國立臺灣大學管理學院會計學研究所

博士論文

Department of Accounting

College of Management

National Taiwan University

Doctoral Dissertation

機器學習在財務預測的應用：

月營收預測與超額報酬表現之比較研究

The Power of Machine Learning in Financial Forecasting: A Comparative Study of Monthly Revenue Prediction and Alpha Generation

官歆芸

Hsin-Yun Kuan

指導教授：王泰昌 博士

Advisor: Taychang Wang, Ph.D.

中華民國 114 年 07 月

July 2025

doi:10.6342/NTU202501949

謝辭

研究如同栽種樹木，需經過播種、施肥、灌溉與時間的累積，才能結出豐碩的果實。在校園裡，感謝學校開設多元的課程，從經濟、財金、會計到英文寫作與程式設計，都是奠定專業知識的基礎，唯有不斷學習並將所學實踐，才能轉化為研究的成果。完成這篇博士論文，要感謝的人實在太多，首先要特別感謝我的指導教授王泰昌老師，老師的悉心指導，讓我在每次遇到研究上的困難時，總能如同一盞明燈指引正確的方向。感謝口試委員陳業寧老師、曾怡潔老師、林瑞青老師與王曉雯老師給予許多寶貴的研究建議，讓研究更加完善。也感謝四位論文寫作課程的老師，林嬋娟老師、劉啟群老師、葉疏老師及陳坤志老師在論文撰寫過程中的提點。

進入博士班並順利完成論文，對我而言是一段奇妙的旅程。就像是宮崎駿《魔女宅急便》的琪琪，渴望尋找一個屬於自己的城市，我也同樣在追尋自己期許的生活與學術道路。非常感謝碩士班指導教授邱士宗老師，以及師培中心的林維能老師，正因為有老師們的鼓勵，讓我更有信心地前行。感謝我的爸爸、媽媽及妹妹，因為您們的支持讓我能專心致志於研究。感謝我的好友們，怡萱、翎瑋、柔儀、昃恩、雅雯、莉婕、鈺溱，有妳們的鼓勵讓我更有動力繼續向前。最後感謝這五年一直陪伴我的濬瑋，每當我遇到瓶頸時，總有你和我一起思考解決方案，讓問題迎刃而解。總覺得自己很幸運，身邊有這麼多人替我加油，未來我亦將秉持初心，繼續不斷學習與成長。

doi:10.6342/NTU202501949

中文摘要

本研究探討機器學習（ML）模型在營收預測與投資策略上的表現，特別聚焦於每月揭露的資料。我們評估六種 ML 模型，其中隨機森林（Random Forest）在預測準確度上表現最佳，且優於分析師預測。根據其預測建構的投資策略，在扣除交易成本後可產生年化超額報酬 51.29%。我們提出的大多數 ML 模型在報酬表現上優於大型語言模型（LLMs）與自我回歸整合移動平均（ARIMA）模型，顯示這些方法在提升投資績效方面具有明顯優勢。

關鍵字: 營收預測;機器學習;大型語言模型;分析師預測;財務分析;台灣股市;ARIMA 模型

doi:10.6342/NTU202501949

英文摘要

This study examines the predictive performance of machine learning (ML) models in revenue forecasting and investment strategies, focusing on monthly disclosures. Six ML models are evaluated, with Random Forest achieving the highest accuracy and exceeding analyst forecasts. Strategies based on its predictions yield an annualized excess return of 51.29% after transaction costs. Most of the ML models we propose generate higher returns than large language models (LLMs) and Autoregressive Integrated Moving Average (ARIMA) models, demonstrating their effectiveness in improving investment performance.

Keywords: Predicted revenue, Machine learning, Large language model, Analyst forecast, Financial analysis, Taiwan stock market, ARIMA model

doi:10.6342/NTU202501949

目次

doi:10.6342/NTU202501949

doi:10.6342/NTU202501949

doi:10.6342/NTU202501949

圖次

doi:10.6342/NTU202501949

表次

doi:10.6342/NTU202501949

1. Introduction

Data is the new oil of the 21st century, as initially stated by Clive Humby[1], and Andrew Ng further emphasizes that machine learning (ML) is the key to unlocking its value.[2] This view highlights ML's ability to extract meaningful patterns from modern financial markets' vast and intricate data. To improve the accuracy and adaptability of revenue forecasts, we employ high-frequency data, which captures short-term fluctuations and provides a more detailed view of firm performance than traditional lower-frequency datasets. As revenue has become increasingly important as a forecasting signal, we build on the approach of Kureljusic and Reisch (2022), who apply machine learning (ML) techniques to predict annual revenue. Our study extends this line of research by incorporating high-frequency monthly data to improve forecasting accuracy and by examining the relationship between predicted revenue growth and stock returns. We also benchmark the performance of ML models against large language models (LLMs) and autoregressive integrated moving average (ARIMA) models.

Revenue represents the direct outcome of a company's core operations, and its sustained growth often signals competitive advantages and market potential. Its significance stems from several factors. First, the Conceptual Framework for Financial

[1] https://en.wikipedia.org/wiki/Clive_Humby
[2] https://mitsloan.mit.edu/ideas-made-to-matter/why-its-time-data-centric-artificial-intelligence

doi:10.6342/NTU202501949

Reporting has further de-emphasized the income statement, limiting earnings' ability to fully convey information to investors due to mismatches between revenues and expenses resulting from differences in recognition timing (Barker et al. (2020)). Second, revenue is generally more resistant to discretionary manipulation, making it a more reliable indicator of firm performance. Its persistence and incorporation of future earnings and cash flow information enhance its value relevance and informational content (Chandra and Ro (2008), Core, Guay, and Van Buskirk (2003)). Finally, the growing dominance of the technology sector has amplified the importance of revenue in financial reporting. Firms driven by intellectual capital and innovation devote substantial resources to research and development (R&D) and the acquisition of intangible assets, often yielding long-term benefits not immediately captured by financial statements. This delay distorts traditional earnings-based metrics, diminishing their reliability as measures of firm value (Chen and Wu (2020), Lev (2018), Srivastava (2014)). Unlike earnings, which are influenced by the accounting treatment of capitalized expenditures, revenue reflects a more immediate and consistent measure of firm performance (Barth, Li, and McClure (2023)).

In empirical finance, revenue surprises have long been recognized for their informational value, shaping market perceptions of firm performance. Empirical research consistently documents their influence on stock prices and investor behavior (Chen and

doi:10.6342/NTU202501949

Yu (2022), Ertimur, Livnat, and Martikainen (2003), Jegadeesh and Livnat (2006)). However, advances in information dissemination have fundamentally transformed revenue surprise strategies. As financial markets become more efficient, market reactions to revenue announcements have accelerated (Chordia, Subrahmanyam, and Tong (2014), Martineau (2019)), reducing opportunities for investors to exploit these events post-announcement. In this context, the ability to accurately forecast revenue growth prior to public disclosure becomes increasingly valuable to investors.

Analyst forecasts have long been integral to capital markets, serving as essential performance benchmarks and shaping investor expectations (Houston, Lev, and Tucker (2010)). These forecasts facilitate communication between firms and external stakeholders, with earnings and revenue being the primary areas of analysis (Graham, Harvey, and Rajgopal (2005)). Revenue, in particular, is a central component of analysts' assessments, offering a direct measure of a firm's operational efficiency, growth prospects, and product differentiation strategies (Ertimur, Mayew, and Stubben (2011)). However, analyst forecasts are not universally available, and their periodic nature—typically issued quarterly or annually—limits the timeliness of information accessible to investors. Traditional analyst-driven forecasting approaches may fail to capture short-term revenue fluctuations, particularly in rapidly evolving industries where early revenue signals are essential for investment strategies.

doi:10.6342/NTU202501949

This study employs ML techniques to produce more timely revenue forecasts, offering a data-driven framework that enhances market responsiveness. Given the double-entry bookkeeping structure of accounting and the inherent interdependencies among financial variables, the field is well-suited for automated ML assessments (Libbrecht and Noble (2015), Penman (2013), Soliman (2008)). ML algorithms process large-scale, complex datasets, identify subtle patterns that may elude human analysts, and continuously refine predictions as new information becomes available. These attributes make ML particularly effective in enhancing both the frequency and accuracy of revenue forecasts. Our study incorporates high-frequency monthly revenue data to improve forecast precision and better capture short-term revenue dynamics. We further examine whether these forecasts generate tradable signals that lead to economically significant abnormal returns, addressing investors' primary focus on stock performance.

This study focuses on Taiwan's stock market for several reasons that make it an ideal setting for examining the predictive capacity of ML in financial forecasting. First, Taiwan’s unique practice of disclosing monthly revenue distinguishes it from other markets, where revenue and quarterly reports are typically released simultaneously, allowing only an assessment of the incremental informational value of revenue. Due to Taiwan Stock Exchange (TSE) regulations, firms must release monthly revenue data at least 20 days before quarterly reports (Chen and Yu (2022)). This regulatory requirement

doi:10.6342/NTU202501949

clarifies the distinction between revenue and other financial disclosures, enabling a more precise evaluation of its relationship with stock returns while mitigating distortions from other accounting variables.

Second, granular monthly revenue data availability offers a rich dataset that enhances ML model accuracy. Frequent updates improve pattern recognition, enable models to track short-term revenue fluctuations, and enhance adaptability to evolving market conditions. A greater volume of observations also mitigates overfitting by expanding the training sample, leading to more robust and generalizable predictions. Additionally, more frequent data points allow ML models to detect nonlinear relationships and subtle shifts in revenue patterns that coarser datasets may fail to capture, ultimately improving predictive performance.

Finally, Taiwan's stock market is heavily influenced by the technology sector, which accounts for approximately 50% of total market capitalization. The sector consists primarily of semiconductor, electronics manufacturing, and high-tech firms, where intangible assets and R&D expenditures are fundamental drivers of value creation. Since these expenditures are typically expensed as incurred, their benefits take time to materialize, making earnings a less reliable indicator of firm performance (Chen and Wu (2020), Wang et al. (2013), Yang and Chen (2003)). As a result, investors in technology-driven industries tend to emphasize revenue as a more timely and reliable measure of

doi:10.6342/NTU202501949

financial health.

Artificial intelligence (AI) advancements have introduced alternative approaches to financial forecasting. Generative AI, particularly LLMs such as GPT-4, has demonstrated notable strengths in text analysis, interpretation, and generation. Recent studies suggest that these models can rival financial analysts in numerical evaluation and judgment (Lopez-Lira and Tang (2024)). While LLMs attain earnings prediction accuracy on par with Neural Networks, their effectiveness in revenue forecasting, particularly in high-frequency settings, remains underexamined. This study provides a systematic comparison of LLMs and ML models in revenue forecasting.

The empirical results indicate that most ML models outperform LLMs, with Random Forest achieving the highest predictive accuracy and generating superior risk-adjusted returns in portfolio applications. The corresponding *t*-statistic exceeds the threshold of three proposed by Harvey et al. (2016), supporting both statistical and economic significance. These findings demonstrate the effectiveness of ML in financial forecasting and quantitative asset management.

There are three primary contributions of this research. First, building on Kajüter et al. (2022), who reviewed 112 studies on interim reports and acknowledged their benefits while noting that several important aspects have yet to be fully explored, we find that monthly revenue disclosures enhance financial information relevance by providing both

doi:10.6342/NTU202501949

predictive and confirmatory value, as set forth in the Conceptual Framework for Financial Reporting. Our findings indicate that ML-based revenue predictions improve forecasting accuracy, strengthening the predictive value of revenue disclosures by helping market participants anticipate firm performance. Additionally, these updates provide feedback information, enabling investors to reassess prior expectations and refine their evaluations.

Second, more frequent disclosures enhance the informational content of financial reports (Smith (2024)). While Kureljusic and Reisch (2022) focus on annual revenue projections, we employ ML models to generate monthly forecasts, providing investors with more timely financial data. A higher forecasting frequency contributes to market efficiency by improving the responsiveness of estimates. Our results indicate that Random Forest achieves a lower mean absolute percentage error (MAPE) of 8.0%, compared to 13.29% reported by Kureljusic and Reisch (2022), suggesting improved predictive accuracy under a higher-frequency setting.

Finally, with the growing interest in ML and LLMs, recent research has increasingly examined their effectiveness in financial forecasting. While prior studies focus on their ability to predict earnings per share (EPS) direction, we extend this analysis by incorporating the latest GPT-4o model and high-frequency data to evaluate whether these models can more precisely capture revenue magnitude. Our findings indicate that most of our proposed ML models perform superior to LLMs in revenue forecasting, confirming

doi:10.6342/NTU202501949

their advantage in capturing revenue trends.

The remainder of this paper is structured as follows: Section 2 reviews the literature on revenue, analyst forecasts, and ML applications in financial analysis. Section 3 provides an overview of the sample selection and research model. Section 4 presents empirical results on ML predictive accuracy and profitability. Section 5 provides additional analyses for robustness. Finally, Section 6 summarizes key findings, discusses implications, and suggests directions for future research.

2. Literature Review

2.1 The Market Impact of Monthly Revenue Disclosures

Research on revenue disclosures has predominantly focused on the U.S. market (Butler, Kraft, and Weiss (2007)) and international settings (Mensah and Werner (2008)), with an emphasis on how quarterly and semi-annual reports influence capital market behavior (Tsao, Lu, and Keung (2018)). In contrast, Taiwan's regulatory framework is distinctive, as it is the only market globally where listed and Over-the-Counter (OTC) companies must disclose monthly revenue from the prior month by the 10th of the following month. This requirement provides investors with revenue data significantly earlier than in markets that rely solely on quarterly financial statements. Consequently, these disclosures supplement quarterly earnings reports, offering additional financial insight and a core indicator of a firm's financial health. The early availability of revenue

doi:10.6342/NTU202501949

data enables investors and analysts to assess profitability and stock performance with greater immediacy, offering a timelier perspective on operational efficiency.

Since the introduction of monthly revenue announcements in 1988 and their refinement with the implementation of International Financial Reporting Standards (IFRS) in 2013, the focus has shifted from individual to consolidated revenue disclosures. As mandated by the Securities Exchange Act, firms must publish monthly operating data on the TSE website, including revenue, year-over-year comparisons, cumulative revenue, and percentage changes (Chen and Yu (2022)). These detailed disclosures enhance market transparency and protect investor interests, facilitating continuous monitoring of corporate performance and enabling more precise stock price adjustments, strengthening market responsiveness and informational integrity.

Recent academic studies have increasingly analyzed the impact of revenue information in stock price formation across different markets. Studies show that revenue surprises are associated with substantial upward movements in stock prices, illustrating its importance for investment decisions (Chen et al. (2014), Ertimur, Livnat, and Martikainen (2003), Jegadeesh and Livnat (2006)). Rees and Sivaramakrishnan (2004) explore the influence of revenue forecasts in shaping investor valuation processes. Research on Taiwan's market indicates that investor reactions to earnings and revenue disclosures differ, particularly between quarterly earnings and monthly revenue growth.

doi:10.6342/NTU202501949

Stocks exhibiting strong revenue growth are more likely to deliver superior future returns compared to those with weaker growth, emphasizing the relevance of monthly revenue disclosures in improving price efficiency and addressing behavioral biases (Wang and Lien (2022)).

