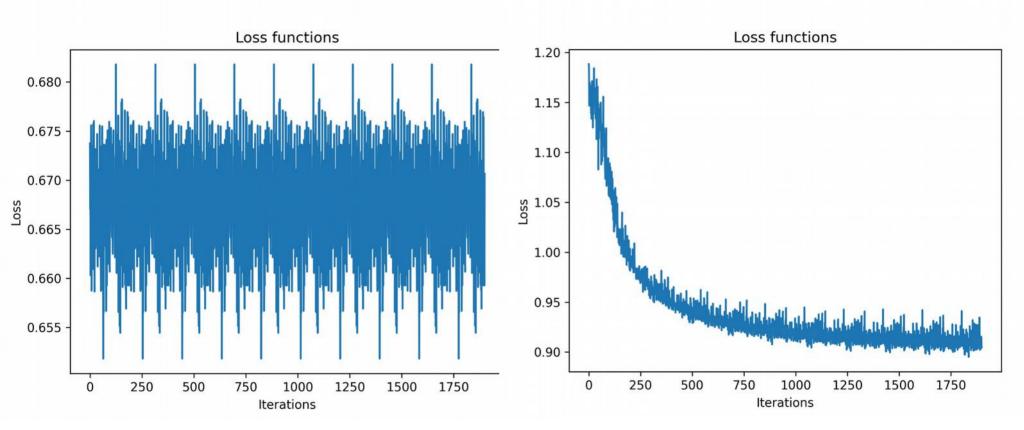
Using nn.BCEWithLogitsLoss()

Standard Clustering Model

Transformer Model



Using nn.BCELoss()

Standard Clustering Model

Loss functions 4.5 4.0 3.5 3.0 ssoj 2.5 2.0 1.5 1.0 1750 1000 1500 250 500 750 1250 Iterations

Transformer Model

