Article

Exploring Low-Risk Anomalies: A Dynamic CAPM Utilizing a Machine Learning Approach

Jiawei Wang [1],* and Zhen Chen [2]

1   School of Finance, Shanghai University of Finance and Economics, Shanghai 200433, China
2   School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University,
    Shanghai 200240, China
*   Correspondence: jiaweiwang@163.sufe.edu.cn

Abstract: Low-risk pricing anomalies, characterized by lower returns in higher-risk stocks, are prevalent in equity markets and challenge traditional asset pricing theory. Previous studies primarily relied on linear regression methods, which analyze a limited number of factors and overlook the advantages of machine learning in handling high-dimensional data. This study aims to address these anomalies in the Chinese market by employing machine learning techniques to measure systematic risk. A large dataset consisting of 770 variables, encompassing macroeconomic, micro-firm, and cross-effect factors, was constructed to develop a machine learning-based dynamic capital asset pricing model. Additionally, we investigated the differences in factors influencing time-varying beta between state-owned enterprises (SOEs) and non-SOEs, providing economic explanations for the black-box issues. Our findings demonstrated the effectiveness of random forest and neural networks, with the four-layer neural network performing best and leading to a substantial rise in the excess return of the long–short portfolio, up to 0.36%. Notably, liquidity indicators emerged as the primary drivers influencing beta, followed by momentum. Moreover, our analysis revealed a shift in variable importance during the transition from SOEs to non-SOEs, as liquidity and momentum gradually replaced fundamentals and valuation as key determinants. This research contributes to both theoretical and practical domains by bridging the research gap in incorporating machine learning methods into asset pricing research.

Keywords: asset pricing; beta estimation; machine learning; stock market

MSC: 62P20; 68T09

1. Introduction

Systematic risk measurement and the trade-off between risk and return are central topics in modern asset pricing research [1,2]. The capital asset pricing model (CAPM), a well-established equilibrium asset pricing theory, is commonly used to assess risk exposure in equity markets [3]. According to CAPM, stocks with high risk should yield higher returns, while stocks with low beta are expected to deliver lower returns. However, recent empirical evidence has cast doubt on this relationship, indicating the presence of "low-risk anomalies" [4–6]. These anomalies refer to the phenomenon where low-beta stocks outperform high-beta stocks in terms of returns, challenging traditional notions of risk and return.

To address the low-risk pricing anomalies, beta measurement within the CAPM requires attention in two aspects. Firstly, beta cannot be directly observed and necessitates precise estimation techniques. Secondly, beta dynamically changes over time. In response to the second issue, existing research can be classified into two main streams: static CAPM and dynamic CAPM. The static CAPM derives an unchanging beta through linear regression of stock returns. On the other hand, the dynamic CAPM recognizes the instability of beta and

utilizes conditional information to optimize beta measurement. For example, Boguth et al. [7] proposed that time-varying betas with economic information improved insights into systematic risk and possessed stronger explanatory power for stock returns. However, most existing studies rely on simple linear regressions and consider only specific macroeconomic variables with limited corporate financial factors. As a result, they neglect the potential impact of a large number of macroeconomic and micro-firm factors on the variation in beta over time.

The Chinese stock market, with a total value of USD 12.5 trillion (RMB 88 trillion) by the end of 2022, has solidified its position as the second-largest market globally. Apart from its sheer size, unique characteristics make it attractive for both academia and industry, contributing to the exploration of emerging markets [8]. Notably, the Chinese market is predominantly influenced by retail investors. According to the 2022 yearbook of the Shanghai Stock Exchange, there are 299.5 million investors in China, with individual investors accounting for a staggering 99.7%. The speculative behaviors and short-term trading activities of retail investors result in increased trading volume, emphasizing the importance of investigating how retail investment behaviors affect risk assessment. Furthermore, the Chinese market exhibits complex market structures, with a large number of state-owned enterprises (SOEs) in infrastructure and livelihood sectors. SOEs often face criticism for inadequate information disclosure and divergence of political objectives from corporate value maximization [9].

The unique structures of the Chinese stock market often render models developed for mature capital markets less effective. Therefore, highly flexible methods are required. Machine learning techniques have demonstrated superior predictive performance compared to traditional methods in various domains, including the prediction of stock returns [10], bond premiums [11], and loan defaults [12]. The high-dimensional nature of machine learning enhances its flexibility, providing better access to the complex generation process of share price. Given that beta exhibits a higher signal-to-noise ratio compared to stock returns, machine learning techniques can incorporate a broader range of predictors and richer functional forms, potentially surpassing traditional methods in beta forecasts as well.