2.2 The Effectiveness and Limitations of Analyst Revenue Forecasts

Sell-side financial analysts play a central role in bridging the information gap between companies and market participants, synthesizing data from public and private sources to produce research reports that include earnings projections, revenue estimates, and valuation targets (Ramnath, Rock, and Shane (2008)). These assessments often reflect prevailing market sentiment, assisting investors in portfolio decisions. Analysts' perspectives influence investor behavior and enhance price formation and informational efficiency. Their ability to distill complex information into meaningful forecasts makes their projections an essential component of investment strategy formulation, facilitating the flow of information between firms and the market.

In financial markets, revenue and earnings forecasts are widely used to assess a company's strength (Gilliam (2014), Keung (2010)). The emphasis on revenue projections illustrates their relevance in evaluating corporate value, as they offer a fundamental measure of business expansion and competitive positioning. Revenue disclosures become particularly significant when earnings reliability is compromised, such as in firms heavily

doi:10.6342/NTU202501949

engaged in R&D intensity. In these firms, investors react more strongly to revenue than earnings surprises (Kama (2009)). As a result, analysts' revenue forecasts provide investors with an additional reference point beyond earnings-based measures, enhancing the assessment of financial performance (Bilinski and Eames (2019), Huang and Hairston (2023)).

Despite their importance, the opacity of analysts' valuation processes raises questions about the reliability of their estimates (Bradshaw (2011), Brown et al. (2015)). Lorenz and Homburg (2018) identify several factors affecting the precision of revenue forecasts, including the projection horizon, timing of revisions, analysts' experience, update frequency, coverage scope, reputation, the volume of earnings estimates issued, the boldness of predictions, and past forecasting performance. Moreover, a lack of independence may introduce optimistic biases, as analysts seek to maintain favorable relationships with corporate management or stimulate brokerage trading activity (Brown, Lin, and Zhou (2022), Cowen, Groysberg, and Healy (2006), Lim (2001)), raising doubts about their objectivity and accuracy.

To improve the transparency and reliability of revenue forecasting, this study applies multiple ML models to generate data-driven predictions, offering a more structured and replicable alternative to traditional analyst estimates. By evaluating the predictive

doi:10.6342/NTU202501949

performance of different ML approaches relative to analyst forecasts, we provide insights into the effectiveness of algorithmic forecasting for market participants.

2.3 The Role of Machine Learning in Financial Forecasting

The emergence of AI has positioned ML as a transformative tool in modern finance. ML techniques are generally categorized into supervised and unsupervised learning, with the primary distinction being the presence of labeled data in training sets. Supervised learning refers to settings in which each observation includes both input features and a known output—commonly referred to as a label—allowing the model to learn the relationship between inputs and outcomes (Kureljusic and Karger (2024)). In contrast, unsupervised learning methods operate without labeled outputs and aim to uncover hidden structures or patterns within the data, such as clustering firms based on financial characteristics. This study focuses on supervised ML, which aims to minimize prediction errors when forecasting actual outcomes, making it particularly effective in assessing corporate performance.

The double-entry bookkeeping system, established by Luca Pacioli[3], provides the foundational accounting structure, capturing the interrelationships among financial statement items through a well-defined logic. With their capacity to process high-dimensional data, ML algorithms effectively model the complexities and dependencies

[3] https://en.wikipedia.org/wiki/Luca_Pacioli

doi:10.6342/NTU202501949

within financial statements and other financial data, leading to more robust and accurate corporate outcome predictions.

ML has demonstrated strong predictive capabilities for key financial market indicators, including stock returns (Gu, Kelly, and Xiu (2020)), earnings (Cao and You (2024); Chen et al. (2022); Hunt, Myers, and Myers (2022)), and revenue (Kureljusic and Reisch (2022)). Compared to earnings, revenue is less affected by cost allocation methods and accounting choices, resulting in greater data stability and lower volatility (Ku, 2011). This stability enhances the predictive accuracy of ML models in revenue forecasting. While Kureljusic and Reisch (2022) analyze revenue prediction using annual data, their study does not account for the higher frequency of revenue disclosures and their implications for market dynamics. As financial markets undergo rapid change and regulations increasingly shape strategic corporate responses, our study extends this line of research by employing monthly data to capture more granular revenue fluctuations and enhance forecasting precision. Furthermore, by linking ML-generated revenue forecasts to stock returns, we examine their role in shaping investor expectations and influencing price dynamics.

3. Sample Selection and Research Methodology

3.1 Data and Sample Selection

doi:10.6342/NTU202501949

We examine companies listed on the TSE and OTC markets from 2013 to 2022. The dataset includes daily stock returns, monthly revenues, quarterly financial statements, annual report disclosures, and yearly analyst forecasts[4], all sourced from the Taiwan Economic Journal (TEJ) database, a primary source on Taiwan's corporate activities, securities market operations, and economic indicators. The sample started in 2013 to ensure consistency in financial reporting following Taiwan's adoption of IFRS. To account for broader market risk factors, we supplement this dataset with Fama-French five-factor data from Kenneth French's developed markets factors website[5] and q-factor data from the q-factor website.[6]

Table 1 outlines the sample selection criteria. TSE regulations require companies to disclose the previous month's revenue by the 10th of each month, with extensions granted to the next business day if the deadline falls on a holiday. We exclude firms that miss the revenue announcement deadline to ensure that investors can reliably access the disclosed information. Additionally, we omit financial firms due to their distinct characteristics, which complicate comparisons with non-financial firms. Applying these selection criteria

[4] Analyst forecast data consolidate research reports from multiple brokerage firms, including Yuanta, Capital, SinoPac, JihSun, KGI, First, Fubon, Uni-President, Cathay, and Hua Nan.

[5] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[6] https://global-q.org/factors.html

doi:10.6342/NTU202501949

yields a final sample of 914,592 firm-day observations from 1,217 companies, providing a broad and representative dataset for analysis.

Table 1 Sample Selection Criteria

Descriptions	Observations	Securities
Firm-daily for all publicly held companies available from 2013 to 2022	1,025,126	1,349
Delete observations for late reports	(5,859)	(90)
Delete observations in the financial industry	(104,675)	(42)
Total number of firm-daily observations	914,592	1,217

Note: This table outlines the sample selection process. The initial dataset consists of 1,025,126 firm-day observations across 1,349 securities from 2013 to 2022. We excluded 5,859 observations (90 securities) due to late reports, ensuring data accuracy. Additionally, 104,675 observations (42 securities) from the financial industry were removed to avoid potential biases. The final sample comprises 914,592 firm-day observations across 1,217 securities.

Table 2 presents the descriptive statistics of our sample. Panel A summarizes the statistical properties of overall firm characteristics. The median daily return (r_{it}) of 0 indicates a symmetrical distribution of daily stock price fluctuations. The average actual revenue (YR) of $532.518 billion is consistent with predictions from various ML models, including Decision Tree (YTR), Random Forest ($YRFR$), Gradient Boosting (YBR), Neural Network (YNR), Nearest Neighbor ($YNNR$), and Elastic Net (YER), as well as analyst forecasts (YAF), all measured annually in billions. This similarity highlights the need for deeper analysis to distinguish differences among these predictive methods. The mean natural logarithm of market value (lnME) of 8.623 reflects variation in firm sizes, ensuring sample representativeness (see Appendix A for variable definitions). Panel B reports the annual sample distribution, where the lower number of observations in 2019

doi:10.6342/NTU202501949

(169,515) reflects the inclusion of only nine months of data.[7] In contrast, sample sizes from 2020 to 2022 remained stable, averaging approximately 248,359 observations per year, supporting the robustness of the dataset and its relevance to market conditions over the study period.

Table 2 Descriptive Statistics

Panel A: Overall firm characteristics

Variables	Mean	S.D.	P25	Median	P75
r_{it}	0.064	2.425	-0.899	0.000	0.881
YR	460.273	2600.715	25.388	60.068	187.833
YTR	456.172	2594.624	24.552	59.569	184.820
YRFR	454.674	2576.314	24.414	59.281	184.070
YBR	456.064	2591.391	24.288	57.868	184.732
YNR	465.787	2593.261	27.192	66.115	192.423
YNNR	439.688	2465.453	24.590	59.795	182.024
YER	465.919	2638.026	24.283	63.160	191.068
YAF	454.554	2465.691	27.601	63.650	194.151
ln*ME*	8.623	1.506	7.560	8.459	9.453

Panel B: Year-by-year sample size

Year	Firm-daily	Frequency (%)	Cumulative (%)
2019	169,515	18.53	18.53
2020	240,272	26.27	44.81
2021	249,980	27.33	72.14
2022	254,825	27.86	100.00

Note: This table summarizes descriptive statistics for the main variables are reported in this table. Panel A presents the overall firm characteristics, where r_{it} denotes daily stock returns. '*YR*' is the actual revenue, while '*YTR*,' '*YRFR*,' '*YBR*,' '*YNR*,' '*YNNR*,' and '*YER*' represent revenue predictions from ML models: Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net, respectively. '*YAF*' stands for analyst forecasts. All revenue ('*YTR*,' '*YRFR*,' '*YBR*,' '*YNR*,' '*YNNR*,' '*YER*,' and '*YAF*') are reported in billions, providing a consistent scale for comparison. 'ln*ME*' represents the natural logarithm of the market value, offering insight into firm size. Panel B details the firm-daily observations,

[7] From January to March, only the annual reports from the preceding two years are accessible, resulting in the unavailability of current-year information.

doi:10.6342/NTU202501949

frequency, and cumulative percentage from 2019 to 2022.

3.2 Machine Learning Approaches for Revenue Forecasting

This study applies six ML models—Decision Tree, Random Forest, Gradient Boosting, Neural Networks, Nearest Neighbors, and Elastic Net—to forecast revenue by capturing both linear and nonlinear patterns in financial data. These models range from the interpretable Decision Tree to the complex Neural Networks, allowing for a comprehensive evaluation of predictive accuracy. Random Forest and Gradient Boosting process high-dimensional data and capture feature interactions, Nearest Neighbors identifies localized patterns, Neural Networks model deep nonlinear relationships, and Elastic Net mitigates multicollinearity (Jiang, Gradus, and Rosellini (2020)). Appendix C provides an intuitive explanation of these models for greater clarity.

3.2.1 Decision Tree

A Decision Tree partitions data through sequential binary splits, maximizing target variable homogeneity within each region. Given an input x, the prediction function is:

$$f_T(x) = \sum_m c_m I(x \in R_m) \tag{1}$$

where R_m denotes the m-th partition, c_m is the constant prediction value for each region, and $I(x \in R_m)$ is an indicator function that equals 1 if x belongs to R_m and 0 otherwise. While Decision Trees are interpretable but prone to overfitting, necessitating

doi:10.6342/NTU202501949

ensemble methods such as Random Forest and Gradient Boosting for improved stability (Breiman et al. (1984), Gu, Kelly, and Xiu (2020)).

3.2.2 Random Forest

Random Forest enhances predictive accuracy by aggregating multiple Decision Trees trained on bootstrapped samples, reducing overfitting and improving generalization (Breiman (2001)). Each tree is trained on a randomly selected subset of features, reducing inter-tree correlation and enhancing robustness. The final prediction is obtained by averaging across B individual trees:

$$f_{RF}(x) = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \tag{2}$$

where $f_b(x)$ represents the output of the b-th tree.

3.2.3 Gradient Boosting

Gradient Boosting iteratively refines predictions by adding weak learners that correct residual errors (Friedman (2001), Schapire (1990)). The prediction function is:

$$f_{GB}(x) = F_M(x) = \sum_{m=1}^{M} \gamma_m h_m(x) \tag{3}$$

where M is the number of boosting iterations, γ_m is the learning rate, and $h_m(x)$ is the base learner at iteration m.

doi:10.6342/NTU202501949

3.2.4 Neural Networks

Neural Networks approximate nonlinear relationships through multiple processing layers (Aggarwal (2023), Cybenko (1989), Hornik, Stinchcombe, and White (1989), McCulloch and Pitts (1943)). Each layer transforms input data through weighted connections and activation functions. The forward propagation equation is:

$$a^{(l)} = \sigma\left(W^{(l)} a^{(l-1)} + b^{(l)}\right) \tag{4}$$

where $W^{(l)}$ and $b^{(l)}$ are the weight matrix and bias vector for layer l, and σ is an activation function (e.g., ReLU, Sigmoid). Backpropagation optimizes the weight parameters using gradient descent or Adam. While Neural Networks effectively capture complex patterns, they require significant computational resources and careful regularization to mitigate overfitting.

3.2.5 Nearest Neighbors

Nearest Neighbors predicts outcomes based on the similarity between data points in feature space (Chung, Williams, and Do (2022), Cover and Hart (1967), Fix (1985)). The prediction function is:

$$f_{NN}(x) = \frac{1}{k} \sum_{i \in N_k(x)} y_i \tag{5}$$

where $N_k(x)$ denotes the set of k nearest neighbors of x, and y_i represents their corresponding target values. The choice of k affects performance: a small k increases

doi:10.6342/NTU202501949

sensitivity to noise, while a large k smooths predictions but may obscure local patterns. Despite its simplicity, Nearest Neighbors becomes computationally expensive in high-dimensional settings due to the need for pairwise distance calculations.

3.2.6 Elastic Net

Elastic Net integrates Lasso (L1) and Ridge (L2) regularization to address multicollinearity and feature selection in high-dimensional settings (Gu, Kelly, and Xiu (2020), Zou and Hastie (2005)). Its objective function is:

$$\hat{\beta}_{EN} = \arg\min_{\beta} \left\{ \frac{1}{2n} \sum_{i=1}^{n} (y_i - X_i\beta)^2 + \lambda \left[\alpha \sum_{j=1}^{p} |\beta_j| + (1-\alpha) \sum_{j=1}^{p} \beta_j^2 \right] \right\} \tag{6}$$

where λ regulates the penalty strength, and α determines the relative contribution of L1 and L2 regularization. By incorporating both penalty terms, Elastic Net can handle correlated predictors, improve feature selection efficiency, and enhance model stability. In financial forecasting, it is applied to identify influential variables in large datasets, mitigating collinearity and improving model interpretability.

3.2.7 Model Training and Optimization

We use 60 financial statement variables covering corporate growth, profitability, asset utilization, cash flow, and risk management. The same feature set is applied consistently across all six ML models, ensuring comparability in predictive performance.

doi:10.6342/NTU202501949

The full set of variables is documented in Appendix B, highlighting their scope and relevance. A four-year rolling window is employed to train the models, ensuring that only the most recent data is used for forecasting. The window advances progressively, maintaining temporal separation between training and test sets, enhancing out-of-sample accuracy. This design is consistent with prior studies—Chen et al. (2022) adopt a three-year window, while Hunt et al. (2022) use five years of data. We employ a four-year horizon as a practical compromise and conduct robustness tests using alternative window lengths to evaluate the sensitivity of forecasting performance.

For model evaluation, we employ K-fold cross-validation following Cerulli (2022). The dataset is divided into K equal subsets, where each subset is used once as the validation set while the remaining subsets are used for training. All models are estimated using consistent hyperparameter settings to ensure comparability. Standardization is applied to maintain feature consistency. After cross-validation, each model is retrained on the full training set and validated on a holdout sample to assess robustness.

3.3 Predictive Performance of Machine Learning and Analysts

In evaluating the predictive accuracy of our ML models, we compare their revenue forecasts with those issued by analysts. Because multiple analyst forecasts exist for a given firm within a year, and some analysts initiate forecasts in the preceding year, early forecasts may deviate from actual outcomes due to significant events or economic shifts.

doi:10.6342/NTU202501949

To improve sample representativeness, we base our approach on Lai, Lin, and Liu (2011) and select the first revenue forecast issued within the fiscal year. Using December 31 as the cutoff date, we assess forecast accuracy based on annual revenue estimates. Since analysts provide only annual forecasts, we train ML models on data from the preceding four years to predict the subsequent year's revenue, ensuring comparability between the two approaches.

