

sddec18-13: Asset management - Financial Factor Discovery - "Value"

Week 7 Report

March 8 - March 21

Team Members

Carter Scheve — *Communications Lead*

Nathan Hanson — *Project Progress Tracker/Manager*

Caleb Utesch — *Meeting Scribe*

Jack Murphy — *Research Analyst*

Samuel Howard — *Lead Engineer*

Alex Mortimer — *Project Manager*

Weekly Summary

This is the seventh weekly report for our team, covering the two weeks since the last report was submitted. These two weeks covered Spring Break of course, so the amount of progress made was not as much as previous weeks. We focused our time on reviewing the changes that have been made in the last few weeks, and testing them to ensure their validity. In addition, this involved researching the techniques involved in the processes so that we could thoroughly test what we have. Our client has continued to provide more and more data to use, which is generating concerns on our part since we aren't completely comfortable with the original data, and are unsure as to what the client wants us to do with the new data.

Summary of Progress this Report

- Started implementing rolling and expanding windows methods for data input into our models, as this was a strongly recommended approach to take by our client.
 - Tested Data Importing Library to improve robustness
 - Re-examined basics of regressive models to begin implementing in the future
-

Pending Issues

- Reading into and understanding the new datasets provided by the client
 - Deciding which dataset will be most beneficial to a quality predictive model
 - Researching how to use Rolling and Expanding Window techniques with small amount of data
 - We are getting poor results predicting outperform and underperform, so trying to predict accurate performance values seems very difficult
-

Plans for Upcoming Reporting Period

- Revise Project Plan for resubmission
 - Continue looking into tree-based feature selection to try and narrow down amount of relevant/important features to input into models
 - Put together a series of comparison models for our client to display the uses, benefits, and drawbacks of each model, along with the purpose for which we are using them
-

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Carter Scheve	<p>Read over academic papers. The topics consisted of feature selection, boosting accuracy of naive bayes, and other machine learning concepts</p> <p>Created new functionality for test train split function that allowed to grab a specific window</p> <p>Used function to do basic testing of rolling and expanding window testing of models</p> <p>Updated documentation and README</p> <p>More tests and bug fixes to library</p> <p>Data recording for testing</p>	10	70
Nathan Hanson	<p>Explored some feature selection methods. Looked into standardizing algorithm output for batch job handling. Implemented CSV output on SVM code for experimentation purposes. Began experimentation with different training window types (i.e. rolling/expanding window)</p>	5	62
Caleb Utesch	<p>Familiarized myself with the new updates to Carter's data library and how to use it, as well as looked over his rolling and expanding window examples posted on the repo.</p> <p>Began looking into tree-based feature selection methods by using documentation from the scikit-learn website. Started writing code for different trials using this methods which should hopefully yield helpful results once completed.</p>	5	57
Jack Murphy	<p>Did some minor experimentation with Bernoulli Bayes Classifier. Still researching and gathering data on how to effectively predict outperformance with a reasonable accuracy score. Did some research into feature selection as well, but mainly focused on learning and working with Bayes this week.</p>	5	59
Samuel Howard	<p>Tested previous suggestion of using the autoregressive techniques on the factors to generate new ones. Since none of the factors are autocorrelated in any significant degree, this avenue of investigation has been halted. I am instead looking towards using principal component analysis to help with feature</p>	5	60

	selection.		
Alex Mortimer	Worked with Carter to improve the readability and robustness of the instructions for his Data Importing Library. Followed the instructions in the README and gave feedback on what was confusing, caused errors on my machine, or could be explained better in the eyes of someone with no experience with the tool. Iteratively improved the instruction set.	7	70