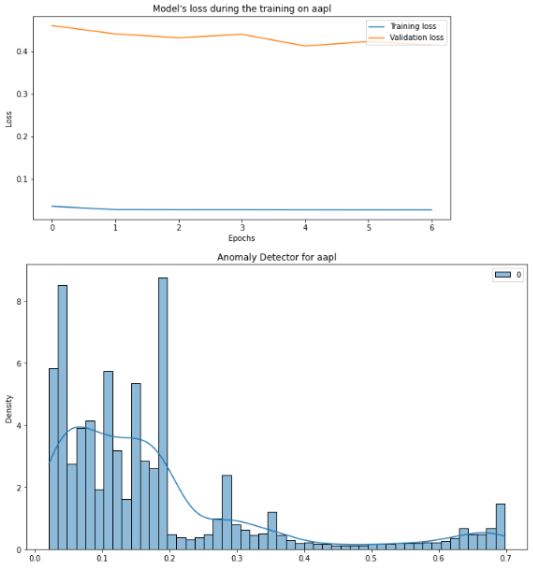


Figure 26: NN Layers = [500, 400, 300, 200, 100, 50]

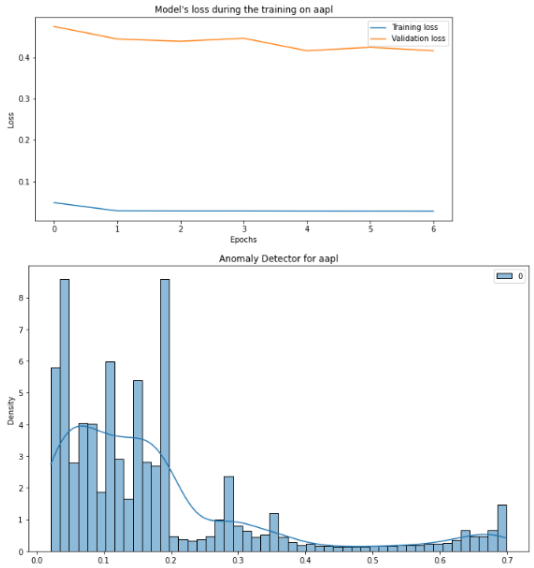


Figure 27: Threshold = 0.3

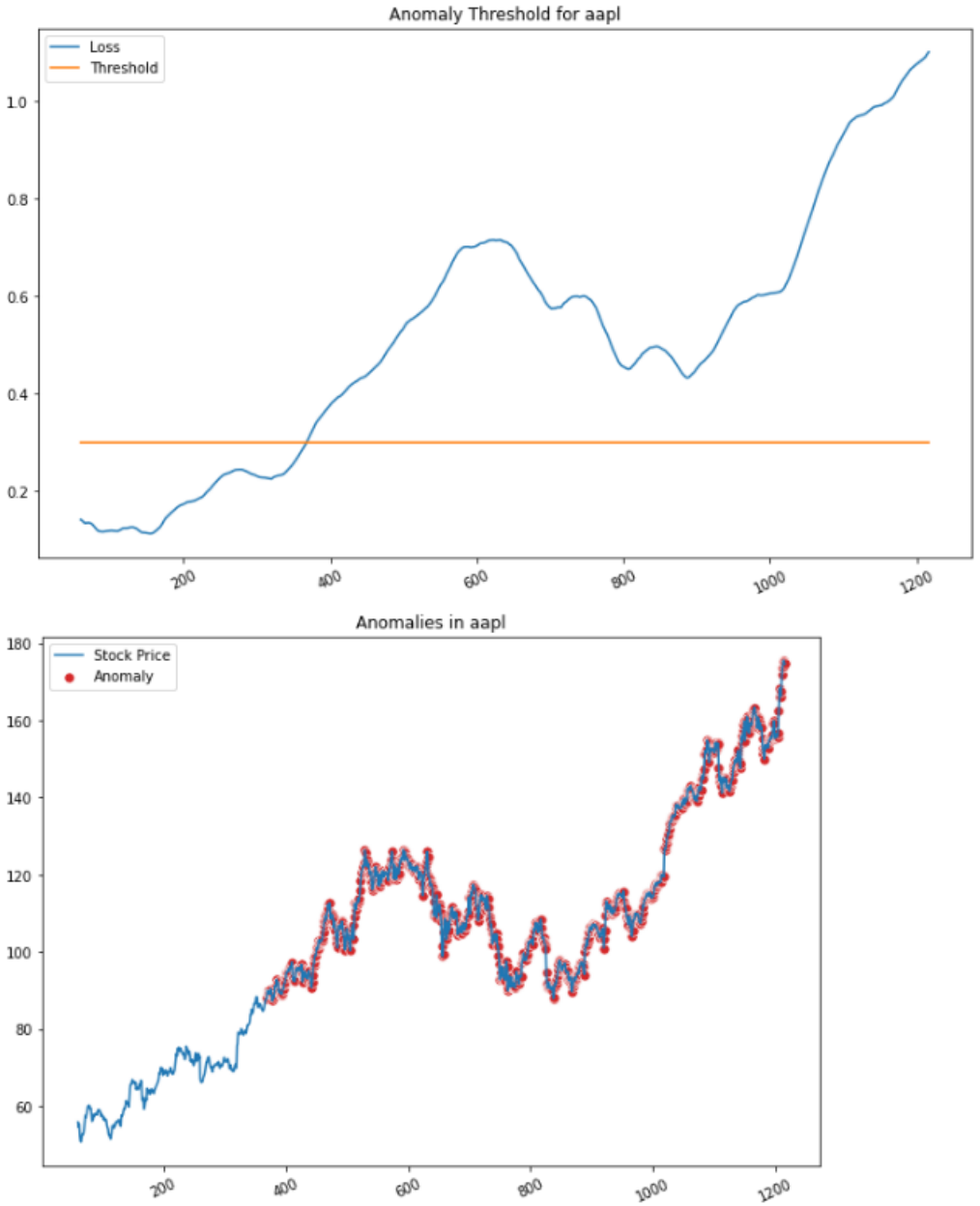