To assess predictive performance, we employ four widely used error metrics: Mean Absolute Error (MAE), MAPE, Mean Squared Error (MSE), and Root Mean Square Error (RMSE). These measures capture different aspects of forecast accuracy, enabling a comprehensive evaluation of model performance. MAE reflects the average absolute error in predictions, making it a clear and reliable metric that is relatively unaffected by outliers. MAPE measures the average absolute percentage error, allowing for relative comparisons across datasets, though it is sensitive to observations where actual values approach zero. MSE computes the average squared errors, penalizing larger deviations more heavily, while RMSE preserves the original measurement scale and allows for more intuitive understanding.

The error metrics are computed as follows:

doi:10.6342/NTU202501949

$$MAE = \frac{1}{n}\sum_{i-1}^{n} |y_i - \hat{y}_i| \tag{7}$$

$$MAPE = \frac{100\%}{n}\sum_{i-1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{8}$$

$$MSE = \frac{1}{n}\sum_{i-1}^{n} (y_i - \hat{y}_i)^2 \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n} (y_i - \hat{y}_i)^2} \tag{10}$$

where y_i represents the actual observed revenue (YR_i) for observation i, and $\hat{y}_i$ denotes the corresponding predicted value. For ML models, $\hat{y}_i$ corresponds to forecasts generated by Decision Tree (YTR_i), Random Forest $(YRFR_i)$, Gradient Boosting (YBR_i), Neural Networks(YNR_i), Nearest Neighbors $(YNNR_i)$, or Elastic Net (YER_i). The analyst forecast is denoted as YAF_i.

3.4 Forecasting Revenue Changes Using Machine Learning Models

Announced revenue, often called Monthly Revenue Surprise, is a key market indicator that can trigger significant price movements. Given the efficiency of market reactions, this study develops an ML-based revenue forecasting strategy to predict revenue changes ahead of announcements. Employing a rolling forecast framework, the model updates with newly released revenue data each month, ensuring forecasts remain adaptive and current. After January's revenue is announced, the model incorporates this information to predict February's revenue, continuing this process every month. To assess

doi:10.6342/NTU202501949

year-over-year revenue changes, investors can then compare forecasted revenue with the same period in the prior year, denoted as Revenue Last Year (RLY).

To quantify this comparison, we define ΔRLY as the percentage change between the predicted revenue and the corresponding period's revenue from the previous year:

$$\Delta RLY = \frac{\hat{y}_i - RLY}{|RLY|} \tag{11}$$

where ΔRLY represents the predicted year-over-year revenue change, $\hat{y}_i$ denotes the forecasted revenue from our ML models ($TR_i, RFR_i, BR_i, NR_i, NNR_i$, or ER_i), and RLY is the revenue from the same period last year. This formulation provides a systematic and objective benchmark for evaluating revenue growth expectations, enabling a direct comparison between forecasted and historical revenue trends.

3.5 Stock Portfolio Strategies Based on ML Revenue Forecasts

Stocks are allocated into decile portfolios based on predicted revenue changes following the cutoff date of the prior revenue announcement. The portfolio with the largest anticipated revenue growth is designated Portfolio 10, while the one with the smallest predicted change is labeled Portfolio 1. Our investment strategy taking a long position in the top decile portfolio based on predicted growth and a short position in the bottom decile portfolio.

The main purpose of this study is to assess the impact of revenue forecasts on stock market performance. To capture the full effect of revenue predictions, we track stock

doi:10.6342/NTU202501949

price movements from the release of the prior revenue announcement to the market reaction following the subsequent disclosure. Given the market's efficient response to revenue announcements, we forecast revenue one month in advance and hold positions until the day after the subsequent announcement. Specifically, the portfolio is formed one month prior to the announcement based on the predicted revenue. For instance, when forecasting January's revenue, the portfolio is formed based on predictions made by January 10 and held until February 11, the day after the revenue disclosure. This approach isolates the impact of revenue announcements while mitigating confounding effects from other market events (Taylor and Tong (2023)).

To evaluate the investment outcomes of these portfolios, we estimate alpha—defined as the intercept term from standard asset pricing regressions that captures abnormal returns unexplained by systematic risk factors. Specifically, we employ two widely used benchmark models: the Fama and French (2015) five-factor model and the Hou, Xue, and Zhang (2015) q-factor model. These models assess whether ML-based revenue forecasts are systematically linked to stock returns and whether predicted revenue growth translates into positive and statistically significant alpha.

4. Empirical Results

This section evaluates the predictive performance of six ML models—Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic

doi:10.6342/NTU202501949

Net—in revenue forecasting. We also assess their ability to generate excess returns and conduct robustness tests to validate their predictive capability. The analysis examines the applicability of ML models to financial forecasting.

4.1 Revenue Forecast Accuracy of Machine Learning and Analyst Estimates

We first compare the predictive accuracy of six ML models with analyst forecasts in annual revenue prediction. Model performance is assessed using four standard metrics: MAE, MAPE, MSE, and RMSE, where lower values indicate greater precision. These measures are widely used to evaluate forecasting accuracy. We refer to Lewis (1982), who notes that MAPE facilitates model comparability. A MAPE below 10% indicates high accuracy, 10–20% suggests good predictions, 20–50% reflects reasonable predictions and values exceeding 50% imply poor accuracy.

Table 3 reports the revenue prediction accuracy of ML models and analyst forecasts. The Random Forest (*YRFR*) model demonstrates the highest predictive accuracy, yielding the lowest errors across all metrics, including MAE, MSE, RMSE, and an MAPE of 10.030%. In contrast, analyst forecasts (*YAF*) exhibit substantially larger error magnitudes in terms of MAPE (24.664%), but still outperform four ML models—Gradient Boosting (*YBR*), Neural Network (*YNR*), Nearest Neighbor (*YNNR*), and Elastic Net (*YER*)—across several error measures. The differences in predictive accuracy are statistically significant for MAPE ($p = 0.0355$) and marginally significant for MAE and RMSE ($p = 0.0586$),

doi:10.6342/NTU202501949

suggesting that the superior accuracy of tree-based ML models, particularly Random Forest, likely stems from their ability to capture nonlinear relationships and complex interactions in financial data, enabling more precise revenue forecasts compared to analyst estimates.

4.2 Assessing the Forecast Accuracy of Machine Learning Models for Revenue

We evaluate the predictive accuracy of ML models in forecasting next-month revenue using a four-year rolling window for training. Model performance is assessed using MAPE[8], as it provides an intuitive and comparable measure of forecasting accuracy, remains unaffected by differences in firm size, and effectively evaluates model performance across varying revenue scales. Figure 1 presents the MAPE for six ML models. The Decision Tree records zero training error but a nonzero test error, indicating overfitting. This observation aligns with prior literature, as Decision Trees are highly flexible and can fully capture patterns in training data, but they also learn noise, leading to weak generalization (Kotsiantis (2013)). Ensemble learning mitigates this issue by aggregating multiple trees to enhance robustness. Among the models, Random Forest achieves the lowest MAPE in training and test sets, demonstrating superior predictive

[8] MAPE provides a standardized measure of prediction accuracy across firms by capturing relative error proportions. Since each monthly revenue forecast corresponds to a distinct MAPE value, we use the median MAPE as the primary accuracy metric to mitigate the influence of extreme values.

doi:10.6342/NTU202501949

performance. In contrast, other ML models exhibit higher MAPE values, suggesting their limited ability to capture revenue patterns effectively.

Next, we compare predicted and actual average revenue, directly assessing their deviations. Figure 2 illustrates the predicted versus actual revenue trends from 2019 to 2022 across six ML models. The dashed orange line represents revenue forecasts, while the solid gray line denotes actual revenue. Random Forest demonstrates substantial predictive accuracy, closely tracking revenue movements over time, highlighting its reliability even in complex revenue environments. In contrast, Neural Network, Nearest Neighbor, and Elastic Net models display greater deviations from actual revenue but still capture the overall trend.

We extract its feature importance rankings—calculated based on the matrix of variable importance used when building the classifier. The values are scaled proportional to the largest value in the set—to further analyze Random Forest's predictive advantage. The results identify accounts payable and notes payable, accounts receivable and notes receivable, current liabilities, net operating income, and operating expenses as the most influential factors in revenue forecasting (Figure 3). Net operating income and expenses reflect a firm's fundamental business performance, capturing cash flow and profitability. Accounts receivable and accounts payable indicate sales and procurement activities, driving future working capital and revenue fluctuations. Current liabilities represent

doi:10.6342/NTU202501949

short-term financial obligations affecting a firm's ability to sustain growth. Combining these features enables Random Forest to leverage historical financial data, cash flow patterns, and operational strategies to enhance monthly revenue predictions

4.3 Portfolio Performance from Machine Learning Revenue Forecasts

While some ML models do not achieve the highest revenue forecasting accuracy, investment applications do not necessarily require precise point estimates of individual firms' revenue. Instead, the ability to distinguish between firms with stronger and weaker revenue growth prospects is more relevant. Gu, Kelly, and Xiu (2020) emphasize that ML models' ranking ability is more critical in investment applications than point forecast accuracy. Following this principle, we construct decile portfolios based on the relative revenue changes predicted by ML models and implement a long-short (10−1) strategy. Even if specific ML models exhibit larger absolute prediction errors, as long as they capture the directional trend, they can still generate excess returns. Additionally, we seek to examine whether Random Forest's predictive accuracy translates into superior investment performance.

This study evaluates the abnormal returns generated by portfolios constructed based on revenue forecasts from various ML models, as detailed in Table 4. The table presents cumulative abnormal returns (*CAR*) associated with forecasted annual revenue changes (ΔRLY). Panel A reports *CAR* from the day following the last revenue announcement to

doi:10.6342/NTU202501949

Table 3 Revenue Prediction Accuracy of Machine Learning Models and Analyst Forecasts

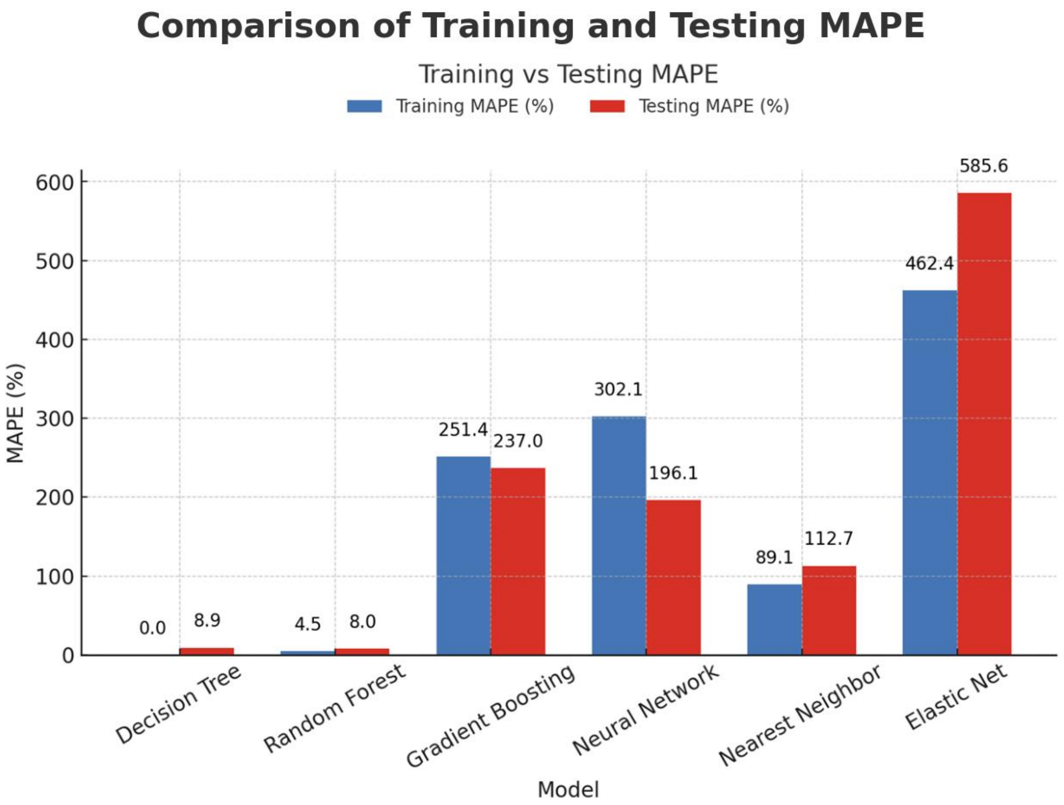
Prediction Quality Measure	*YTR*	*YRFR*	*YBR*	*YNR*	*YNNR*	*YER*	*YAF*	*Prob>F*
MAE	90.482	35.901	49.550	73.017	73.668	77.089	55.946	0.0586
MSE	1267958.500	44997.851	166442.460	85885.735	132196.140	188293.770	62028.157	0.4174
RMSE	1126.037	212.127	407.974	293.063	363.588	433.928	249.055	0.0586
MAPE	14.954	10.030	32.602	47.890	65.082	101.307	24.664	0.0355

Note: This table compares the revenue prediction accuracy of ML models and analyst forecasts. '*YTR*,' '*YRFR*,' '*YBR*,' '*YNR*,' '*YNNR*,' and '*YER*' represent revenue predictions generated by Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net, respectively. '*YAF*' denotes Analyst Forecasts. Prediction accuracy is evaluated using standard error metrics, where lower values indicate more accurate predictions. The Prob > *F* column reports the significance level from an F-test, assessing whether differences in prediction accuracy across models are statistically significant.

