ABSTRACT

Nowadays, prediction with machine learning plays a significant role in quantitative trading. While many researchers focused on this field, most of them train models at a single stock level, partly because training multiple stocks together requires appropriate normalization methods since volatility and liquidity vary between stocks. This paper studies the prediction of the return of stock collections in three sizes in short-term trading. We find that with suitable normalization, training all the stocks together will achieve better performance. Also, predictability in short-term returns is large, systematic, and pervasive over short horizons and there will be an obvious decay when horizons become longer. We identify the relevant predictors and train models at the single stock level, industry group level, and portfolio level. Finally, we compare how predictability improves with multiple training levels, machine learning methods, and the timeliness of each stock.