Common Fallacies and Pitfalls of Machine Learning in Finance

Bryce Schumacher

ABSTRACT
As machine learning grows in popularity so do black box and turnkey solutions. However state-of-the-art machine learning tools remain research tools and at a minimum require some domain and machine learning expertise. Additionally the ability for machine learning algorithms to explore and describe complex datasets has progressed much more rapidly than their robustness to low quality datasets. These issues play heavily into the difficulties of using machine learning in finance. Several factors make finance difficult for machine learning practitioners. These include: overly complex algorithms, small sample sizes, unbalanced datasets, and high labeling error. These issues must be addressed before building any machine learning solution to a financial problem.

*The publication of this abstract is intended for educational purposes only from an internal symposium and its content has not been peer-reviewed.