



THE UNIVERSITY OF
CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

SHENTAO YANG

Department of Statistics
The University of Chicago

Beating the Market with Machine-Learning

FRIDAY, January 31, 2020, at 4:00 PM
Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Empirical asset pricing studies in firm characteristics mostly focus on univariate stock ranking in an unsupervised manner to find the annual winners. In this project, we introduce a classification scheme to study “characteristics v.s. beating the market”. As an attempt to extend the classical stock-ranking approach, supervised machine-learning algorithms are employed, which conveniently digest multiple firm characteristics simultaneously. We perform a comparative study of machine-learning classifiers, including neural networks, random forest, and support vector machine. At the broadest level, we find that machine-learning algorithms on average offer an improved classification for stock winners in terms of some common evaluation metrics such as recall and F1score, compared with the classical univariate stock ranking. We also identify some highly profitable machine-learning algorithms, such as support vector machine and naïve Bayes, and some useful firm characteristics (e.g., net equity issues and book-to market ratio). Our last finding is that both machine-learning and stock-ranking approaches can construct stock-picking strategies whose returns “beat the market”.

Key Words: Classification, One-Year Return, Machine Learning, Stock Ranking, Firm Characteristics