The need for machine learning techniques arises from the limitations of traditional methods in asset pricing research. Traditional models, such as linear regression, typically assume linearity and overlook the interactions among predictors, resulting in incomplete risk assessment and less accurate predictions. Moreover, traditional models may struggle to handle high-dimensional datasets that include numerous macroeconomic and micro-firm factors. In contrast, machine learning techniques excel at capturing nonlinear patterns and efficiently handling high-dimensional data. By applying machine learning techniques, we can enhance the precision and predictive power of risk assessment.

In this study, we aimed to address the persistent risk–return asymmetry observed in the Chinese stock market. To achieve this, we introduced an innovative dynamic CAPM utilizing machine learning approach. Specifically, by using mainstream machine learning algorithms including partial least squares, random forest, and neural networks, we measured the systematic risk in a more flexible and intelligent way. Additionally, given the substantial market capitalization of SOEs in the Chinese market, we further examined the beta anomaly, for both SOEs and non-SOEs separately, and revealed the underlying reasons driving these differences at the factor level.

This study makes the following contributions:

- We uncovered the factors influencing time-varying beta, including macroeconomic and micro-firm features, as well as their cross-effects. By constructing a comprehensive database comprising 70 micro-firm characteristics and 10 macroeconomic indicators, we enhanced the measurement of systematic risk with improved data dimensionality and granular precision, surpassing previous research;
- We proposed a novel dynamic CAPM that leverages mainstream machine learning algorithms. This innovation, to the best of our knowledge, is among the first applications of machine learning techniques to asset pricing research. By incorporating advanced

methods such as regression trees and neural networks, we effectively tackled the challenges of high-dimensional data, capturing nonlinear and interactive effects, thus providing accurate estimates of systematic risk;

- Our paper unveils the underlying causes of low-risk anomalies and provides valuable implications for academia and industry. We observed that the neural networks, particularly the NN4, yielded the highest excess returns. Liquidity predictors emerged as the most influential factors, followed by momentum indicators. Furthermore, through our subsample analysis, we revealed that, during the transition from SOEs to non-SOEs, the variable importance of fundamental and valuation diminished, making way for liquidity and momentum.

The paper is organized as follows. In Section 2, we present a comprehensive review of the relevant literature. Section 3 provides a detailed description of data sources and methodology. The main empirical findings are presented in Section 4. In Section 5, we further investigate the differences in factors influencing beta through a subsample analysis. In Section 6, we summarize our key findings and conclusions.

2. Literature Review

2.1. The Dynamic CAPM

The CAPM utilizes beta to measure the systematic risk exposure of equities. The traditional static CAPM theory assumes a constant beta, but empirical evidence suggests that actual betas are unstable due to several factors [13]. These factors include changes in a company's strategy, capital structure, and returns, as well as microeconomic aspects such as dividend policies and financial leverage. As a result, variations in firm-level characteristics and the external macroeconomic environment contribute to beta's variability over time.

For several decades, scholars in both academia and industry have been actively engaged in the development of dynamic asset pricing models. Hansen and Richard [14] were among the pioneers in investigating the linear conditional CAPM and time-varying beta to enhance the model's explanatory power, thereby laying the foundation for dynamic CAPM. Ferson and Siegel [15] examined the impact of conditional information on the effectiveness of constructing portfolio strategies within the model. Cederburg and O'Doherty [16] constructed dynamic CAPM incorporating macroeconomic variables such as lagged beta, market dividend rates, and credit spreads, which effectively mitigated risk pricing anomalies. Cosemans et al. [17] introduced macroeconomic indicators, micro-firm variables, and cross-product terms to estimate dynamic stock betas, resulting in superior predictions with significant practical implications.

