In implementing a strategic asset allocation policy, the forecasting of long-term equity market returns is critically important. Although, historically, several econometric models have been employed for forecasting, more recently machine learning methods have been used for that purpose. In “The Best of Both Worlds: Forecasting US Equity Market Returns Using a Hybrid Machine Learning–Time Series Approach,” Haifeng Wang, Harshdeep Singh Ahluwalia, Roger A. Aliaga-Díaz, and Joseph H. Davis explore machine learning methods to forecast 10-year-ahead US stock returns. To compare the relative performance of machine learning methods, the authors compare the accuracy of these methods to the forecasts of one of the most commonly used regression-based forecasting models by asset managers, the traditional Shiller cyclically adjusted price-to-earnings (CAPE) ratio model. The authors find that machine learning techniques can only modestly improve the forecast accuracy of the regression-based CAPE ratio model. Moreover, they actually result in worse performance than a vector autoregressive model (VAR)–based two-step approach introduced in 2018 by three of the authors of this article. However, when the authors implement a hybrid ML-VAR approach (i.e., VAR-based two-step approach with machine learning techniques allowing for unspecified nonlinear relationships), they find up to 56% improvement in real-time forecast accuracy for 10-year annualized US stock returns. They find the ensemble method consistently offers the best out-of-sample forecast.

Machine learning applications in finance have shown benefits over traditional linear models in forecasting stock returns. Edward Leung, Harald Lohre, David Mischlich, Yifei Shea, and Maximilian Stroh quantify these benefits by comparing the forecasting performance of commonly used machine learning algorithms with that of traditional linear methods in their article “The Promises and Pitfalls of Machine Learning for Predicting Stock Returns.” Using well-known equity characteristics, the authors forecast returns for large- and mid-cap stocks from various regional indexes using a gradient boosting machine (GBM) algorithm and standard ordinary least squares (OLS) approaches. In doing so, they shed light on the mechanics and results of the GBM model in order to alleviate its black-box character. While the forecasts from GBM models outperform OLS models based on statistical tests of forecasting performance, the economic gains from such nonlinear models depend on the ability to take the appropriate risks and efficiently implement trades.

Jochen Papenbrock, Peter Schwendner, Markus Jaeger, and Stephan Krügel in their article “Matrix Evolutions: Synthetic Correlations and Explainable Machine Learning for Constructing Robust Investment Portfolios” use evolutionary algorithms to simulate correlation matrixes useful for financial market applications. Referring to their novel approach to generate realistic correlation matrixes as “matrix evolutions,” they explain how it can be used for many applications in asset management, such as generating risk scenarios for portfolios, pricing of multi-asset derivatives, backtesting investment strategies, and hedging correlation-dependent investment strategies and financial products. The potential application of matrix evolutions is demonstrated by the authors in a machine learning case study for robust portfolio construction in a multi-asset universe that shows how an explainable machine learning program
links the synthetic matrixes to the portfolio volatility spread of hierarchical risk parity allocation versus equal risk contribution allocation. They find that the hierarchical risk parity allocations of a multi-asset futures portfolio is robust and that their approach can identify the underlying driving variables for correlation matrixes. Furthermore, it is superior to the equal risk contribution approach, which exhibits higher portfolio risk. The authors explain how the entire workflow involving matrix evolutions scales well with technologies of acceleration such as GPUs and quantum-inspired algorithms, allowing asset managers to run millions of realistic samples to simulate correlated markets.

Cross-sectional strategies are a popular style of systematic trading. Unlike time-series approaches, which only consider each asset independently, cross-sectional strategies capture risk premiums by buying assets with the highest expected returns and selling those with the lowest. What is critically important in implementing a cross-sectional systematic strategy is accurately ranking assets prior to portfolio construction. Techniques for ranking use either simple heuristics or sorted outputs from standard regression or classification models. However, these techniques have been shown to be suboptimal for ranking in other domains such as in information retrieval. In “Building Cross-Sectional Systematic Strategies by Learning to Rank,” Daniel Poh, Bryan Lim, Stefan Zohren, and Stephen Roberts propose a framework for addressing this deficiency of ranking techniques. Their proposed framework enhances cross-sectional portfolios by incorporating learning-to-rank algorithms, improves ranking accuracy by learning the pairwise and listwise structures across financial instruments. The authors show, using cross-sectional momentum as a demonstrative case study, that the use of modern machine learning ranking algorithms can substantially improve the trading performance of cross-sectional strategies. Specifically, they find that compared with traditional ranking approaches, machine learning ranking algorithms provide an approximately threefold increase in Sharpe ratios.

