

Contents

1	Introduction	1
2	Denoising and Detoning	24
3	Distance Metrics	38
4	Optimal Clustering	52
5	Financial Labels	65
6	Feature Importance Analysis	74
7	Portfolio Construction	92
8	Testing Set Overfitting	105
	Appendix A: Testing on Synthetic Data	125
	Appendix B: Proof of the "False Strategy" Theorem	128
	Bibliography	130
	References	136