Intraday Trading Strategy Selection Under Transaction Costs with Machine Learning

Alan Iván Gaona López

Universidad Anáhuac México

1 Abstract

The use of machine learning models in financial markets has recently gained relevance due to their ability to identify patterns in historical data and improve investment decision-making. This project sought to develop and compare several machine learning models capable of predicting short-term price direction of a financial asset at different intraday horizons, using historical data and technical indicators at one-minute frequency. Additionally, transaction costs were considered to ensure that profitability assessments reflected realistic market conditions. The analysis used one-minute data for Microsoft (MSFT) retrieved via the TwelveData API, covering January 1, 2022 to December 31, 2023, a two-year window with diverse market conditions that provided a representative test bed for intraday signals.

A comparative analysis was conducted using Decision Tree, Random Forest, and XGBoost, which were trained in a rolling-window scheme to capture evolving market dynamics and tested across multiple intraday horizons. Evaluation included precision, recall, F1-score, F_β-score, and cost-adjusted profit and loss (P&L) to balance predictive accuracy with economic performance. Given the high-frequency setting and the need for repeated retraining, tree-based methods were prioritized, since prior studies reported that ensembles such as Random Forest and gradient-boosted trees often matched or exceeded Support Vector Machines and deep architectures on price-direction tasks while requiring less preprocessing and computation. In addition, under the available resources, alternatives such as SVMs and Neural Networks were comparatively slower at this frequency, and therefore, they were omitted from the study.

Empirical results showed that, after evaluating the models across multiple intraday horizons with explicit transaction costs, net outcomes were sensitive to the choice of horizon and to the cost assumptions. Within this setup, the study settled on an XGBoost model at a 15-minute horizon, which achieved a positive net P&L of \$1,405.40 after accounting for transaction costs, with an average share price of \$292.71 during the test period and an average per-trade profit of 0.0079%, despite modest classification performance, with a precision of 51.70%.

This configuration provided a favorable balance between trade frequency and the precision of those trades.

Taken together, these findings indicated that explicitly modeling transaction costs was decisive for judging viability, since it prevented the overstatement of P&L that would have arisen without costs. They also showed that predictive accuracy alone was not a sufficient evaluation criterion, because profitability could diverge from classification metrics, and that, within a cost-aware setup, machine learning strategies could still achieve positive net profits.