doi:10.6342/NTU202501949

Figure 1 Comparison of Training and Testing MAPE Across Machine Learning Models

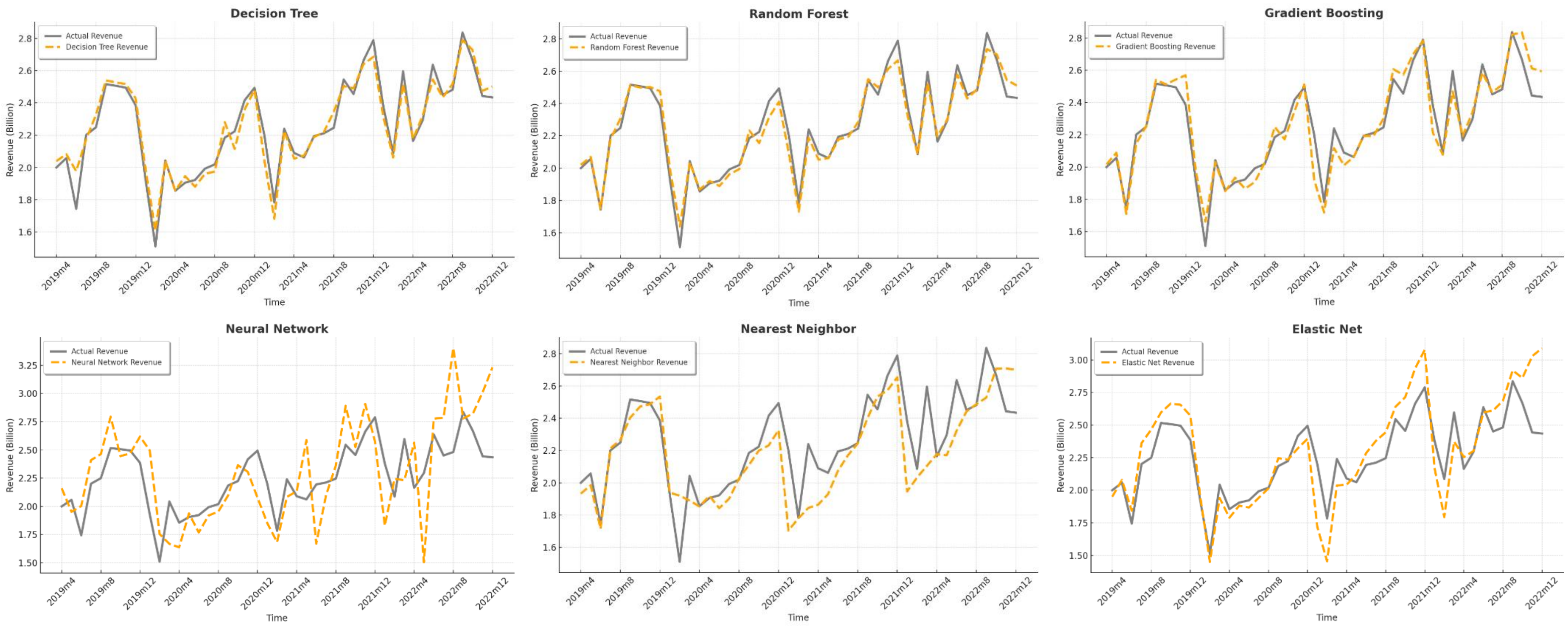
Figure 1 illustrates the Mean Absolute Percentage Error (MAPE) for six ML models: Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net. MAPE serves as an indicator of prediction accuracy, with lower scores corresponding to better outcomes The blue bars represent MAPE on the training dataset, while the red bars correspond to MAPE on the testing dataset. A substantial disparity between training and testing MAPE suggests potential overfitting, where strong in-sample performance does not translate into accurate out-of-sample predictions.



doi:10.6342/NTU202501949

Figure 2 Predicted and Actual Revenues Across Machine Learning Models

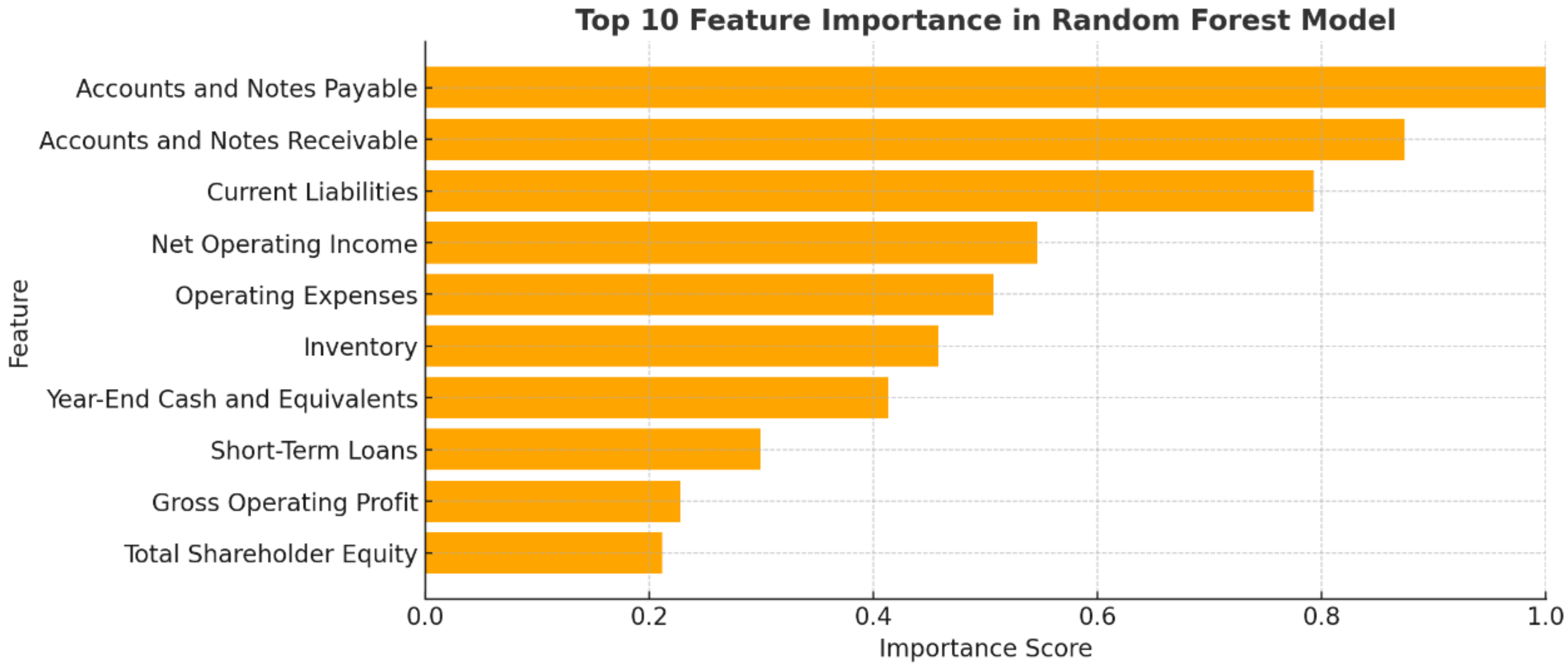
Figure 2 compares predicted and actual revenues from 2019 to 2022 across six ML models: Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net. The orange dashed line represents the predicted revenue generated by each model, while the gray solid line displays the actual observed revenue over time. The x-axis captures the time progression quarterly, and the y-axis reflects revenue in billions of dollars. This figure demonstrates the performance of different models in forecasting revenue patterns over time.



doi:10.6342/NTU202501949

Figure 3 Top 10 Feature Importance in the Random Forest Model

Figure 3 highlights the ten variables with the highest importance scores as determined by the Random Forest model. On the x-axis are the importance scores, capturing the extent to which each feature contributes to the model, while the y-axis labels the features themselves. A higher importance score indicates a more significant influence on the model's output.



doi:10.6342/NTU202501949

the subsequent announcement month, while Panels B and C adjust for risk using the Fama-French five-factor and q-factor models. The results reveal a strong association between ML-based revenue forecasts and abnormal returns.

For unadjusted returns, the 10−1 portfolio strategy based on the Random Forest model generated an abnormal return of 5.539% (*t*-statistic = 4.55); even after risk adjustments, Random Forest retained abnormal returns of 5.513% (*t*-statistic = 4.20). using the Fama-French five-factor model and 5.581% (*t*-statistic = 4.13) with the q-factor model. These consistent findings demonstrate the robustness of Random Forest in generating excess returns, even after accounting for various risk factors.

While Decision Tree, Gradient Boosting, Nearest Neighbor, and Elastic Net do not surpass Random Forest in predictive accuracy, they still yield statistically significant abnormal returns. These results suggest that while their point forecasts may be less precise, their ability to rank stocks based on revenue changes remains effective. The positive abnormal returns indicate that even models with relatively higher forecasting errors can still contribute to profitable investment strategies, provided they capture fundamental revenue trends.

By contrast, Neural Network fails to generate meaningful investment returns, with unadjusted, Fama-French five-factor, and q-factor adjusted returns of 0.429% (*t*-statistic = 0.45), 0.670% (*t*-statistic = 0.76), and 0.739% (*t*-statistic = 0.78), respectively.

doi:10.6342/NTU202501949

This underperformance may stem from overfitting, high model complexity, or limited training data, which restrict its ability to generalize revenue predictions into tradable signals. These findings reveal the varying effectiveness of ML models in translating revenue forecasts into stock market performance.

To evaluate the robustness of these investment strategies across firms of different sizes, we conduct a subsample analysis (Panel D) and examine equally weighted portfolios (Panel E). The results indicate that all models exhibit consistent performance across large-cap and small-cap stocks and in equally weighted portfolios. This consistency confirms the effectiveness of ML models in forecasting revenue changes and highlights their applicability across different market segments.

To assess the investment implications of financial forecasts, we also construct portfolios based on both analyst forecasts and ML-predicted annual revenues. The results show that *CARs* are statistically insignificant, regardless of the forecast source. This outcome is attributable to two factors. First, analyst forecasts are available only on an annual basis, resulting in a limited sample size that constrains the ability to generate statistically significant portfolio returns. Second, the availability of monthly revenue disclosures enables the market to continuously update its expectations, diminishing the timeliness of annual forecasts. To further evaluate predictive performance, we compare ML models with alternative forecasting approaches, including LLMs and ARIMA

doi:10.6342/NTU202501949

models, as discussed in Sections 5.2 and 5.3.

The results in Table 4 indicate that ML models have strong potential in predicting revenue fluctuations and constructing profitable investment portfolios. Among the evaluated models, Random Forest consistently generates the highest abnormal returns with strong statistical significance. Its superior predictive performance and investment profitability are likely attributable to ensemble learning, which aggregates the outputs of multiple decision trees to improve predictive performance, bootstrap sampling, which mitigates overfitting, and aggregation, which improves prediction stability by averaging multiple decision trees, ensuring robustness against the influence of individual variables.

Figure 4 presents the annual fluctuations in abnormal returns generated by various ML models from 2019 to 2022, evaluating their predictive robustness in dynamic market conditions. Mclean and Pontiff (2016) observe that the effectiveness of predictive signals often weakens as arbitrage activities intensify and market liquidity improves, raising the question of whether ML models can sustain consistent profitability over time.

As shown in Figure 4, most ML models consistently generated positive abnormal returns, demonstrating their resilience to evolving market conditions. Although Neural Network and Nearest Neighbor recorded slight negative returns in 2019 and 2022 (-

doi:10.6342/NTU202501949

1.031% and -0.712%, respectively), they remained positive in other years. Despite these fluctuations, the overall trend suggests that ML models effectively predict abnormal returns across different periods. This pattern reflects the adaptability of ML techniques to shifting market dynamics and supports the use of ML-based revenue forecasting as an investment strategy.

Figure 5 presents the monthly average abnormal returns generated by various ML models from 2019 to 2022, highlighting their performance fluctuations under different market conditions. The results indicate that ML models produced positive abnormal returns for most of the period, albeit with varying degrees of volatility. Among them, Random Forest recorded the highest peak return, reaching 10.21% in May, while its lowest return was observed in January at -0.809%. Other models exhibit greater variability, experiencing pronounced fluctuations throughout the year. A more granular analysis reveals that market dynamically, suggesting that making predictions only once per year may fail to capture these variations and adapt to an evolving investment environment.

doi:10.6342/NTU202501949

Table 4 Cumulative Abnormal Returns Based on Revenue Predictions from Machine Learning Models						
Panel A: Cumulative Excess Returns from Previous to Current Revenue Announcement						
Return	Decision Tree	Random Forest	Gradient Boosting	Neural Network	Nearest Neighbor	Elastic Net
1 (low)	-0.040	-0.260	-0.101	1.453	0.860	1.163
2	-0.033	-0.043	0.334	0.673	0.181	0.684
3	0.035	0.309	0.665	2.121	1.111	0.448
4	0.578	0.351	0.612	1.420	1.519	0.718
5	1.125	1.150	1.216	1.651	0.929	0.769
6	1.777	0.611	1.999	1.686	2.047	1.322
7	1.946	2.801	2.755	2.315	1.929	2.319
8	2.679	2.325	3.082	1.422	2.120	2.416
9	2.183	2.909	3.675	2.017	2.068	2.133
10 (high)	4.890	5.279	2.627	1.881	2.742	4.270
10−1	4.930***	5.539***	2.728***	0.429	1.882**	3.107***
t-statistic	(4.75)	(4.55)	(2.69)	(0.45)	(2.57)	(3.47)
Panel B: Risk-Adjusted Portfolio Returns Based on the Fama-French Five-Factor Model						
FF5	Decision Tree	Random Forest	Gradient Boosting	Neural Network	Nearest Neighbor	Elastic Net
1 (low)	-0.060	-0.401	-0.537	2.118	1.428	1.309
2	-0.333	0.255	0.104	0.570	0.494	0.865
3	-0.192	-0.043	0.104	2.168	1.341	0.308
4	0.678	0.282	0.395	0.778	1.289	0.657
5	1.031	1.197	1.345	2.053	0.777	1.035

doi:10.6342/NTU202501949

6	1.542	0.907	2.213	1.792	1.798	1.375
7	2.174	2.642	2.919	2.674	2.191	2.112
8	2.516	2.378	3.019	1.822	2.468	2.180
9	2.335	3.616	4.139	2.098	2.601	2.853
10 (high)	4.589	5.747	2.586	2.062	2.448	4.324
10−1	4.877***	5.513***	2.066*	0.670	1.944**	2.987***
t-statistic	(4.17)	(4.20)	(1.82)	(0.76)	(2.44)	(3.21)
Panel C: Risk-Adjusted Returns Using q-Factor Model						
q-factor	Decision Tree	Random Forest	Gradient Boosting	Neural Network	Nearest Neighbor	Elastic Net
1 (low)	0.024	-0.165	-1.130	2.019	1.247	0.860
2	-0.084	-0.096	0.086	0.617	0.135	0.394
3	-0.457	-0.211	0.619	2.299	1.255	0.142
4	0.957	0.281	0.629	0.725	2.184	0.511
5	1.382	1.140	1.485	2.001	1.168	0.638
6	1.834	0.607	1.610	1.791	2.130	1.530
7	3.181	2.832	2.822	2.013	1.751	2.214
8	2.904	2.926	3.079	1.448	2.015	1.692
9	2.757	3.035	3.846	2.277	3.117	2.335
10 (high)	4.481	5.120	2.644	2.021	2.425	4.399
10−1	4.799***	5.581***	1.723	0.739	1.906**	3.436***
t-statistic	(3.97)	(4.13)	(1.55)	(0.78)	(2.50)	(3.61)
Panel D: Subsample Analysis for Large-Cap and Small-Cap Stocks						

doi:10.6342/NTU202501949

Large-cap		Decision Tree	Random Forest	Gradient Boosting	Neural Network	Nearest Neighbor	Elastic Net
Return	10−1	4.862***	5.117***	3.270***	0.410	2.146***	3.041***
	t-statistic	(4.52)	(4.45)	(3.11)	(0.48)	(2.78)	(3.38)
FF5	10−1	5.065***	5.142***	3.398***	-0.159	1.514*	2.414***
	t-statistic	(4.58)	(4.11)	(3.15)	(-0.19)	(1.78)	(2.80)
q-factor	10−1	5.521***	4.996***	3.252***	0.376	1.960**	3.026***
	t-statistic	(4.84)	(3.90)	(2.98)	(0.44)	(2.35)	(3.25)
Small-cap		**Decision Tree**	**Random Forest**	**Gradient Boosting**	**Neural Network**	**Nearest Neighbor**	**Elastic Net**
Return	10−1	5.014***	5.228***	3.737***	0.385	0.835**	4.117***
	t-statistic	(10.30)	(10.56)	(7.50)	(0.92)	(2.22)	(7.50)
FF5	10−1	4.971***	5.357***	3.677***	0.382	0.854**	4.289***
	t-statistic	(9.49)	(10.63)	(7.44)	(0.77)	(2.13)	(6.63)
q-factor	10−1	5.106***	5.951***	3.541***	0.172	0.845**	4.186***
	t-statistic	(9.20)	(12.72)	(7.25)	(0.37)	(2.13)	(6.83)

