Online Non-linear Prediction of Financial Time Series Patterns

Joel da Costa

Supervisor: A/Prof. T. Gebbie

A dissertation submitted in partial fulfilment of the requirement

for the Master of Science degree in Advanced Analytics,

School of Statistical Sciences,

University of Cape Town

2019

Abstract

We consider a mechanistic non-linear machine learning approach to learning signals in financial time series data. A modularised and decoupled algorithm framework is established and is proven on daily sampled closing time-series data for JSE equity markets. The input patterns are based on input data vectors of data windows preprocessed into a sequence of daily, weekly and monthly or quarterly sampled feature measurement changes (log feature fluctuations). The data processing is split into a batch processed step where features are learnt using a Stacked AutoEncoder (SAE) via unsupervised learning, and then both batch and online supervised learning are carried out on Feedforward Neural Networks (FNN) using these features. The FNN output is a point prediction of measured time-series feature fluctuations (log differenced data) in the future (ex-post). Weight initializations for these networks are implemented with restricted Boltzmann machine pretraining, and variance based initializations. The validity of the FNN backtest results are shown under a rigorous assessment of backtest overfitting using both Combinatorially Symmetrical Cross Validation and Probabilistic and Deflated Sharpe Ratios. Results are further used to develop a view on the phenomenology of financial markets and the value of complex historical data under unstable dynamics.

Keywords: online learning, feedforward neural network, restricted Boltzmann machine, variance weight initialization, stacked autoencoder, pattern prediction, JSE, non-linear, financial time series, combinatorially symmetrical cross validation, backtest overfitting, deflated Sharpe ratio, probabilistic Sharpe ratio.