Figure 28: Threshold = 0.5

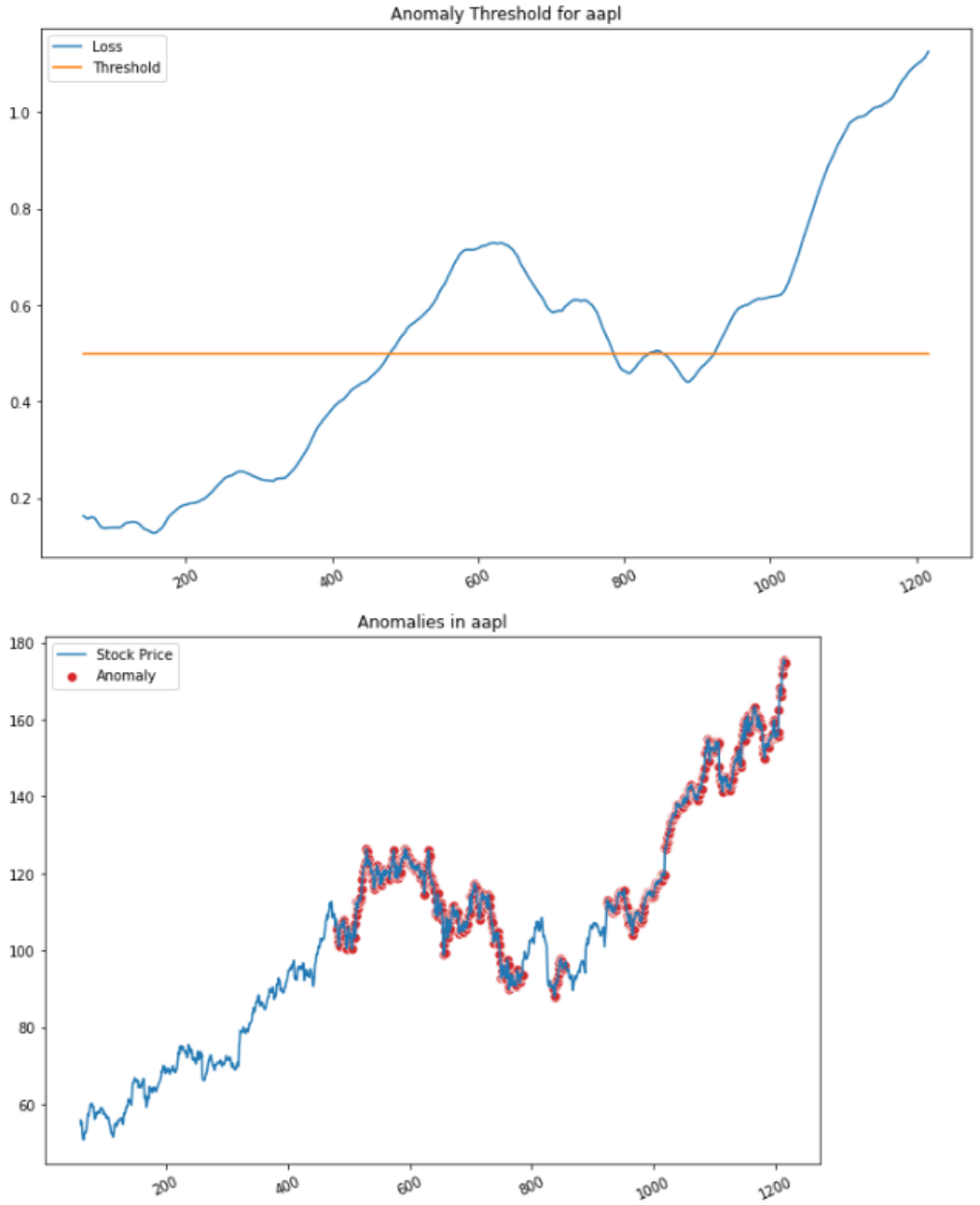


Figure 29: Threshold = 0.8