Panel E: Equal-Weighted Portfolio Returns

Equal-weighted		Decision Tree	Random Forest	Gradient Boosting	Neural Network	Nearest Neighbor	Elastic Net
Return	10−1	416.833***	437.011***	263.380***	32.632	263.380***	333.771***
	t-statistic	(9.61)	(9.95)	(7.08)	(0.88)	(7.08)	(9.88)
FF5	10−1	386.209***	453.686***	253.467***	30.262	226.396***	339.388***
	t-statistic	(9.07)	(9.64)	(6.37)	(0.82)	(5.75)	(9.10)
q-factor	10−1	390.489***	443.433***	263.669***	27.319	248.261***	353.508***
	t-statistic	(8.42)	(9.71)	(6.99)	(0.72)	(6.37)	(9.44)

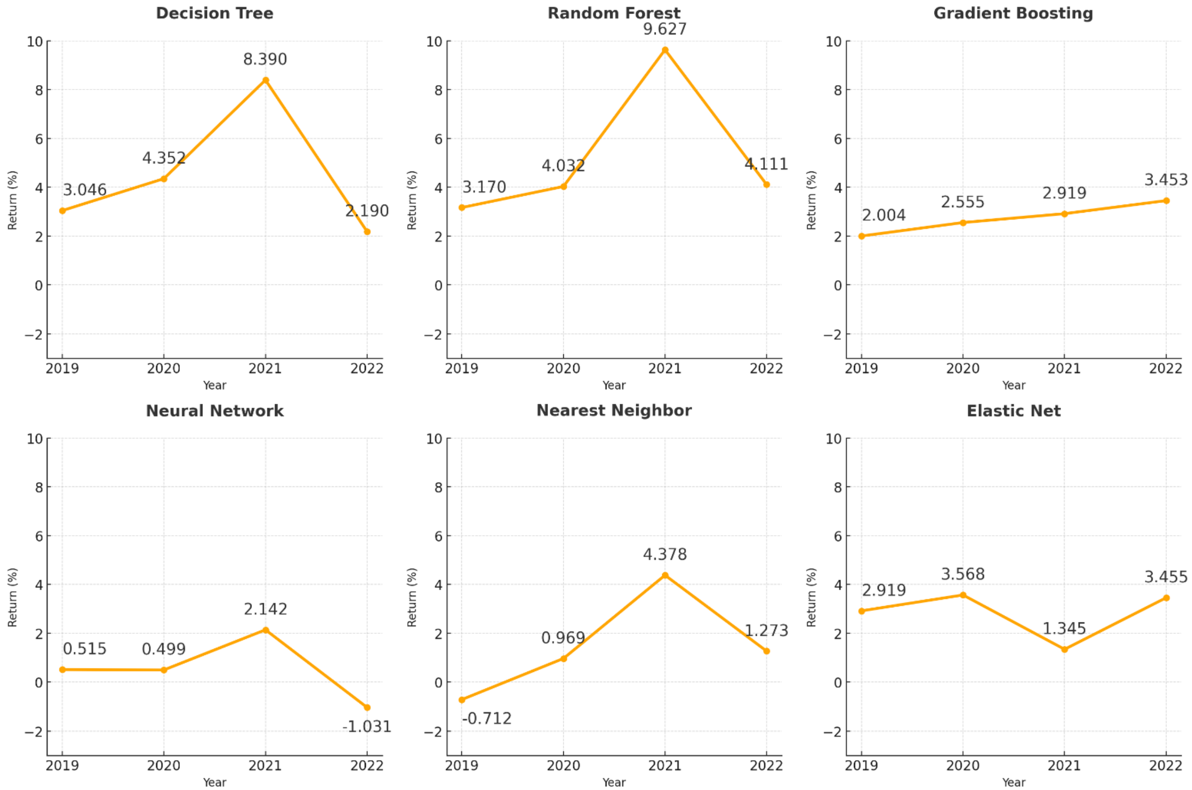
doi:10.6342/NTU202501949

Note: This table presents the cumulative abnormal returns based on revenue predictions from various ML models. Panel A displays cumulative excess returns from the day following the previous revenue announcement to the current month's revenue announcement. Data is sorted into deciles based on predicted revenues, with "10−1" representing the difference between the highest and lowest deciles. Panels B and C report risk-adjusted returns using the Fama-French five-factor and q-factor models. Panel D presents a subsample analysis for large-cap and small-cap stocks, and Panel E shows results for equal-weighted portfolios.

doi:10.6342/NTU202501949

Figure 4 Average Excess Returns for Machine Learning Models Over the Years

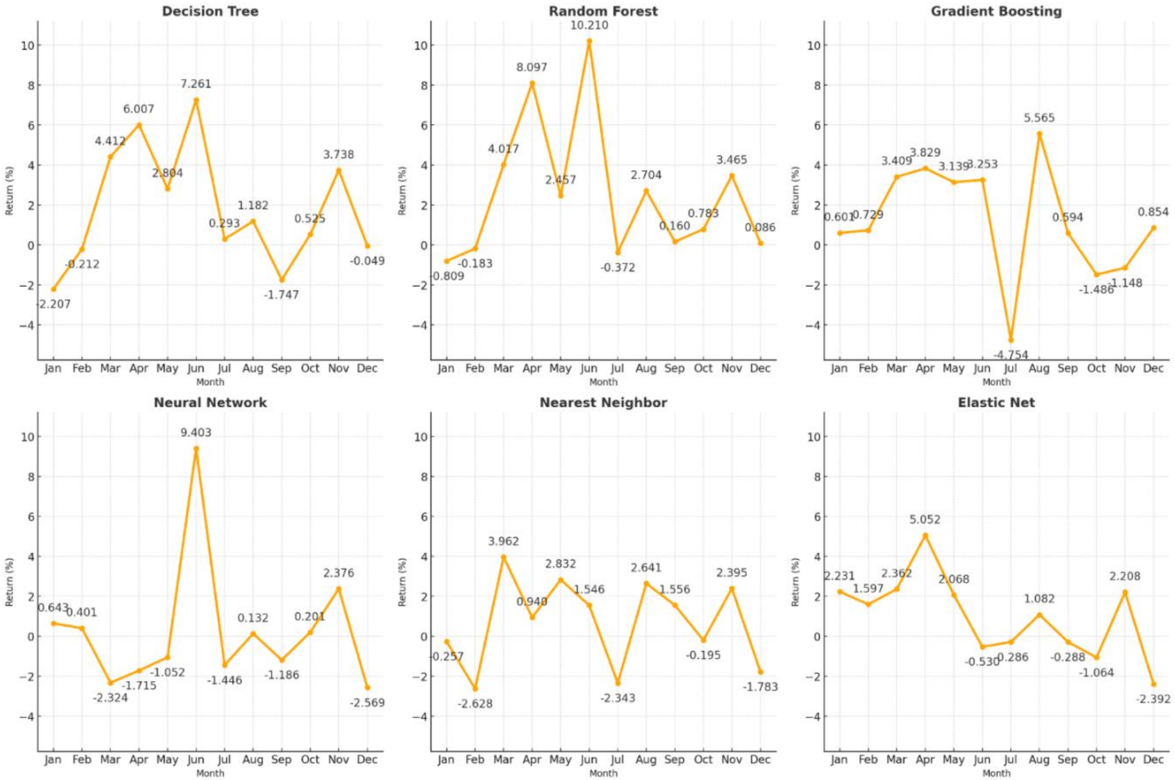
Figure 4 shows the annual fluctuations in excess returns for various ML models—Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net—from 2019 to 2022. The x-axis displays the time in years, whereas the y-axis reflects the corresponding percentages return rates, capturing the performance of each ML model in generating abnormal returns annually.



doi:10.6342/NTU202501949

Figure 5 Average Excess Returns for Machine Learning Models Across Months

Figure 5 displays the monthly variations in excess returns for various ML models—Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net—from 2019 to 2022. The x-axis displays the time in months, while the y-axis reflects the corresponding percentages return rates, capturing the fluctuations in abnormal returns across different models each month.



doi:10.6342/NTU202501949

4.4 Isolating Predictive Performance from Revenue Announcement Effects

To ensure that the investment window purely reflects the predictive ability of ML-based revenue forecasts, we further isolate its impact from any market reactions following the actual revenue disclosure. For instance, the January sales announcement is released on February 10. To prevent potential distortions from post-announcement market responses, the portfolio is held from January 10 (the prior month's revenue announcement date) to February 9 (the day before the January sales announcement). This adjustment eliminates potential bias from including February 10 and 11, as returns on these days may reflect market responses to the actual revenue disclosure.

Table 5 reports *CAR* under this alternative holding period, confirming that the results remain robust. ML-based revenue forecasts generate significant abnormal returns even when strictly excluding actual revenue announcements. This analysis further validates that the observed returns stem from the informational content of revenue predictions rather than market reactions to realized revenues.

Table 5 Excluding the Effect of Revenue Announcements

Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	4.375***	4.039***	4.671***
	(3.78)	(3.29)	(4.24)
Random Forest	4.822***	4.340***	4.885***
	(3.75)	(2.96)	(3.64)
Gradient Boosting	2.347**	2.215**	2.054**
	(2.49)	(2.29)	(2.15)
Neural Network	0.997	1.429	0.516
	(0.95)	(1.24)	(0.80)
Nearest Neighbor	1.988***	1.883***	1.316*

doi:10.6342/NTU202501949

	(3.06)	(2.58)	(1.86)
Elastic Net	2.341***	2.034***	2.655***
	(3.11)	(2.63)	(3.23)

Note: This table reports the cumulative abnormal returns after excluding the effect of revenue announcements, based on portfolios constructed using six ML models. The "10−1" portfolio represents the return spread between the top and bottom deciles of predicted revenue growth.

4.5 Short-Term Abnormal Returns from Machine Learning Revenue Forecasts

Understanding short-window returns is crucial for assessing the immediate market response to revenue announcements. This analysis focuses on a narrow event window, capturing investors' direct reactions to new information while minimizing the influence of confounding factors. We evaluate *CAR* over three short-term event windows—[-1, +1], [-2, +2], and [-3, +3]—to measure the effectiveness of different ML models in predicting short-term market movements.

Table 6 presents the short-term window analysis results, examining market fluctuations before and after revenue announcements. Except for the Nearest Neighbor, most ML models continue to generate significantly positive abnormal returns within short windows, further demonstrating the predictive value of ML-based revenue forecasts in short-horizon market reactions. Among these models, Random Forest consistently produces significantly positive returns across most event windows, demonstrating its ability to capture short-term price movements and generate excess returns.

Table 6 Cumulative Abnormal Returns in Short Windows

Panel A: Excess Return			
Portfolio (10−1)	[-1,+1]	[-2,+2]	[-3,+3]

doi:10.6342/NTU202501949

Decision Tree	0.902*	1.278**	1.551**
	(1.87)	(2.18)	(2.50)
Random Forest	0.864	1.428**	1.632**
	(1.60)	(2.27)	(2.38)
Gradient Boosting	0.652	0.888**	1.001**
	(1.54)	(2.54)	(2.21)
Neural Network	-0.928*	-0.913*	-0.964*
	(-1.82)	(-1.80)	(-1.88)
Nearest Neighbor	-0.059	-0.184	0.243
	(-0.15)	(-0.43)	(0.55)
Elastic Net	0.808*	1.715***	1.981***
	(1.80)	(3.56)	(3.37)
Panel B: Risk-Adjusted Returns Using Fama-French Five-Factor Model			
Portfolio (10−1)	[-1,+1]	[-2,+2]	[-3,+3]
Decision Tree	1.174**	1.103*	1.495**
	(2.51)	(1.73)	(2.09)
Random Forest	1.149**	1.214*	1.776**
	(2.19)	(1.76)	(2.23)
Gradient Boosting	0.721	0.899**	1.349***
	(1.59)	(2.28)	(2.77)
Neural Network	-0.785	-0.749	-0.830
	(-1.56)	(-1.30)	(-1.59)
Nearest Neighbor	0.135	-0.362	0.006
	(0.37)	(-0.74)	(0.01)
Elastic Net	0.865*	1.596***	1.796***
	(1.81)	(3.03)	(2.98)
Panel C: Risk-Adjusted Returns Using q-Factor Model			
Portfolio (10−1)	[-1,+1]	[-2,+2]	[-3,+3]
Decision Tree	0.966**	1.166*	1.736**
	(2.07)	(1.92)	(2.45)
Random Forest	0.925*	1.312**	2.062***
	(1.73)	(2.03)	(2.68)
Gradient Boosting	0.594	0.812**	1.441***
	(1.32)	(2.18)	(3.13)
Neural Network	-0.915*	-0.917*	-0.749
	(-1.75)	(-1.70)	(-1.37)
Nearest Neighbor	0.017	-0.260	0.098
	(0.05)	(-0.57)	(0.21)

doi:10.6342/NTU202501949

Elastic Net	0.758	1.786***	1.919***
	(1.62)	(3.48)	(3.14)

Note: This table presents the cumulative abnormal returns for short-term event windows, including [-1,+1], [-2,+2], and [-3,+3], based on portfolios constructed using six ML models. The "10−1" portfolio represents the return spread between the highest and lowest predicted revenue portfolios. Panel A reports raw excess returns, while Panels B and C report returns adjusted for risk based on the Fama-French five-factor model (FF5) and q-factor models.

4.6 Post-Revenue Announcement Drift and Machine Learning Forecasts

This section examines whether revenue announcements lead to a drift effect in the month following their release, analogous to the post-earnings announcement drift (PEAD) that Ball and Brown (1968), Bernard and Thomas (1989), Foster, Olsen, and Shevlin (1984), Jegadeesh and Titman (1993) identify. We analyze *CAR* from the second day after the announcement to the end of the month ([+2, +EOM]) to assess the predictive effectiveness of various ML models. For example, when the revenue for January is announced, we use the forecasts to predict February revenue and form portfolios accordingly. These portfolios are held from the second trading day after the official February revenue announcement (typically released on February 10) through the end of February.

Table 7 reports *CAR* for each ML model over this period. The results indicate that, except for Random Forest, most ML models generate insignificant abnormal returns post-announcement. This limited drift effect may indicate improvements in the information environment, which allow investors to process and incorporate new information more efficiently, reducing the persistence of abnormal returns. Fink (2021)

doi:10.6342/NTU202501949

finds that PEAD has weakened recently, suggesting a more efficient market response to financial disclosures. Additionally, revenue announcements in Taiwan are typically concise, providing monthly revenue, year-over-year comparisons, and cumulative data, which may further accelerate market reactions and limit post-announcement drift. Among the models examined, Random Forest is the only one that continues to generate statistically significant positive abnormal returns, suggesting that it captures revenue-related signals that persist beyond the initial market reaction.

Table 7 Cumulative Abnormal Returns from Post-Announcement to Month-End

Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	1.615	1.561	1.432
	(1.59)	(1.34)	(1.34)
Random Forest	1.814*	2.238*	1.807
	(1.70)	(1.85)	(1.61)
Gradient Boosting	-0.215	-0.504	-0.217
	(-0.45)	(-0.98)	(-0.44)
Neural Network	-0.219	0.254	0.100
	(-0.34)	(0.38)	(0.16)
Nearest Neighbor	0.418	0.265	0.463
	(0.69)	(0.40)	(0.73)
Elastic Net	-0.057	-0.951	-0.099
	(-0.08)	(-1.34)	(-0.13)

Note: This table presents cumulative abnormal returns from the post-announcement period to the end of the month. The analysis employs six ML models. The "10−1" represents the return difference between the highest and lowest predicted revenue portfolios.

