Empirical Asset Pricing via Machine Learning for Cryptocurrency

Li-Han Chang
Department of Banking and Finance, Tamkang University

Po-Wei Chen*
Institute of Statistics, Academia Sinica

Hung-Wen Cheng
Department of Data Science, Soochow University

Pei-Jie Hsiao
Department of Finance, National Sun Yat-sen University

Abstract

We apply machine learning to the empirical asset pricing problem of forecasting information coefficients (ICs) in the cryptocurrency market. Using a sample of 998 cryptocurrencies from 2018 to 2022, we evaluate linear regression, tree-based models, boosted regression trees, kernel approaches, nearest neighbors, and deep neural networks in predicting future IC values and implementing IC-based long–short strategies across multiple quantile portfolios. Linear regression (OLS) shows weak predictive power and often produces poor return performance. In contrast, machine learning models consistently outperform OLS and the market benchmark. Random forests and support vector regression deliver the most consistent performance across quantile portfolios, achieving both higher cumulative returns and superior win rates compared to the equal-weighted market portfolio. These findings highlight the role of machine learning as a viable approach to enhance portfolio construction and generate superior performance in cryptocurrency markets.

Key words: Information Coefficients, Machine Learning, Cryptocurrency, Factor Investing, FinTech