

Independent Study Proposal

Fall 2023

Zhengguang Wang -zw4re

Professor Judy Fox

Topic: Interpretable Time Series and Sensitivity

Overview:

We have come to an era in which we want our model not only has high prediction performance but also being trustworthy, which leads to the study of Interpretability of Black-box large neural network/Transformer models. In this Independent Study with Professor Judy Fox, I plan to delve deeper into this fascinating field. Since basic knowledge of attention mechanism and encoder/decoder architecture is a prerequisite to study interpretable, I will also put my effort in studying those as well.

Specific Study Materials:

- Prerequisites

Machine Learning with PyTorch and Scikit-Learn

The Illustrated Transformer

I will use these two resources to study Attention mechanism and Encoder/Decoder Structure.

- Interpretability Methods and Models

Factorial Sampling Plans for Preliminary Computational Experiment

I will use this paper to study Morris method.

Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting

I will use this paper to study the TFT model.

A Unified Approach to Interpreting Model Predictions

I will use this paper to study Shap method.

"Why Should I Trust You?": Explaining the Predictions of Any Classifier

I will use this paper to study LIME method

- Deep Learning Package

<https://pytorch.org/tutorials/beginner/basics/intro.html>

https://www.youtube.com/playlist?list=PL_lsbAsL_o2CTlGHgMxNrKhzP97BaG9ZN

I will use these two to study Pytorch in a systematic fashion.

Schedule:

Week	Concept/Knowledge	Assistance with Age Sensitivity Publication
10.2	Study Attention Mechanism	Run benchmarking scripts
10.9	Study Encoder/Decoder Architecture	Subject to Khairul's instructions every Monday
10.16	Study TFT Model	Subject to Khairul's instructions every Monday
10.23	Study other benchmarking Model DLinear, Autoformer, etc	Subject to Khairul's instructions every Monday
10.30	Study Morris Method	Subject to Khairul's instructions every Monday
11.6	Study Shap Method	Subject to Khairul's instructions every Monday
11.13	Study Feature Occlusion Method	Subject to Khairul's instructions every Monday
11.20	Study LIME/Feature Ablation Method	Subject to Khairul's instructions every Monday
11.27	Prepare for the paper submission	Prepare for the paper submission

Other:

I appreciate Professor Fox and Khairul for their support and guidance for us since the Summer.