4.7 Machine Learning Predictions in the Technology Sector

Revenue plays a central role in valuing technology firms, given the earnings volatility and uncertainty associated with R&D-intensive businesses (Kothari, Laguerre, and Leone (2002)). Prior research finds that markets react more strongly to revenue

doi:10.6342/NTU202501949

surprises in technology firms than those with lower R&D intensity (Chandra and Ro (2008), Kama (2009)), reinforcing the importance of revenue as a key valuation metric. We classify firms into technology and non-technology industries based on TSE industry definitions to assess industry-specific effects. The technology industry includes semiconductors, computers and peripherals, optoelectronics, communications and internet, electronic components, electronic distribution, and information services, while all other industries fall into the non-technology category.

Table 8 presents differences in ML model performance across industries. Panel A reports that ML models generate higher abnormal returns in the technology industry, whereas Panel B shows lower returns in non-technology firms. ML models exhibit statistically significant abnormal returns in most technology firms, though their predictive performance declines in non-technology industries. These findings support differentiated investment strategies across industries, reflecting the greater importance of revenue in technology firms, where rapid revenue growth often translates into higher stock returns.

Further analysis reveals that Random Forest consistently delivers strong performance across both industries, generating the highest risk-adjusted abnormal returns. In this analysis, Nearest Neighbor models become statistically insignificant in

doi:10.6342/NTU202501949

both industries, possibly due to diminished predictive effectiveness resulting from the reduced sample size.

We find that the Elastic Net model yields significantly stronger revenue forecasting performance for technology firms than for non-technology firms, suggesting that the financial and operational features of tech companies are more effectively utilized by machine learning models. The top five predictors are Last Year's Monthly Revenue, Inventory, Cumulative Revenue, Last Year's Cumulative Revenue, and Net Operating Income. In technology firms, past revenue and inventory levels serve as reliable predictors of future revenue. In rapidly innovating industries, higher inventory typically reflects expectations of strong demand, rather than excess stock.

Table 8 Portfolio Returns Across Technology and Non-Technology

Panel A: Technology Industry			
Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	4.707***	4.545***	5.186***
	(5.94)	(5.59)	(6.23)
Random Forest	4.735***	4.819***	4.599***
	(5.00)	(4.81)	(4.66)
Gradient Boosting	5.385***	5.440***	4.777***
	(4.91)	(4.52)	(4.02)
Neural Network	0.286	0.369	0.900
	(0.33)	(0.36)	(0.92)
Nearest Neighbor	0.869	0.474	0.653
	(1.02)	(0.53)	(0.69)
Elastic Net	5.977***	6.094***	6.167***
	(5.33)	(5.26)	(5.46)
Panel B: Non-Technology Industry			
Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	4.539***	4.601***	4.531***

doi:10.6342/NTU202501949

	(3.26)	(3.14)	(3.09)
Random Forest	4.678***	4.291***	4.498***
	(3.40)	(2.81)	(3.01)
Gradient Boosting	1.929*	1.732	1.429
	(1.73)	(1.43)	(1.27)
Neural Network	-0.199	-0.329	-0.202
	(-0.19)	(-0.34)	(-0.18)
Nearest Neighbor	0.639	0.510	0.442
	(0.78)	(0.64)	(0.51)
Elastic Net	-0.250	-1.450	-0.912
	(-0.20)	(-1.11)	(-0.67)

Note: This table presents the portfolio performance based on ML predictions for technology (Panel A) and non-technology (Panel B) industries.

4.8 Robustness tests

4.8.1 Effect of Training Window Length in Machine Learning Forecasts

To assess the robustness of our ML models' predictive performance, we extend the analysis by incorporating rolling windows of 2-year, 3-year, and 5-year periods to evaluate the stability of ML models across different training sets. This approach allows us to examine whether the length of the training window affects forecasting accuracy and the ability to predict abnormal returns.

Table 9 presents the results, showing that ML models maintain stable predictive performance across all time horizons. Furthermore, as the training period lengthens, most models generate higher abnormal returns, suggesting that ML effectively integrates historical data to enhance forecasting precision. This improvement may stem from two factors. First, a longer training window provides more diverse and representative observations, enabling models to capture structural patterns rather than

doi:10.6342/NTU202501949

overfitting to short-term noise. Second, because financial data are often cyclical and volatile, shorter samples may fail to span different macroeconomic regimes, limiting the model's ability to generalize over time. This observation suggests that longer training windows contribute to greater model stability and statistical significance.

More specifically, the Random Forest model consistently delivers the highest abnormal returns across different training windows, reporting 4.739%, 4.465%, and 6.007% for the 2-year, 3-year, and 5-year periods, respectively. These findings are consistent with its strong performance in the 4-year rolling window used in the primary analysis. Likewise, other ML models exhibit comparable trends, confirming the robustness of the primary analysis and demonstrating that the length of the training window does not materially affect the overall predictive effectiveness of ML models.

Table 9 Abnormal Returns from Portfolios Across Different Rolling Windows

Portfolio	Excess return	2yrs	3yrs	5yrs
Decision Tree	10−1	3.809***	3.890***	5.031***
	t-statistic	(4.52)	(4.11)	(3.59)
Random Forest	10−1	4.739***	4.465***	6.007***
	t-statistic	(5.31)	(4.04)	(4.03)
Gradient Boosting	10−1	2.596***	1.870*	2.465*
	t-statistic	(2.88)	(1.81)	(1.72)
Neural Network	10−1	0.261	0.736	-0.179
	t-statistic	(0.48)	(1.04)	(-0.21)
Nearest Neighbor	10−1	2.065***	1.609**	2.258**
	t-statistic	(3.09)	(2.20)	(2.49)
Elastic Net	10−1	2.383***	2.298**	5.287***
	t-statistic	(3.12)	(2.55)	(4.30)

Note: This table presents the abnormal returns of portfolios based on various rolling windows, utilizing six ML models. Results are reported for 2-year, 3-year, and 5-year rolling windows.

doi:10.6342/NTU202501949

4.8.2 Portfolio Performance and Revenue per Share Analysis

This section examines how revenue per share (RPS) growth influences portfolio performance while adjusting for the effect of outstanding shares. RPS, defined as monthly revenue divided by the number of outstanding shares, provides a revenue-based measure that adjusts for potential dilution. By isolating the impact of changes in outstanding shares, RPS enables a clearer distinction between genuine improvements in operating efficiency and superficial revenue growth driven by equity issuance or asset expansion. This adjustment is particularly important when evaluating firms with varying capital structures or aggressive financing policies.

Table 10 presents portfolio performance based on RPS forecasts. Most ML models generate statistically significant abnormal returns after adjusting for outstanding shares. However, Nearest Neighbors does not produce significant excess returns, implying that distance-based methods may have limited effectiveness in capturing revenue signals when share dilution is considered. Among all models, Random Forest achieves the highest performance, with excess returns of 3.737%, which remained statistically significant after adjusting for risk using the Fama-French five-factor and q-factor models. These findings confirm the robustness of Random Forest as an effective predictive model.

Table 10 Portfolio Performance Based on Revenue per Share Predictions

Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	3.712***	3.579***	3.728***

doi:10.6342/NTU202501949

	(4.89)	(4.36)	(4.30)
Random Forest	3.737***	3.235***	3.734***
	(3.94)	(3.15)	(3.40)
Gradient Boosting	3.027***	3.356***	3.413***
	(3.57)	(3.56)	(3.82)
Neural Network	0.029	0.055	0.059
	(0.04)	(0.07)	(0.08)
Nearest Neighbor	-0.167	-0.562	-0.314
	(-0.24)	(-0.68)	(-0.43)
Elastic Net	1.892***	1.843***	2.163***
	(2.91)	(2.68)	(3.17)

Note: This table presents the portfolio performance based on revenue per share predictions.

4.8.3 Portfolio Performance After Excluding the Construction Industry

Chen, Liu, and Chiao (2022) emphasize the distinct revenue recognition method used in the construction industry, where firms primarily adopt the completed contract method. This approach results in significant revenue fluctuations, introducing potential distortions in financial forecasting models. To validate our findings, we conduct a robustness analysis by excluding the construction sector from our sample, enabling a more accurate evaluation of potential biases and enhancing the reliability of our findings. Removing construction firms ensures that our investment strategy's effectiveness is not influenced by industry-specific accounting treatments that could artificially affect revenue predictions and portfolio performance.

Table 11 reports portfolio performance based on predicted monthly revenue changes after excluding the construction industry. The results show that our investment strategy remains effective following this exclusion, confirming the robustness and

doi:10.6342/NTU202501949

generalizability of our results. Among the models, Random Forest delivers the strongest performance, generating excess returns of 5.809%, with statistical significance maintained even after risk adjustments using the Fama-French five-factor and q-factor models.

Table 11 Portfolio Performance Excluding the Construction Sector

Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	5.329***	5.950***	5.662***
	(4.98)	(5.42)	(4.98)
Random Forest	5.809***	5.446***	5.988***
	(4.55)	(3.99)	(4.58)
Gradient Boosting	3.079***	3.237***	3.655***
	(2.71)	(3.61)	(3.41)
Neural Network	0.604	0.314	0.486
	(0.63)	(0.32)	(0.50)
Nearest Neighbor	1.584**	1.604*	1.516*
	(2.01)	(1.94)	(1.83)
Elastic Net	3.418***	3.124***	3.472***
	(3.76)	(3.20)	(3.69)

Note: This table presents the portfolio performance of machine learning models after excluding the construction industry.

5. Additional Analysis

5.1 Strategy Profitability After Accounting for Transaction Costs

Novy-Marx and Velikov (2016) find that many investment strategies experience a substantial decline in profitability once transaction costs are incorporated, often rendering abnormal returns negligible. To assess the practical feasibility of our strategy, we incorporate transaction costs into the evaluation framework, ensuring that the observed excess returns remain statistically and economically meaningful. This

doi:10.6342/NTU202501949

adjustment enables a more realistic assessment of whether the strategy retains profitability under real-world trading frictions.

We employ a comprehensive transaction cost framework to provide a conservative yet realistic evaluation of the strategy's viability. The cost structure includes a 0.6% securities transaction tax, a 0.57% brokerage fee, and a 0.08% short-selling fee. Additionally, we include a one-month 0.13% funding cost based on the average loan interest rate from the top five banks. We also incorporate a one-month -0.02% interest revenue from securities lending, where a negative value reflects a positive return. These components collectively amount to 1.36% of total transaction costs, establishing a structured basis for evaluating the strategy's economic sustainability.

Figure 6 reports the post-cost abnormal returns across different ML models. After transaction costs, the Random Forest strategy generates an abnormal return of 4.179%[9], followed by Decision Tree: 3.570%, Gradient Boosting: 1.368%, Nearest Neighbors: 0.522%, and Elastic Net: 1.747%. In contrast, the Neural Network model, already statistically insignificant in preliminary tests, yields a negative return of -0.931%.

These results demonstrate the varying degrees of resilience among ML-driven investment strategies when subjected to real-world trading frictions. Despite transaction

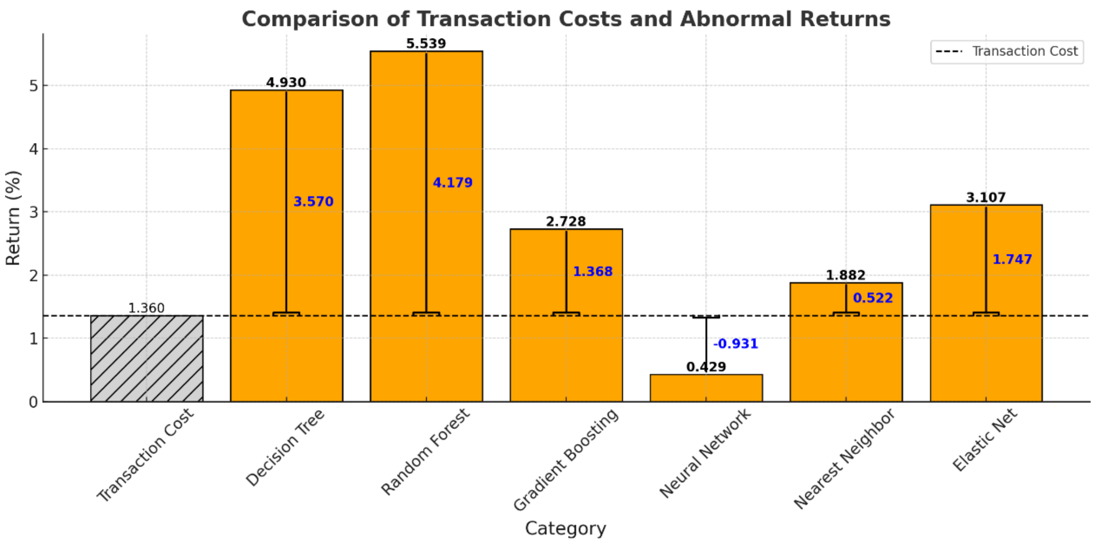
[9]Given a monthly return of 3.570%, the annualized return is approximately 51.29%, calculated as $(1 + 0.0357)^{12} - 1$.

doi:10.6342/NTU202501949

costs, most ML-based portfolios deliver positive abnormal returns, confirming these approaches' robustness and practical relevance. Accounting for transaction costs in investment strategy assessments is essential, as it allows for a more accurate evaluation of a strategy's long-term profitability under realistic trading conditions. Moreover, these findings illustrate the potential of ML-driven models to sustain excess returns even in the presence of market frictions.

Figure 6 Transaction Costs and Excess Returns for Machine Learning Models

Figure 6 compares transaction costs and excess returns across six ML models: Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net. The orange bars represent the excess returns achieved by each strategy, while the gray bar denotes the associated transaction cost.



5.2 Investment Performance of Large Language Model Forecasts

Generative AI, notably LLMs such as GPT-4, has achieved significant progress in text analysis, interpretation, and generation. Recent research suggests that these models can approximate financial analysts' capabilities in numerical analysis and decision-

doi:10.6342/NTU202501949

making (Lopez-Lira and Tang (2024)). Additionally, some studies show that LLMs perform exceptionally well in settings with high analyst disagreement, accurately predicting quarterly earnings Kim, Muhn, and Nikolaev (2025). Building on this foundation, our study extends prior methodologies by employing LLMs to forecast higher-frequency monthly revenues and comparing their predictive performance with ML models.

To refine previous approaches, we introduce several methodological improvements. First, we utilize GPT-4o, an advanced iteration of GPT-4 Turbo. GPT-4o improves response quality and replicates financial analysts' reasoning processes more accurately, enhancing predictive accuracy. Additionally, while prior studies predicted binary directions, confidence levels, and three levels of magnitude, our study categorizes revenue growth rates into ten deciles, allowing for a more detailed and granular analysis.

Our methodology follows a structured process to ensure robust predictions. We anonymize and standardize company financial statements to mitigate biases stemming from the model's prior knowledge. To ensure a fair comparison, both ML and LLM models are trained on identically processed datasets, where all company-specific information is anonymized, and financial variables are standardized. We then apply chain-of-thought (CoT) prompting to guide GPT-4o in identifying financial trends,

doi:10.6342/NTU202501949

computing key ratios, and deriving economic insights. The model subsequently predicts monthly revenue changes, assigning them to ten tiers, where 1 represents the lowest predicted revenue growth and 10 the highest (Wei et al. (2022)).

