NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

GEMM-SAFIN(FRIE)++

Explainable Artificial Intelligence with Episodic Memory

Motivation

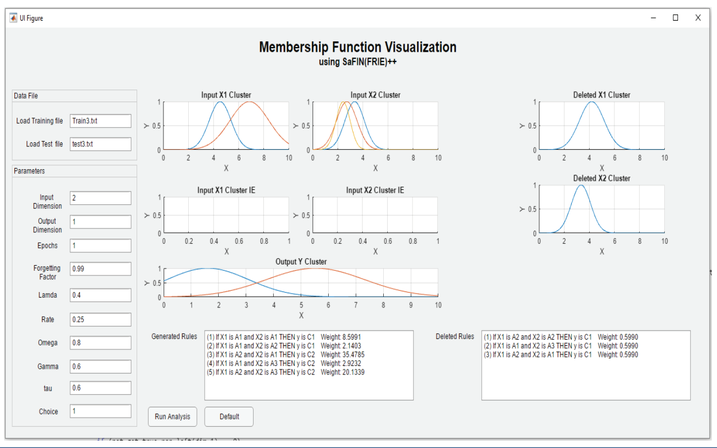
A neuro-fuzzy system is a combination of neural network and fuzzy logic that resolved the black-box nature of a neural network. Despite neuro-fuzzy has interpretability, the current model face two issues:

1. Lack of transparency between the neuro-fuzzy system and human experts – details on what makes the system arrive and formulate its rules is unknown.
2. Interpolated / extrapolated rules may be deleted – for long periods, rules that are inactive will be deleted.

Objective & Methodology

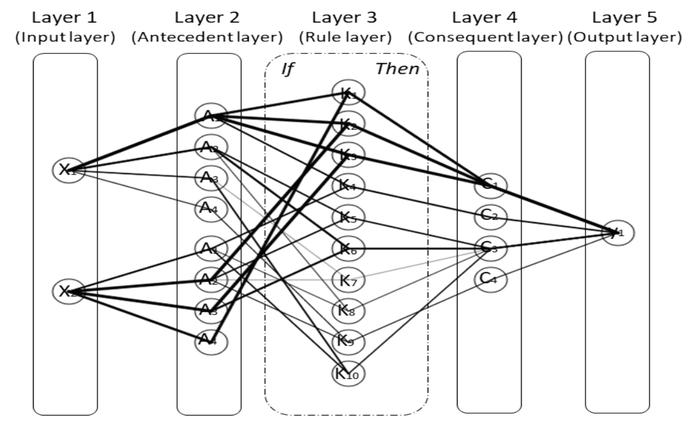
The project propose two effective explainable artificial intelligence interface for visualization and a general episodic memory mechanism for caching rules of important event.

Fuzzy Inference Learning Visualization



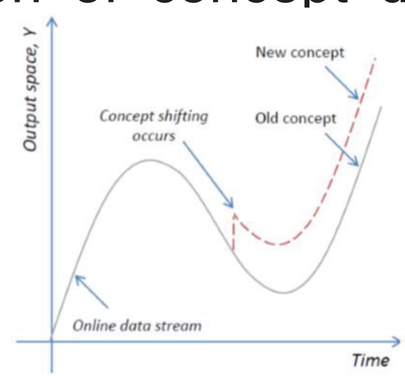
Visualize of fuzzy sets in input-output dimension and rule generation

Neuro-Fuzzy Visualization



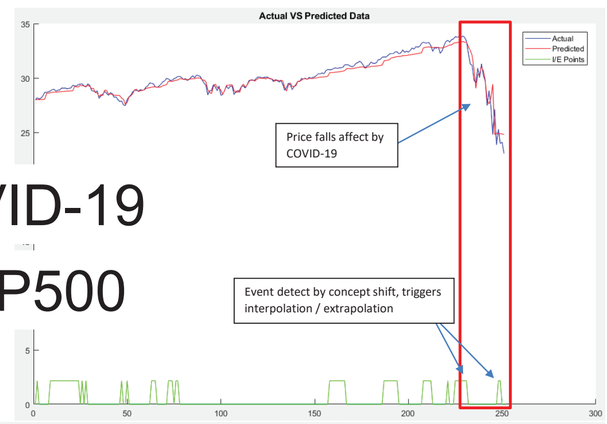
Visualize of firing rules

Interpolated / extrapolated rules are created when there is a detection of concept drift /shift. These rules are stored separately from the rulesbase as part of an event.



Analysis of Event Detection

COVID-19



S&P500

Subprime 2007

S&P500

Rules generated from such event are stored in episodic memory. When similar event occurs in the future, rules will be recall without the worry of being deleted due to inactiveness

By: Nelson Ko Mingwei | Supervisor: Assoc Prof Quek Hiok Chai