One-Class Classification in Neural Networks

E. Maboudou[1], ***P. Senaratne***[1]

[1]Department of Statistics and Data Science, University of Central Florida, USA

The performance of Artificial Neural Networks (ANNs) can decline when the neural network parameters shift over time. To maintain reliable predictions, it is crucial to monitor the neural network architecture and update or retrain the model when necessary to adapt to new data patterns.

To address this issue, we propose utilizing one-class classification techniques to monitor the latent feature representations, or "embeddings," produced by the ANN. One-class classification methods can detect changes in the data stream. If new data points begin to fall outside these boundaries, it indicates a potential change in the underlying data distribution or the parameters of the neural network, which may impact model accuracy and highlight the need for retraining.

Keywords: Embedding layer, One-class classification, Anomaly detection, LS-SVDD