There is considerable evidence that abrupt changes in investor behavior occur in financial markets. To address these conditions, A. Sinem Uysal and John M. Mulvey in their article “A Machine Learning Approach in Regime-Switching Risk Parity Portfolios” propose a regime-based risk parity framework for dynamic asset allocation in which regime information is obtained via machine learning algorithms. The framework has two primary components: (1) regime modeling and prediction and (2) identifying a regime-based strategy to enhance the performance of a risk-parity portfolio. For the first component, Uysal and Mulvey apply supervised learning algorithms, based on a large macroeconomic database, to estimate the probability of an upcoming recession or a stock market contraction. They find for their out-of-sample tests that the predictions are reliable. This is particularly the case for US recessions during the period 1973 to 2020. Their probability estimates are linked to a dynamic investment overlay strategy. They report a substantial improvement in risk-adjusted, out-of-sample performance compared with a nominal risk parity in two-asset and multi-asset test cases when using the combined approach, even during the rising interest rate period in their investigation period.

Style rotation strategies, the investment thesis that active managers can enhance returns by rotating equity styles, have become increasingly popular since the 1990s. In “Style Rotation Revisited,” John Galakis, Ioannis Vrontos, and Spyridon Vrontos develop a forecasting framework for the effective implementation of style timing-rotation strategies for the US equity market. The proposed forecasting framework applies standard
univariate and multiple binary econometric models, such as multiple stepwise logit regression models and several machine learning techniques, such as linear regularization techniques (ridge, least absolute shrinkage and selection operator, and elastic net), discriminant analysis methods, Bayesian classifiers (naïve Bayes and Bayesian generalized linear models), and classification and regression tree models. To assess the predictive ability of the proposed models, the authors use several valuation metrics. The efficacy of several style rotation strategies is assessed not only in terms of their statistical significance but their economic significance as well. Their empirical analysis suggests that certain univariate logit models and machine learning techniques, such as naïve Bayes, bagging, Bayes generalized linear models, discriminant analysis models, and $k$-nearest neighbors, enhance the accuracy of the generated forecasts that lead to profitable investment strategies.

Valuable information and insights can be extracted from analyzing text data. It is important to understand the beliefs, sentiment, and disagreements among investors during times of market turbulence. In “Inside the Mind of Investors During the COVID-19 Pandemic: Evidence from the StockTwits Data,” Hasan Fallahgoul analyzes the multimodal data provided by the investing platform StockTwits during the COVID-19 pandemic. In doing so, he explores the evolution of sentiment within and across investors and sectors to gain valuable insights on how the pandemic shaped market consensus, or lack thereof. Fallahgoul finds that there is a sharp increase in disagreement and decline in sentiment at the onset of the COVID-19 pandemic between February 19, 2020, and March 23, 2020, where a historical market high was followed by a record drop. Additionally, the data show that the financial sector experienced the most pessimism, whereas the healthcare sector was the most optimistic during this time period.

In “Neural Networks, the Treasury Yield Curve, and Recession Forecasting,” Michael Puglia and Adam Tucker investigate the performance of neural network classifiers using probit regressions for forecasting US recessions with term spreads and other macro-financial data panels. They propose an econometric method for cross-validating and conducting statistical inference on machine learning classifiers and explaining forecasts. The method that Puglia and Tucker propose involves three steps. First, to address issues posed by sparse economic data when conducting analysis using machine learning methods, they employ a nested time-series cross-validation strategy. To select models and algorithms from the many candidates, in the second step they use pairwise post hoc McNemar’s tests. To aid in the economic interpretation of their results, they apply Shapley value decomposition of forecasts. They find that although the probit regression does not underperform a neural network classifier—in contrast to a growing body of literature demonstrating that machine learning methods outperform alternative classification algorithms—this machine learning algorithm does identify important features of the joint distribution of recession over term spreads and other macro-financial variables that a probit regression cannot identify (e.g., skewness and fat tails). After offering reasons for their results, they employ their three-step procedure to study US recessions over the post-Volcker period, analyzing feature importance across business cycles.

Francesco A. Fabozzi
Managing Editor