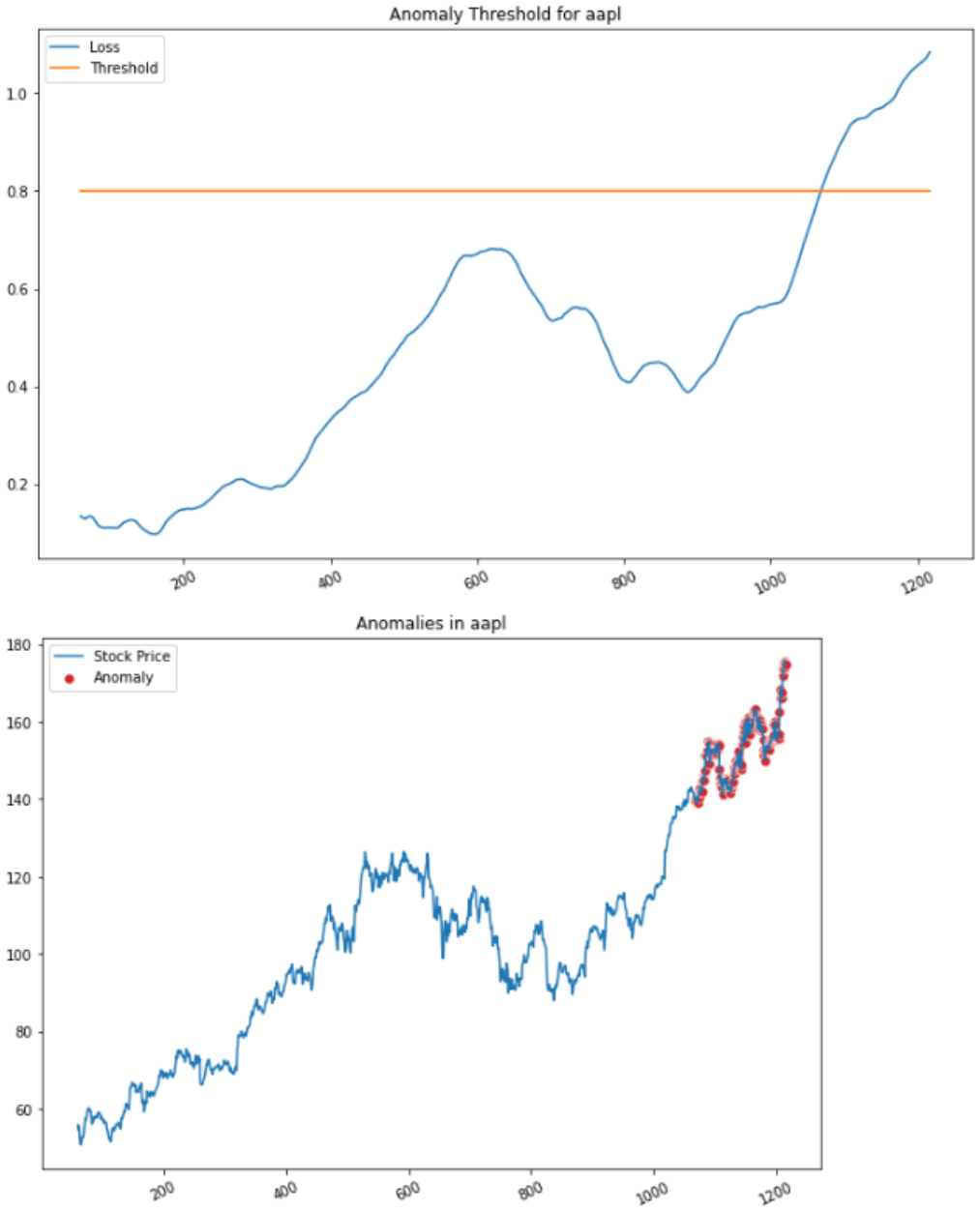


Figure 30: Threshold = 1.0

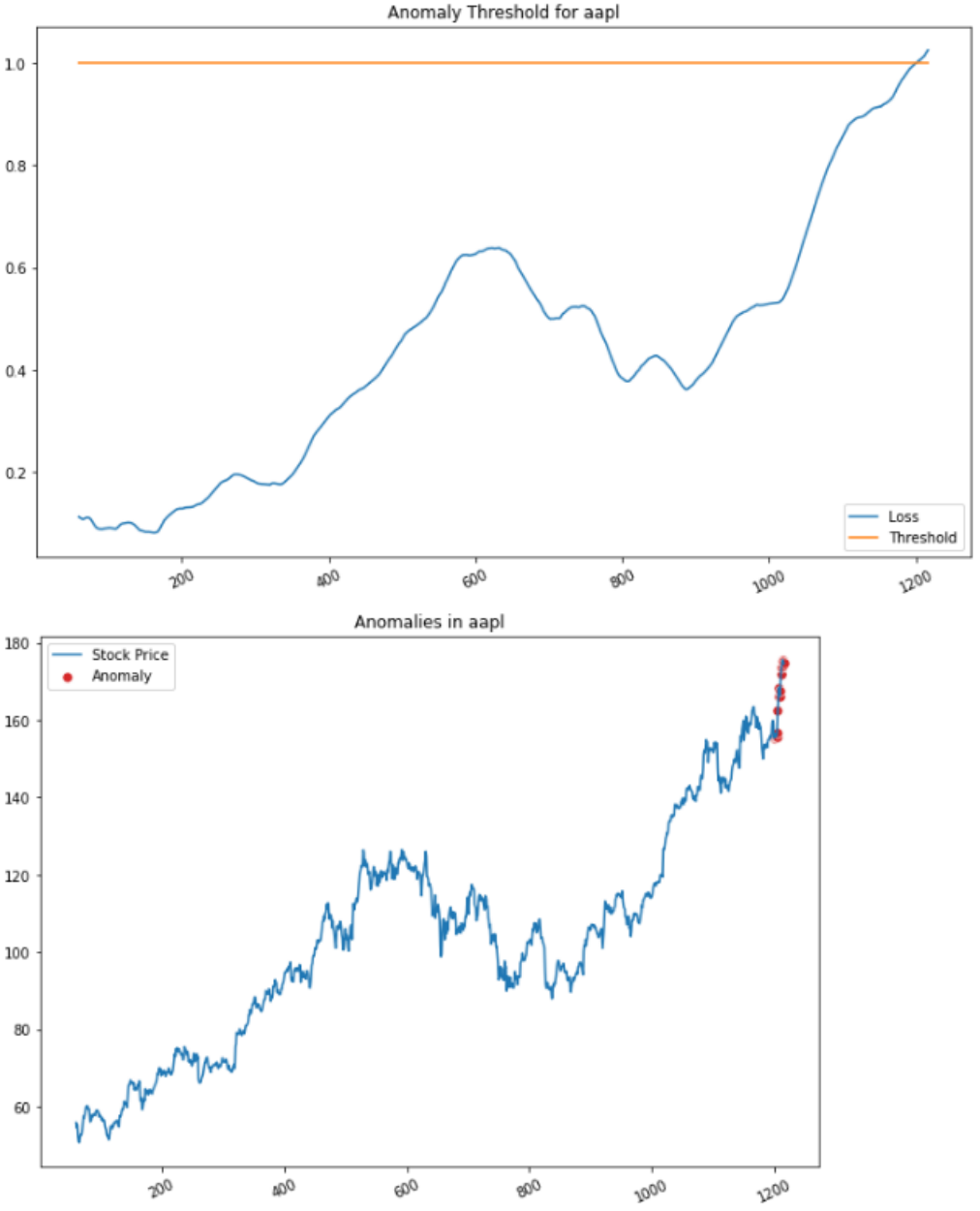


