

# Surrogate Loss Learning for DTW

## Abstract

In Time Series analysis Dynamic Time Warping (DTW) is a technique used to quantify similarity between two time sequences which might be varying in length and speed. Because of its ability to align the two time series, DTW is a better evaluation measure for the multi-horizon time series forecasting tasks. Research has manifested that DTW, which is a metric can also be used as a loss function for training deep learning networks. However, loss functions like DTW are non-differentiable and cannot be directly implemented as the loss function because error can not be back propagated through non-differential functions.

Considering these factors, surrogate loss functions turns out to be best alternative for non-differential or computationally complex loss functions. These functions mimic the desired evaluation metrics (or loss functions), and are highly differentiable. In this thesis, a surrogate model for computing the DTW score for the output of multi-horizon time series forecasting tasks will be developed. The developed surrogate model will be compared with the state-of-the-art loss functions and will be evaluated for the DTW score.