Table 12 presents portfolio performance based on LLM-generated forecasts. The 10−1 portfolio constructed from LLM-based predictions yields a return of 1.609% (*t*-statistic = 2.18), which remains statistically significant after adjusting for risk factors, with excess returns of 1.416% (*t*-statistic = 1.94) and 1.294% (*t*-statistic = 1.66) under the Fama-French five-factor and q-factor models, respectively.

Compared to ML models (Table 4), investment strategies based on LLMs generate lower returns, exceeding only those derived from Neural Networks. These findings illustrate fundamental differences in predictive approaches. While LLMs can process structured financial data, the returns achieved through these models remain lower than those of most quantitative methods. Among ML approaches, Random Forest demonstrates greater effectiveness in capturing financial patterns and trends, making it more suitable for revenue forecasting.

Table 12 Portfolio Performance Based on Large Language Model Predictions

Portfolio	Return	FF5	q-factor
1 (low)	0.616	0.929	0.918
2	1.488	1.331	1.363
3	1.580	1.541	1.326
4	1.577	1.486	1.612
5	2.005	2.126	1.863
6	1.525	1.613	1.325

doi:10.6342/NTU202501949

7	2.002	1.932	1.816
8	1.544	1.206	1.065
9	1.615	1.558	1.951
10 (high)	2.225	2.197	2.434
10−1	1.609**	1.416*	1.294*
t-statistic	(2.18)	(1.94)	(1.66)

Note: This table presents the portfolio performance based on predictions from a Large Language Model (LLM). Portfolios are sorted into deciles based on LLM-predicted revenue, with Portfolio 1 (low) representing the lowest predicted revenue and Portfolio 10 (high) representing the highest.

5.3 Investment Performance of ARIMA-Based Forecasts

This section evaluates the forecasting performance of the ARIMA model in revenue prediction and compares it with ML approaches. ARIMA has been widely applied in time series forecasting, including stock returns (Dong et al. (2020)), EPS (Bao et al. (1983), Brown (1993), Hopwood and Newbold (1980)), and revenue estimation (Huang et al. (2017), Liu and Sun (2020)). While ARIMA is effective in short-term forecasting by capturing historical patterns, trends, and cyclical fluctuations, its predictive accuracy depends on the stability of these patterns (Ripley (2002), Wang et al. (2018)). Consequently, its performance may deteriorate in environments characterized by structural breaks or regime shifts. Building on prior research, we apply ARIMA to monthly revenue forecasting and assess its predictive accuracy relative to ML models.

We estimate ARIMA models using a rolling window approach, training on 12-, 24-, 36-, and 48-month periods to predict the subsequent month's revenue. This method ensures forecasts incorporate recent data patterns, making it well-suited for short-term

doi:10.6342/NTU202501949

prediction. However, unlike ML models that employ the full dataset, ARIMA relies on a fixed-length historical window, which may constrain its ability to capture long-term trends.

To optimize ARIMA specifications across firms, we employ Stata's xtarimau command, which selects the best-fitting model based on the Hyndman and Khandakar (2008) algorithm. The selection process evaluates candidate models using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to balance predictive accuracy and model complexity.

Table 13 reports portfolio performance based on ARIMA-generated revenue forecasts. The $10-1$ strategy produces statistically significant abnormal returns. However, ARIMA generates lower returns than most ML methods, exceeding only those of the Neural Networks model. Moreover, its *t*-statistics generally fall below 3, suggesting weaker statistical significance and indicating the greater effectiveness of ML methods in predicting revenue changes.

Table 13 Portfolio Performance Based on Autoregressive Integrated Moving Average

Portfolio	Win(12)	Win(24)	Win(36)	Win(48)
Return				
10−1	1.787***	1.552***	1.803***	1.596**
t-statistic	(3.29)	(2.88)	(3.12)	(2.36)
FF5				
10−1	1.636***	1.434***	1.670***	1.498**
t-statistic	(2.77)	(2.55)	(2.75)	(2.06)
q-factor				

doi:10.6342/NTU202501949

10−1	1.564***	1.577***	1.600***	1.550**
t-statistic	(2.64)	(2.78)	(2.70)	(2.10)

Note: This table presents the performance of portfolios based on predictions from the Autoregressive Integrated Moving Average (ARIMA) model using different rolling windows. Win(12), Win(24), Win(36), and Win(48) represent rolling windows of 12, 24, 36, and 48 months, respectively.

5.4 Investment Performance of EPS Forecasts

This section extends the analysis to evaluate the predictive performance of ML models in forecasting EPS. As a critical financial metric directly tied to stock performance, EPS provides an alternative benchmark for assessing the effectiveness of ML-based forecasting models. This analysis examines whether ML models exhibit similar predictive strength across different financial indicators and whether revenue forecasts offer superior investment signals compared to EPS forecasts.

Table 14 reports portfolio returns based on ML-predicted EPS growth. The 10−1 portfolio remains statistically significant across most ML models, confirming that ML-driven EPS forecasts embed predictive value. Random Forest, the best-performing model, delivers an annualized return of 39.328%, compared to 66.468% for revenue forecasts[10]. Other ML models, including Decision Trees and Gradient Boosting, generate statistically significant abnormal returns based on EPS predictions, albeit at lower levels than revenue forecasts.

[10] The annualized return is calculated using the compound interest formula $(1+r)^n-1$, where r denotes the periodic return and n is the number of periods per year. For example, a quarterly return of 9.832% yields an annualized return of approximately 45.52%, and a monthly return of 5.539% yields an annualized return of approximately 90.97%.

doi:10.6342/NTU202501949

Despite the effectiveness of ML in EPS forecasting, the investment profitability is substantially lower than that derived from revenue-based predictions. This result suggests that while revenue and EPS forecasts contain predictive signals, revenue growth is a more effective trading signal, likely due to its timeliness and direct impact on investor expectations. In contrast, EPS may be subject to greater accounting discretion and reporting frequency, limiting its ability to generate excess returns.

These findings demonstrate that ML models achieve strong predictive performance in revenue forecasting and exhibit efficacy in EPS prediction, though with diminished return-generating potential. This additional analysis reinforces the robustness of the results and further supports the distinct advantage of monthly revenue forecasts in driving investment performance.

Table 14 Machine Learning Forecasting Performance on EPS

Portfolio	Return (10−1)	FF5 (10−1)	q-factor (10−1)
Decision Tree	9.254***	9.656***	10.666***
	(4.39)	(3.48)	(3.76)
Random Forest	9.832***	10.968***	11.100***
	(4.85)	(5.61)	(5.44)
Gradient Boosting	9.773***	9.375***	10.627***
	(4.80)	(5.29)	(5.21)
Neural Network	-0.282	0.507	-0.148
	(-0.17)	(0.25)	(-0.07)
Nearest Neighbor	5.342**	5.981	4.909*
	(2.43)	(1.51)	(1.85)
Elastic Net	9.523***	10.567***	10.327***
	(3.83)	(3.82)	(3.60)

Note: This table evaluates the forecasting performance of ML models on EPS.

doi:10.6342/NTU202501949

6. Conclusion

This study applies six ML models—Decision Tree, Random Forest, Gradient Boosting, Neural Network, Nearest Neighbor, and Elastic Net—to forecast monthly revenues. Among these models, Random Forest exhibits the highest predictive accuracy. To evaluate the economic significance of the forecasts, we form investment portfolios sorted by predicted revenue growth and test their ability to generate abnormal returns. Empirical evidence shows that all machine learning models—except Neural Networks—deliver statistically significant positive alphas, outperforming both LLM and ARIMA benchmarks. The Random Forest model consistently delivers the most potent performance across multiple robustness tests.

The primary advantage of ML lies in its ability to process and analyze large datasets with minimal human intervention, thus reducing potential errors. These models capture complex patterns and nonlinear relationships, particularly in high-dimensional financial data. Given that revenue disclosures occur monthly, ML models can rapidly adapt to evolving trends, enhancing the timeliness of revenue forecasts. The empirical findings support this assertion, demonstrating that ML-based revenue predictions provide more timely and accurate signals than traditional forecasting methods.

This study has two primary limitations. First, the analysis is conducted within the context of the Taiwanese securities market, where revenue disclosure is subject to

doi:10.6342/NTU202501949

unique regulatory requirements. Although regulatory regimes differ across markets, the findings may still offer relevant implications for economies with comparable reporting structures. The limited accessibility of public analyst forecast reports imposes considerable information costs on individual investors, potentially hindering timely access to revenue expectations. More frequent revenue forecasts could improve market transparency, support better-informed investment decisions, and enhance capital market efficiency.

Second, financial information may be influenced by economic cycles, industry shifts, firm-specific characteristics, and regulatory changes, all of which contribute to the non-stationarity of the data. These factors pose challenges to maintaining stable and accurate forecasting models over time. Further investigation may account for evolving political and economic conditions and adjust feature design when necessary.

Future research could extend these ML frameworks to forecast other important financial indicators, including cash flows and firm-level risk measures. Another promising avenue involves integrating unstructured data, such as news sentiment and social media analytics, with structured financial data to enhance predictive performance. However, such integration presents methodological complexities and requires advanced natural language processing (NLP) methodologies to extract meaningful signals.

doi:10.6342/NTU202501949

Addressing these issues would further enhance the practical relevance of ML models in financial forecasting and investment strategies.

doi:10.6342/NTU202501949

References

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In: Springer.

Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting. *Journal of Accounting Research*, *6(2), 159–178*.

Bao, D. H., Lewis, M. T., Lin, W. T., & Manegold, J. G. (1983). Applications of time-series analysis in accounting: A review. *Journal of Forecasting*, *2*(4), 405-423.

Barker, R., Penman, S., Linsmeier, T. J., & Cooper, S. (2020). Moving the Conceptual Framework Forward: Accounting for Uncertainty. *Contemporary Accounting Research*, 37(1), 322–357.

Barth, M. E., Li, K., & McClure, C. G. (2023). Evolution in value relevance of accounting information. *The Accounting Review*, *98*(1), 1–28.

Bernard, V. L., & Thomas, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research*, *27*, 1-36.

Bilinski, P., & Eames, M. (2019). Analyst revenue forecast reporting and the quality of revenues and expenses. *Journal of Business Finance & Accounting*, *46*(1-2), 136-158.

Bolotov, I. (2022). XTARIMAU: Stata module to find the best [S] ARIMA [X] models in heterogeneous panels with the help of arimaauto.

Bradshaw, M. T. (2011). Analysts' forecasts: what do we know after decades of work? *Available at SSRN 1880339*.

Breiman, L. (2001). Random forests. *Machine Learning*, *pp. 45*, 5–32.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). Cart. *Classification and Regression Trees*.

Brown, A. B., Lin, G., & Zhou, A. (2022). Analysts' forecast optimism: The effects of managers' incentives on analysts' forecasts. *Journal of Behavioral and Experimental Finance*, *p. 35*, 100708.

Brown, L. D. (1993). Earnings forecasting research: its implications for capital markets research. *International Journal of Forecasting*, *9*(3), 295–320.

Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2015). Inside the “black box” of sell-side financial analysts. *Journal of Accounting Research*, *53*(1), 1–47.

Butler, M., Kraft, A., & Weiss, I. S. (2007). The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports. *Journal of Accounting and Economics*, *43*(2-3), 181-217.

Cao, K., & You, H. (2020). Fundamental analysis via machine learning. *HKUST Business School Research Paper* (2020-009).

Cerulli, G. (2022). Machine learning using stata/python. *The Stata Journal*, *22*(4), 772–810.

doi:10.6342/NTU202501949

Chandra, U., & Ro, B. T. (2008). The role of revenue in firm valuation. *Accounting Horizons*, *22*(2), 199-222.

Chen, H.-Y., Chen, S.-S., Hsin, C.-W., & Lee, C.-F. (2014). Does revenue momentum drive or ride earnings or price momentum? *Journal of Banking & Finance*, *38*, 166-185.

Chen, J.-C., Liu, Y.-R., & Chiao, C. (2022). Unexpected Monthly Revenue in the Taiwan Stock Market: A Comprehensive Analysis. *Journal of Management & Business Research*, *39*(4), 503-545.

Chen, T.-C., & Wu, Y. J. (2020). The influence of R&D intensity on financial performance: The mediating role of human capital in the semiconductor industry in Taiwan. *Sustainability*, *12*(12), 5128.

Chen, W.-D., & Yu, C.-T. (2022). The information value of interim accounting disclosures: evidence from mandatory monthly revenue reports. *Review of Quantitative Finance and Accounting*, *58*(1), 245-295.

Chen, X., Cho, Y. H., Dou, Y., & Lev, B. (2022). Predicting future earnings changes using machine learning and detailed financial data. *Journal of Accounting Research*, *60*(2), 467–515.

Chordia, T., Subrahmanyam, A., & Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, *58*(1), 41-58.

Chung, I. H., Williams, D. W., & Do, M. R. (2022). For better or worse? revenue forecasting with machine learning approaches. *Public Performance & Management Review*, *45*(5), 1133-1154.

Core, J. E., Guay, W. R., & Van Buskirk, A. (2003). Market valuations in the new economy: An investigation of what has changed. *Journal of Accounting and Economics*, *34*(1-3), 43-67.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21-27.

Cowen, A., Groysberg, B., & Healy, P. (2006). Which types of analyst firms are more optimistic? *Journal of Accounting and Economics*, *41*(1-2), pp. 119–146.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, *2*(4), 303–314.

Dong, H., Guo, X., Reichgelt, H., & Hu, R. (2020). Predictive power of ARIMA models in forecasting equity returns: a sliding window method. *Journal of Asset Management*, *21*(6), 549-566.

Ertimur, Y., Livnat, J., & Martikainen, M. (2003). Differential market reactions to revenue and expense surprises. *Review of Accounting Studies*, *8*, 185-211.

Ertimur, Y., Mayew, W. J., & Stubben, S. R. (2011). Analyst reputation and the issuance

doi:10.6342/NTU202501949

of disaggregated earnings forecasts to I/B/E/S. *Review of Accounting Studies*, *pp. 16*, 29–58.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fink, J. (2021). A review of the post-earnings-announcement drift. *Journal of Behavioral and Experimental Finance*, *p. 29*, 100446.

Fix, E. (1985). *Discriminatory Analysis: Nonparametric Discrimination, Consistency Properties* (Vol. 1). USAF school of Aviation Medicine.

Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns. *Accounting Review*, 574-603.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232.

Gilliam, T. A. (2014). *Revenue Management: The Use of Order Backlog to Meet Revenue Reporting Targets* The Florida State University.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, *40*(1-3), pp. 3–73.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. *The Review of Financial Studies*, *29*(1), 5–68.

Hopwood, W., & Newbold, P. (1980). Time series analysis in accounting: A survey and analysis of recent issues. *Journal of Time Series Analysis*, *1*(2), 135–144.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366.

Hou, K., Van Dijk, M. A., & Zhang, Y. (2012). The implied cost of capital: A new approach. *Journal of Accounting and Economics*, *53*(3), 504–526.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, *28*(3), 650-705.

Houston, J. F., Lev, B., & Tucker, J. W. (2010). To guide or not to guide? Causes and consequences of stopping quarterly earnings guidance. *Contemporary Accounting Research*, *27*(1), 143–185.

Huang, H.-C., Hwang, S.-Y., Chang, S., & Kang, Y. (2017). Forecasting company revenue trend using financial news.

Huang, T.-C., & Hairston, S. (2023). Analyst revenue forecasts and firm revenue misstatements. *European Accounting Review*, *32*(2), 379–414.