Figure 31: Threshold = 1.2

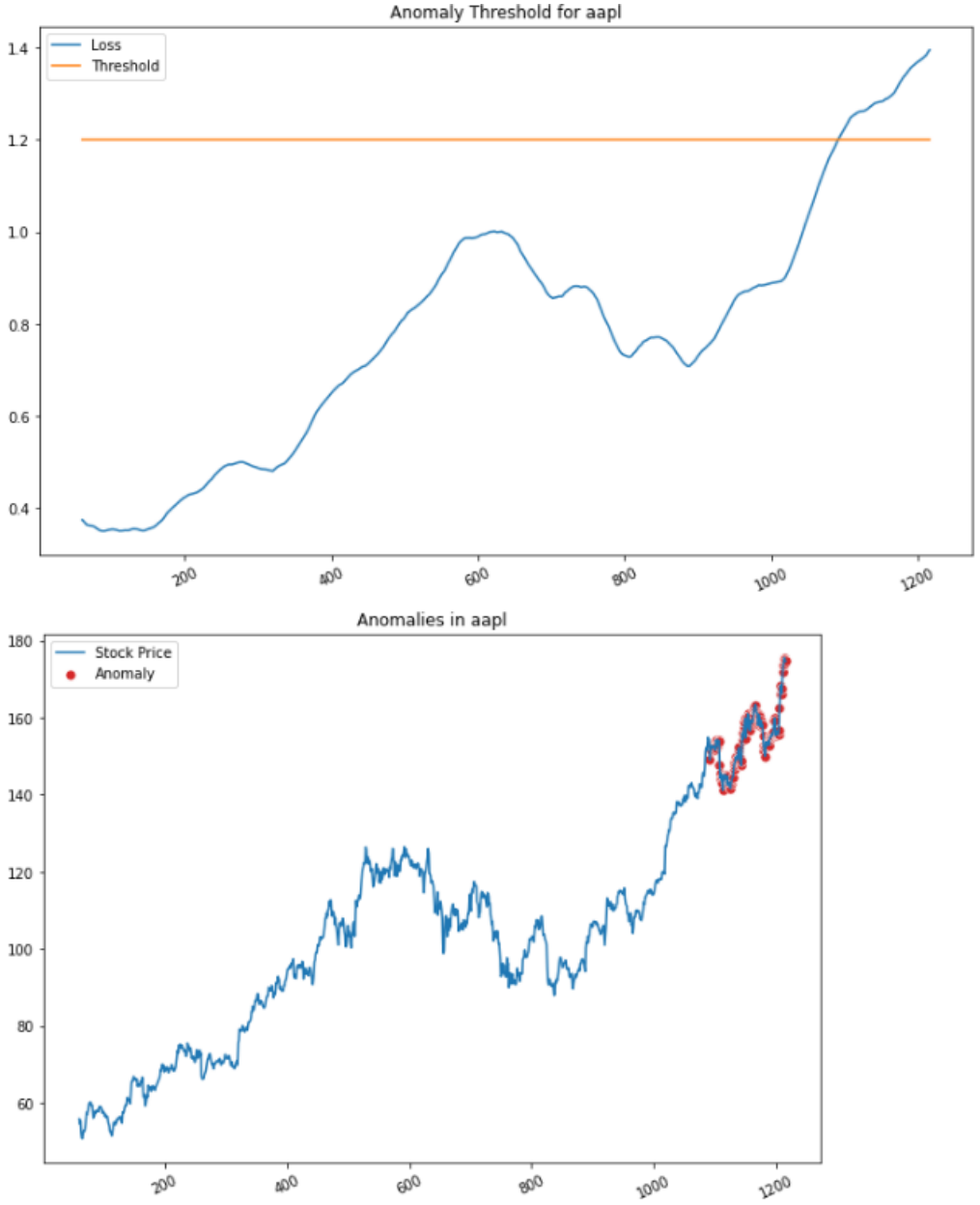


Figure 32: Threshold = 1.5