Results

Dereck Piché   Guillermo Martinez   Jonas Gabirot

April 2, 2023

1 Results

1.1 Baseline: MLP

1.2 LSTM

1.3 Regression Transformer

We built a Transfomer for our regression task. The transformer contains encoder layers. Each of the encoder layers contains 8 attention heads with a ReLU activated linear layer. We trained the transformer on 9000000 DNA sequences of length 30 and tested it's MSE loss on the remaining 100000 training examples. The transformer used the same positionnal encoding as the one proposed in [1]. The DNA sequences came in string form. Each of the 4 tokens were transformed into an integer and fed into a learned embedding layer that transformed them into vectors of 16 dimensions. The number is a bit arbitrary, but since it works really well we left it there. The accuracy troughout the epochs can be seen in figure 1.3

2 Analysis

2.1 Comparisons

3 Conclusion

References

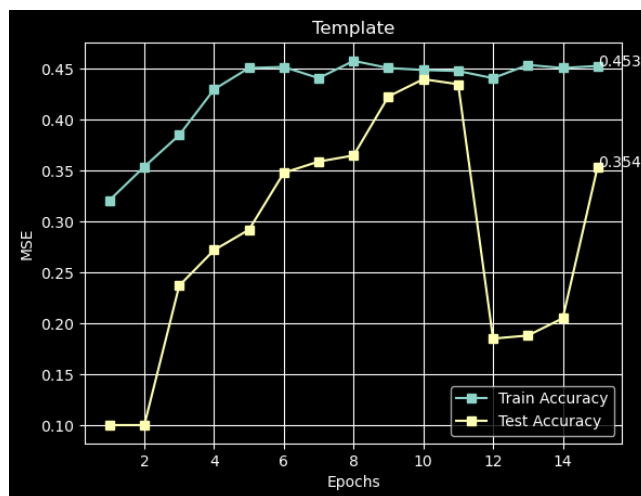[1] Ashish Vaswani et al. Attention is all you need. *CoRR*, abs/1706.03762, 2017.

Figure 1: Voici le template

Figure 2: Voici le template