Evaluation of Machine Learning in Empirical Asset Pricing Proposal

Ze Yu Zhong

Wednesday 24th April, 2019

Remember that the entire document is meant to be no more than 5 pages, 10 max, not including appendix

includes motivation and brief review of key literature

1. Statement of Topic

This thesis aims to evaluate the application of machine learning algorithms in empirical asset pricing. While there has been

2. Background Material

3. Significance of Proposed Study

This study will be the first to

In addition,

4. Research Plan

We first evaluate machine learning algorithms on a collection of simulated datasets which include cross sectionally correlated characteristics which enter the return equation in a range of linear and highly non-linear combinations.

4.1. Models

The thesis will focus only on four different models, chosen for their prevalence and popularity in the literature.

4.1.1. Linear Model

4.1.2. Penalized Linear

4.1.3. Random Forest

4.1.4. Neural Networks

4.2. Simulation

The machine learning algorithms will be tested on these 12 specifications and their performance evaluated.

4.3. Real World Data

Finally, we evaluate machine learning algorithms on real world data.

5. Outline of Results Thus Far

6. Other Considerations

7. Research Timeline

May, June:

- Finalize Simulation of dataset
- Begin fitting models to simulated datasets
- Organise real world dataset

July:

- Finish fitting models to simulated datasets
- Finish cleaning real world data
- Begin fitting models to real world data

August:

- Collate results from models
- Research interpretation of results
- First Draft

September:

- Second Draft

October:

- Submission