Volume 3 | Number 4 | Fall 2021

the journal of financial data science

In asset management, alternative data are diverse nontraditional datasets utilized by quantitative and fundamental institutional investors that is expected to enhance portfolio returns. In the opening article, "Alternative Data in Investment Management: Usage, Challenges, and Valuation," Gene Ekster and Petter N. Kolm elaborate on what alternative data are, how they are used in asset management, key challenges that arise when working with alternative data, and how to assess the value of alternative databases. The key challenges include entity mapping, ticker-tagging, panel stabilization, and debiasing with modern statistical and machine learning approaches. There are several methodologies described for assessing the value of alternative datasets, including an event study methodology (which Ekster and Kolm refer to as the "golden triangle"), the application of report cards, and the relationship between a dataset's structure of information content and its potential to enhance investment returns. The effectiveness of these methods is illustrated using a case study.

In "Fairness Measures for Machine Learning in Finance," by the team of Sanjiv Das, Michele Donini, Jason Gelman, Kevin Haas, Mila Hardt, Jared Katzman, Krishnaram Kenthapadi, Pedro Larroy, Pinar Yilmaz, and Muhammad Bilal Zafar, propose a machine learning (ML) pipeline for fairness-aware machine learning (FAML) in finance that encompasses metrics for fairness (and accuracy). Various considerations for the choice of specific metrics are also analyzed. The authors discuss which of these measures to focus on at various stages in the ML pipeline, pre-training and post-training, as well as examining simple bias mitigation approaches. Using a standard dataset, they show that the sequencing in their FAML pipeline provides a sound approach for obtaining a fair and accurate ML model.

With the goal of constructing more robust portfolios, there is a growing effort to replace the Markowitz mean–variance optimization in allocation models. There are many modeling approaches for asset allocation that have gained much attention in the quantitative community. One such asset allocation model, hierarchical risk parity (HRP), reflects the hierarchical correlation dynamics of financial markets. Peter Schwendner, Jochen Papenbrock, Markus Jaeger, and Stephan Krügel in their article "Adaptive Seriational Risk Parity and Other Extensions for Heuristic Portfolio Construction Using Machine Learning and Graph Theory" present a systematic approach to generate a family of variations of the classical HRP for portfolio construction and asset allocation. The authors refer to their conceptual framework as adaptive seriational risk parity (ASRP). Backtesting the different HRP-type asset allocation variations applied to a multi-asset futures universe, the authors report that most of the static tree-based alternatives to HRP outperform the single-linkage clustering used in HRP on a risk-adjusted basis. Adaptive tree methods show mixed results, and most generic seriation-based approaches underperform.

For asset managers seeking to enhance or replace traditional factor-based investing using machine learning (ML)-driven investment strategies, interpretability, transparency, and auditability are important issues. In "Interpretable, Transparent, and Auditable Machine Learning: An Alternative to Factor Investing," Daniel Philps, David Tilles, and Timothy Law demonstrate that symbolic artificial intelligence (SAI)—a form

the journal of financial data science

of *satisficing* that systematically learns investment decision rules (symbols) for stock selection—provides a solution for dealing with these important issues while providing superior return characteristics compared to traditional factor-based stock selection and allowing for interpretable outcomes. Empirically comparing the performance of the proposed SAI approach with a traditional factor-based stock selection approach for an emerging market equities universe, the authors show that SAI generates superior return characteristics while providing a viable and interpretable alternative to factor-based stock selection. Their approach has significant implications for investment managers, providing an ML alternative to factor investing but with interpretable outcomes that could satisfy internal and external stakeholders.

Momentum is a pervasive anomaly in asset prices over time which has been shown to offer remarkable risk-adjusted performance since its first documentation in the academic literature in 1993. The empirical evidence supporting a momentum strategy has been shown not only for individual stocks but also for factor returns. In "Factor Momentum and Regime-Switching Overlay Strategy," Junhan Gu and John M. Mulvey design strategies that incorporate regime information into the portfolio optimization context by identifying regimes for historical time periods using an ℓ_1-trend filtering algorithm and exploring different machine learning techniques to forecast the probability of an upcoming stock market crash. The proposed regime-based asset allocation is then applied by Gu and Mulvey to a nominal risk parity strategy and show that a time-series factor momentum strategy generates high risk-adjusted returns. Moreover, the proposed strategy exhibits pronounced defensive characteristics during market crashes.

Categorization systems for categorizing a universe of funds has been used by market participants for many purposes, probably the most being for identifying peers for the purpose of performance analysis. The most popular system, and the one most heavily investigated in the academic literature, is the Morningstar categorization. These academic studies, which have employed unsupervised clustering techniques, have found that the Morningstar categorization provides inconsistent clustering. More recently, supervised classification techniques have been used to investigate this issue. The conclusion from these studies is that categorization is learnable with very high accuracy using a purely data-driven approach, causing a paradox. More specifically, clustering was inconsistent with respect to categorization, whereas supervised classification was able to reproduce (near) complete categorization. In "On Robustness of Mutual Funds Categorization and Distance Metric Learning," Dhruv Desai and Dhagash Mehta seek to resolve this apparent paradox by explaining incorrect uses and interpretations of machine learning techniques in the academic literature. The authors show that the Morningstar categorization system can be reproduced by using a data-driven approach if an appropriate list of variables and metrics is employed to identify the optimal number of clusters and by preprocessing the data using distance metric learning. Desai and Mehta conclude that as long as machine learning techniques are correctly implemented, the Morningstar categorization is intrinsically rigorous, consistent, rule-based, and reproducible using data-driven approaches.

The question of whether machines can learn to reliably predict auction outcomes in financial markets is examined by Nikolaj Normann Holm, Mansoor Hussain, and Murat Kulahci in "Classification Methods for Market Making in Auction Markets." To study this question, the authors apply classification methods on auction data

the journal of financial data science

from the request-for-quote (RFQ) protocol used in many multi-dealer-to-client markets. They find that the highest performance is achieved using gradient-boosted decision trees coupled with preprocessing tools to handle class imbalance. The most important features they identify are competition level, client identity, and bid–ask quotes. To demonstrate the usefulness of their findings, the authors create a profit-maximizing agent to suggest price quotes. When compared to real-world dealers, they found that the agent demonstrated significant aggressive behavior, exemplifying that a well-calibrated classifier is capable of identifying the underlying signal in detecting auction outcomes and can be used as an input in response price optimization. The authors conclude that machines can learn to reliably predict auction outcomes in financial markets.

There is a gap between benchmarking a high-frequency limit order books (LOBs) dataset and model for researchers to objectively assess prediction performances. The authors of "Benchmark Dataset for Short-Term Market Prediction of Limit Order Book in China Markets," Charles Huang, Weifeng Ge, Hongsong Chou, and Xin Du bridge that gap. The authors point out that prediction target in previous studies is too simplistic—mid-price direction change for the next few events, which is not suitable for a practical trading strategy. They suggest a more practically effective set of features to capture both LOB snapshots and periodic data. The authors present a benchmark LOB dataset from the Chinese stock market for the period of June to September 2020. Experiment protocols are designed by the authors for model performance evaluation (at the end of every second, to forecast the upcoming volume-weighted average price change and volume over 12 horizons ranging from 1 second to 300 seconds). They compare the performance a baseline linear regression and state-of-the-art deep learning models, based on both accuracy statistics and trading profits. Based on their analysis, they propose a practical short-term trading strategy framework based on the alpha signal generated from their model.

Francesco A. Fabozzi
Managing Editor