Machine Learning in Stock Trading: Clustering Pairs of Stocks

David Baines, Peter Kinder, Josh Nielsen

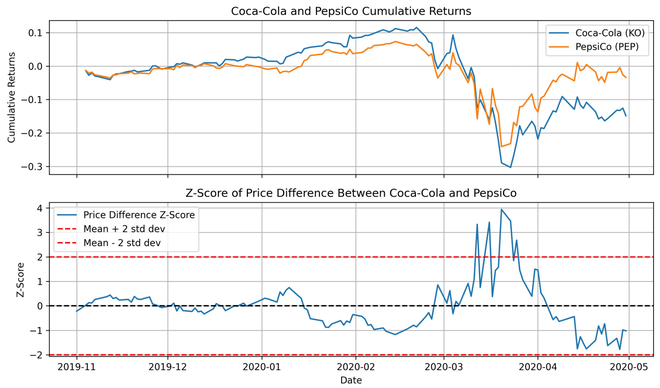
Boulder

Problem

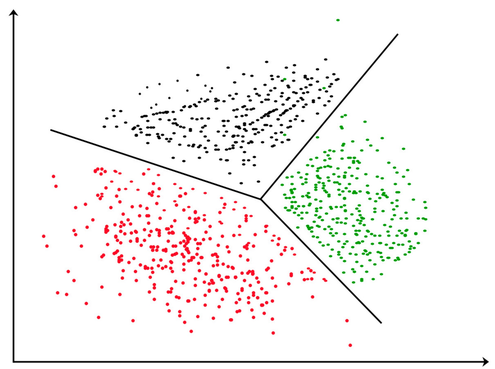
"Pairs trading" is a common stategy in **quantitative investing.**

It requires finding **two stocks that are fundamentally similar** to each other, then hoping they converge if they're trading "apart" from one another.

Can **unsupervised clustering algorithms** find profitable pairs of stocks to trade?



Approach



Extending a framework from Han, He, and Toh (2021), we test four different unsupervised clustering algorithm**s - k-means, Agglomerative, and DBSCAN,** as well an additional algorithm, **HDBSCAN**.

Our approach was to **first replicate the results** in the original paper to learn the methodology, **then introduce two enhancements to their methodology**: portfolio construction hyperperameter tuning and the HDBSCAN clustering algorithm.

We then compute a variety of descriptive analytics to determine **which clustering algorithm worked best in terms of return, volatility, consistency, and downside risk.**

Data

Time Period: 1980:01 to 2021:12
Fundamental Data: COMPUSTAT, IBES, OptionsMetrics, 13Fs
Pricing and Return Data: CRSP

We experimented with three different data sets:

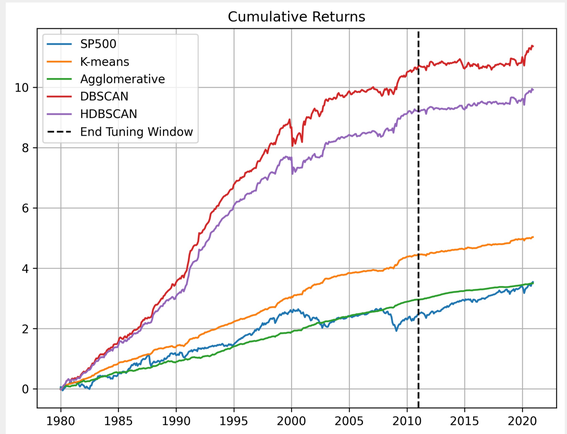
1. Compiled data to achieve feature set found in the original paper
2. Imputed data to reduce sparsity but keep information gain
3. "Raw" input data with sparsity and reduced features

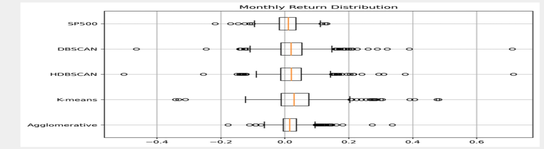
Wharton University of Pennsylvania

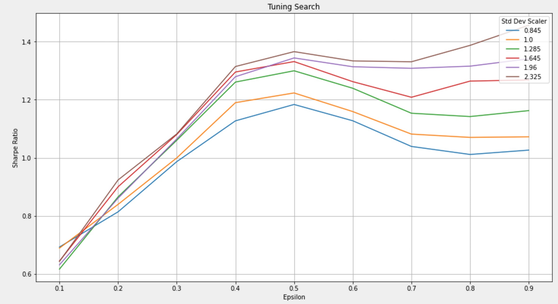
wrds WHARTON RESEARCH DATA SERVICES

Results

	Training Set (1980:01 to 2010:12, n = 372)					Testing Set (2011:01 to 2020:12, n = 120)				
	S&P 500	Agglomerative	K-Means	HDBSCAN	DBSCAN	S&P 500	Agglomerative	K-Means	HDBSCAN	DBSCAN
Cumulative Return	11.65	19.27	84.75	9969.74	41690.96	2.96	1.71	1.81	2.04	2.06
Monthly Return Mean	0.008	0.008	0.012	0.027	0.032	0.010	0.005	0.005	0.007	0.008
Monthly Return Std. Dev.	0.045	0.011	0.016	0.057	0.076	0.039	0.004	0.015	0.057	0.070
Ann. Return	0.092	0.097	0.146	0.320	0.381	0.118	0.054	0.061	0.089	0.100
Ann. Std. Dev.	0.0381	0.0551	0.1964	0.2618	0.1558	0.0135	0.0530	0.1960	0.2425	0.1352
Ann. Sharpe	2.5372	2.6404	1.6314	1.4545	0.5892	2.5372	2.6404	1.6314	1.4545	0.5892
Sortino	4.1727	6.2233	1.9324	1.9545	0.7874	11.7851	2.0669	0.8728	0.8832	1.1720
Max Drawdown	-52.56%	-4.78%	-4.52%	-43.89%	-58.73%	-20.00%	-0.75%	-5.73%	-22.12%	-29.77%
Longest DD Days	2434	304	306	822	731	397	214	275	883	1584
Avg. Drawdown	-8.31%	-0.96%	-1.13%	-4.00%	-4.78%	-5.46%	-0.32%	-1.33%	-8.81%	-13.67%
Avg. Drawdown Days	223	66	58	79	82	105	61	80	271	353







Discussion

We arrived at three main findings:

1. Adding hyperparemeter tuning allows for more performant out-of-sample portfolios

2. Some features (firm-level characterstics) hold more predictive power than others

3. The X model is more effective out-of-sample than its rivals