Hung, W., Lu, C.-C., & Yang, J. J. (2023). Monthly Revenue Growth and Cross-Sectional Stock Returns.

doi:10.6342/NTU202501949

Hunt, J. O., Myers, J. N., & Myers, L. A. (2022). Improving earnings predictions and abnormal returns with machine learning. *Accounting Horizons*, *36*(1), 131–149.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, *27*, 1-22.

Jegadeesh, N., & Livnat, J. (2006). Revenue surprises and stock returns. *Journal of Accounting and Economics*, *41*(1-2), 147-171.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91.

Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised machine learning: a brief primer. *Behavior Therapy*, *51*(5), 675–687.

Kajüter, P., Lessenich, A., Nienhaus, M., & van Gemmern, F. (2022). Consequences of interim reporting: A literature review and future research directions. *European Accounting Review*, *31*(1), 209-239.

Kama, I. (2009). On the market reaction to revenue and earnings surprises. *Journal of Business Finance & Accounting*, *36*(1-2), 31-50.

Keung, E. C. (2010). Do supplementary sales forecasts increase the credibility of financial analysts' earnings forecasts? *The Accounting Review*, *85*(6), 2047-2074.

Kim, A., Muhn, M., & Nikolaev, V. (2024). Financial statement analysis with large language models. *arXiv Preprint arXiv:2407.17866*.

Kothari, S., Laguerre, T. E., & Leone, A. J. (2002). Capitalization versus expensing: Evidence on the uncertainty of future earnings from capital expenditures versus R&D outlays. *Review of Accounting Studies*, *7*(4), 355–382.

Ku, K. P. (2011). Earnings and Sales Momentum (Chinese). *Journal of Management (Chinese)*, *28*(6), 521–544.

Kureljusic, M., & Karger, E. (2023). Forecasting in financial accounting with artificial intelligence–A systematic literature review and future research agenda. *Journal of Applied Accounting Research*(ahead-of-print).

Kureljusic, M., & Reisch, L. (2022). Revenue forecasting for European capital market-oriented firms: A comparative prediction study between financial analysts and machine learning models. *Corporate Ownership & Control*, *19*(2), 159-178.

Lai, Chi-Cheng, Lin, Wen-Yi, and Liu, Y. Angela. 2011. The Conflicts of Interest on Security Analysts' Earnings Forecasts in Taiwan, *NTU Management Review* 22, 357–389.

Lev, B. (2018). The deteriorating usefulness of financial report information and how to reverse it. *Accounting and Business Research*, *48*(5), 465–493.

Lewis, C. D. (1982). Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. In: Butterworth Scientific.

Li, K. K., & Mohanram, P. (2014). Evaluating cross-sectional forecasting models for

doi:10.6342/NTU202501949

implied cost of capital. *Review of Accounting Studies*, *19*, 1152-1185.

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321-332.

Lim, T. (2001). Rationality and analysts' forecast bias. *The Journal of Finance*, *56*(1), 369–385.

Liu, L., & Sun, D. (2020). Research on financial revenue forecast based on optimal ARIMA model. *Advances in Applied Mathematics*, *9*(3), 414-420.

Lopez-Lira, A., & Tang, Y. (2023). Can chatgpt forecast stock price movements? return predictability and large language models. *arXiv preprint arXiv:2304.07619*.

Lorenz, T., & Homburg, C. (2018). Determinants of analysts' revenue forecast accuracy. *Review of Quantitative Finance and Accounting*, *51*, 389-431.

Martineau, C. (2019). The evolution of financial market efficiency: Evidence from earnings announcements, SSRN Scholarly Paper ID 3111607. *Social Science Research Network, Rochester, NY*.

Matsumoto, D. A. (2002). Management's incentives to avoid negative earnings surprises. *The Accounting Review*, *77*(3), 483–514.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *pp. 5*, 115–133.

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, *71*(1), 5–32.

Mensah, Y. M., & Werner, R. H. (2008). The capital market implications of the frequency of interim financial reporting: an international analysis. *Review of Quantitative Finance and Accounting*, *31*, 71-104.

Novy-Marx, R., & Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *The Review of Financial Studies*, *29*(1), 104-147.

Penman, S. H. (2013). *Financial Statement Analysis and Security Valuation*. McGraw-hill.

Ramnath, S., Rock, S., & Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting*, *24*(1), 34-75.

Rees, L. L., & Sivaramakrishnan, S. (2004). The effect of meeting or beating revenue forecasts on the association between quarterly returns and earnings forecast errors. *Available at SSRN 655421*.

Ripley, B. D. (2002). Time series in R 1.5. 0. *The Newsletter of the R Project Volume*, *2*, 2.

Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, *5*, 197-227.

doi:10.6342/NTU202501949

Smith, Kevin C., 2025, Beyond the event window: Earnings horizon and the informativeness of earnings announcements, *The Accounting Review* 100(2), 351–382.

Schipper, K. (1991). Analysts' forecasts. *Accounting Horizons*, *5*(4), 105–121.

Soliman, M. T. (2008). The use of DuPont analysis by market participants. *The Accounting Review*, *83*(3), 823-853.

Srivastava, A. (2014). Why have measures of earnings quality changed over time? *Journal of Accounting and Economics*, *57*(2-3), 196-217.

Taylor, S., & Tong, A. (2023). How important are semi-annual earnings announcements? An information event perspective. *Accounting & Finance*, *63*(3), 3585–3607.

Tsao, S.-M., Lu, H.-T., & Keung, E. C. (2018). Interim reporting frequency and the mispricing of accruals. *Accounting Horizons*, *32*(3), 29-47.

Veenman, D., & Verwijmeren, P. (2018). Do investors fully unravel persistent pessimism in analysts' earnings forecasts? *The Accounting Review*, *93*(3), 349-377.

Vineeth, V. S., Kusetogullari, H., & Boone, A. (2020). Forecasting sales of truck components: a machine learning approach. 2020 IEEE 10th International Conference on Intelligent Systems (IS),

Wang, C.-H., Lu, Y.-H., Huang, C.-W., & Lee, J.-Y. (2013). R&D, productivity, and market value: An empirical study from high-technology firms. *Omega*, *41*(1), 143-155.

Wang, Q., Li, S., Li, R., & Ma, M. (2018). Forecasting US shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model. *Energy*, *160*, 378-387.

Wang, Z.-M., & Lien, D. (2022). Is maximum daily return a lottery? Evidence from monthly revenue announcements. *Review of Quantitative Finance and Accounting*, *59*(2), 545-600.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824-24837.

Yang, C. H., & Chen, J. R. (2003). Innovation and Market Value in Newly-Industrialized Countries: The Case of Taiwanese Electronics Firms. *Asian Economic Journal*, *17*(2), 205-220.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301-320.

doi:10.6342/NTU202501949

Appendix A. Variables Definitions

Variable	Definition
ΔRLY	Year-over-year growth rate of predicted revenue.
$\Delta RPSLY$	Year-over-year growth rate of predicted revenue per share.
BR	Predicted revenue using Gradient Boosting.
CAR	Cumulative abnormal returns are calculated from the day after the previous revenue announcement to the current month's revenue announcement.
EPS	Earnings per share, calculated as net income after taxes minus preferred dividends, divided by the weighted average number of common shares outstanding.
ER	Predicted revenue using Elastic Net.
$\ln ME$	Natural logarithm of market equity, defined as ln(shares outstanding × unadjusted closing price).
NNR	Predicted revenue using Nearest Neighbors.
NR	Predicted revenue using Neural Networks.
r_{it}	Daily stock returns.
RFR	Predicted revenue using Random Forest.
RLY	Revenue from the previous fiscal year.
TR	Predicted revenue using Decision Tree.
YAF	Annual revenue forecast by brokerage analysts (in billions).
YBR	Annual predicted revenue using Gradient Boosting (in billions).
YER	Annual predicted revenue using Elastic Net (in billions).
$YNNR$	Annual predicted revenue using Nearest Neighbors (in billions).
YNR	Annual predicted revenue using Neural Networks (in billions).
YR	Annual actual revenue (in billions).
$YRFR$	Annual predicted revenue using Random Forest (in billions).
YTR	Annual predicted revenue using Decision Tree (in billions).

doi:10.6342/NTU202501949

Appendix B. Feature Variables and Economic Significance

This study selects 60 variables for ML model training to capture a company's financial condition, operational performance, capital structure, and market indicators. Each variable's economic significance is summarized below:

No.	Variable	Category	Economic Significance
1	Cumulative Revenue	Monthly Revenue	Tracks year-to-date progress toward annual goals.
2	Month-over-Month Revenue	Monthly Revenue	Reflects seasonal changes and trends.
3	Last Year's Cumulative Revenue	Monthly Revenue	Serves as a comparative benchmark.
4	Last Year's Monthly Revenue	Monthly Revenue	Benchmarks current performance.
5	Monthly Revenue Growth Rate	Monthly Revenue	Key short-term revenue forecasting indicator.
6	Revenue Growth Rate	Monthly Revenue	Indicator of growth potential, directly influencing revenue forecasts.
7	After-Tax Net Profit Growth Rate	Income Statement	Vital for assessing distributable profits to shareholders.
8	EBIT (Earnings Before Interest and Taxes)	Income Statement	Assesses core profitability and debt repayment capacity.
9	Gross Operating Profit	Income Statement	Direct impact on the company's profitability.
10	Net Operating Income	Income Statement	Core measure of profitability after cost deductions.
11	Non-Operating Income and Expenses	Income Statement	Offers insights into non-core profitability.
12	Operating Expenses	Income Statement	Direct impact on profitability.
13	Operating Gross Profit Growth Rate	Income Statement	Reflects gross profit potential, impacting profitability.
14	Operating Profit	Income Statement	Key indicator of operational efficiency.
15	Operating Profit Growth Rate	Income Statement	Indicates operational efficiency and profitability.
16	Operating Profit Variability	Income Statement	Essential for profitability stability and risk assessment.
17	Ordinary Net Profit Growth Rate	Income Statement	Essential for long-term operational stability predictions.

doi:10.6342/NTU202501949

18	Pre-Tax Profit	Income Statement	Reflects overall profitability before taxes.
19	Pre-Tax Profit Growth Rate	Income Statement	Affects overall financial health prediction.
20	Recurring Net Profit Growth Rate	Income Statement	Reflects the company's regular profitability excluding one-off items.
21	Accounts Payable and Notes	Balance Sheet	Influences cash turnover and short-term debt.
22	Accounts Receivable and Notes	Balance Sheet	Affects cash flow and turnover.
23	Cash and Cash Equivalents	Balance Sheet	Reflects liquidity and short-term debt repayment ability.
24	Current Liabilities	Balance Sheet	Reflects short-term financial pressure.
25	Depreciable FA Growth Rate	Balance Sheet	Influences capital expenditure strategies.
26	Goodwill and Intangible Assets	Balance Sheet	Impacts brand value and market competitiveness.
27	Inventory	Balance Sheet	Influences capital utilization and sales potential.
28	Net Worth Growth Rate	Balance Sheet	Measures financial health from an investor perspective.
29	Property, Plant, and Equipment (PPE)	Balance Sheet	Critical for assessing production capabilities.
30	Short-Term Loans	Balance Sheet	Affects short-term liquidity and repayment ability.
31	Total Asset Growth Rate	Balance Sheet	Indicates potential for expansion and future revenue implications.
32	Total Assets	Balance Sheet	Indicates overall scale and growth potential.
33	Total Liabilities	Balance Sheet	Key to assessing financial stability and risk.
34	Total Shareholder Equity	Balance Sheet	Essential measure of financial stability.
35	Year-End Cash and Equivalents	Balance Sheet	Indicates year-end liquidity and debt repayment ability.
36	Free Cash Flow	Cash Flow Statement	Key for evaluating financial flexibility.
37	After-Tax Net Profit Margin	Financial Ratios	Indicator of ultimate profit efficiency.
38	Cash Flow per Share	Financial Ratios	Indicates cash flow creation ability.
39	Comprehensive Income per Share	Financial Ratios	Reflects total per-share earnings
40	Comprehensive ROA	Financial Ratios	Assesses overall asset efficiency and profitability.

doi:10.6342/NTU202501949

41	Comprehensive ROE	Financial Ratios	Indicates total return, covering all aspects.
42	Dividend Yield	Financial Ratios	Assesses shareholder value through dividend return.
43	Earnings per Share (EPS)	Financial Ratios	Reflects per-share earnings and value.
44	Financial Leverage	Financial Ratios	Reflects financial risk and capital structure.
45	Fixed Asset Turnover Ratio	Financial Ratios	Indicates revenue generated per unit of fixed asset.
46	Gross Profit Margin	Financial Ratios	Reflects profit potential and pricing ability.
47	Interest Coverage Ratio	Financial Ratios	Evaluates ability to cover interest expenses.
48	Monthly Revenue per Share	Financial Ratios	Guides shareholder return assessments.
49	Net Operating Cycle Days	Financial Ratios	Essential for cash flow management.
50	Operating Income per Share	Financial Ratios	Indicator of stock shares and revenue relationship.
51	Operating Leverage	Financial Ratios	Indicates profit fluctuation in response to revenue changes.
52	Operating Profit per Share	Financial Ratios	Shows profitability of operating activities.
53	Pre-Tax Profit Margin	Financial Ratios	Key profitability measure.
54	Pre-Tax Profit per Share	Financial Ratios	Helpful in assessing shareholder returns.
55	Quarter-End Common Stock Market Value	Financial Ratios	Reflects quarter-end market performance.
56	ROA (Return on Assets)	Financial Ratios	Measures asset utilization efficiency.
57	ROE (Return on Equity, Post-Tax)	Financial Ratios	Evaluates shareholder returns and value creation.
58	Revenue Variability	Financial Ratios	Indicates market demand and operational stability.
59	Tobin's Q	Financial Ratios	Indicates growth potential via market value vs. replacement cost.
60	Total Asset Return Growth Rate	Financial Ratios	Reflects asset management efficiency over time.

doi:10.6342/NTU202501949

Appendix C. Intuitive Explanation of Machine Learning Models

To facilitate understanding for readers less familiar with machine learning techniques, this appendix provides an intuitive summary of the six models applied in this study. While the formal definitions and formulas are presented in the main text, this section explains each method using simplified language and analogies.

Decision Tree

A Decision Tree asks a series of yes/no questions to split the data and make predictions. It creates branches based on which variable best separates the data at each step. The model is easy to understand but may overfit the training data.

Random Forest

A Random Forest builds many Decision Trees using random subsets of data and features, then averages their predictions. Like asking multiple experts and taking the average answer, it helps reduce overfitting and improve accuracy.

Gradient Boosting

Gradient Boosting builds trees one at a time. Each new tree focuses on fixing the mistakes made by the previous one, gradually improving the overall prediction. This step-by-step refinement can lead to highly accurate results.

Neural Networks

Neural Networks mimic how the human brain processes information. They use multiple layers of "neurons" to transform input data and learn complex patterns. These models are powerful but require a lot of data and computing power.

Nearest Neighbors

This model predicts outcomes based on similarity. If a company is similar to five others, its future performance is predicted by averaging those five. It works well for local patterns but is less efficient when data has many variables.

Elastic Net

Elastic Net combines two regularization techniques—Lasso (which removes unimportant variables) and Ridge (which keeps coefficients small)—to improve prediction when many variables are correlated. It’s useful for selecting key predictors in financial data.

doi:10.6342/NTU202501949