Empirical findings from various financial markets have consistently demonstrated the enhanced cross-sectional asset returns provided by dynamic models. For example, Mazzola and Gerace [18] conducted comparative analysis of optimal portfolios in the Australian securities market, evaluating the performance of both static CAPM and dynamic CAPM models. Their findings revealed that the dynamic approach, based on pre- and post-transaction cost returns, outperformed the static model due to effective portfolio rebalancing. Chen and Tindall [19] constructed actively managed portfolios of Chicago Board Options Exchange (CBOE) Volatility Index derivatives to reduce portfolio correlation with the equity market. The results indicated that the Kalman filter-based dynamic CAPM produced the best outcomes, generating equity market-neutral portfolios with positive alpha. Hollstein et al. [20] suggested that the dynamic CAPM could effectively explain size, value, and momentum anomalies using high-frequency data, with high-frequency betas providing more accurate predictions compared to those based on daily data. Leal et al. [21] introduced the symmetric CAPM, a dynamic approach that considers distributions with lighter or heavier tails than the normal distribution. They conducted a case study using real data to estimate the systematic risk of financial assets for a Chilean company, and the results revealed that the symmetric CAPM outperformed the traditional CAPM, particularly when dealing with non-Gaussian distributions.

Furthermore, the static CAPM suggests a trade-off between risk and return, while empirical results have shown that high-beta stocks tend to provide lower expected returns than low-beta stocks [22,23]. The reasons behind this low-risk pricing anomaly are complex. Bali et al. [24] attributed the low-risk effect to investors' behavioral biases associated with idiosyncratic risk. Frazzini and Pedersen [25] explained the beta anomaly using the theory of leverage constraints and found that the betting against beta (BAB) factor generated significant alpha. Asness et al. [26] argued that alpha is primarily driven by the betting against correlation factor associated with leverage constraints, rather than the betting against volatility factor related to behavioral effects.

However, in emerging markets, the beta anomaly lacks convincing explanations. China, as the largest emerging market, exhibits unique characteristics, including individual investors as the main players and strict limits on arbitrage. Based on these facts, this paper aims to address the beta anomaly in China by employing machine learning-based dynamic models.

2.2. Application of Machine Learning in Stock Forecast

Given that the stock market is inherently a nonlinear, dynamic, and complex stochastic system, predicting stock prices becomes a challenging task in time-series forecasting. In the past, traditional time-series models such as linear regression, ARIMA, and GARCH were commonly used for stock forecasting [27]. However, with advancements in computer technology and increased computing power, machine learning models such as support vector regression [28], tree-based models [29], and neural networks [30] have shown superior capabilities in handling complex, non-stationary, and nonlinear characteristics compared to traditional models.

Shah [31] published "Machine Learning Techniques for Stock Prediction", which set off a boom in applying machine learning technology to asset pricing area. Subsequently, numerous models have been developed and widely utilized in the prediction of stocks, bonds, options, and other financial fields. Hsu et al. [32] discovered that machine learning methods outperform economic methods in predicting financial markets. They further demonstrated that the prediction accuracy is influenced by market maturity, input variables, forecast benchmark time, and forecast methods. Zhu et al. [33] employed neural networks to forecast stock prices by utilizing 14 trading indicators, including opening and closing prices, as well as technical analysis indicators such as ROC and RSI. Gu et al. [10] compared the predictive power of linear regression with machine learning models for a cross-section of stocks in the US market and reported that the artificial neural network model with three hidden layers exhibited the highest predictive power. Drobetz and Otto [34] evaluated the forecasting performance of machine learning methods in the European stock market and found that support vector machines achieved the best performance and significant profitability.

Several previous studies have applied machine learning methods specifically to Chinese stocks. Yuan et al. [35] developed integrated long-term stock selection models for the Chinese stock market based on various machine learning algorithms. Their analysis revealed that the random forest yielded the best performance for both feature selection and stock price trend prediction. Yu et al. [36] utilized machine learning techniques such as KNN, SVM, and AdaBoost to analyze the correlation between stock returns and their ESG (environmental, social, and governance) scores. The experiment indicated that ESG stocks exhibit better risk performance during normal times compared to non-ESG-related stocks, although they did not deliver excess returns. Leippold et al. [9] examined the prediction ability of machine learning methods in the Chinese stock market and identified liquidity indicators as the most critical factor, followed by price momentum-based signals. Considering the significant role of government signals, they also observed a substantial increase in the predictability of SOEs over longer time horizons.

Overall, in the empirical asset pricing literature, studies employing machine learning-based approaches have addressed many aspects, primarily focusing on the predictability

of stock returns. However, limited research has been conducted on the predictability of systemic risk, which is equally important for firms and investors. Hence, this paper aims to predict the beta of stocks using machine learning techniques and extend the research on the predictability of risk characteristics.

2.3. Related Machine Learning Techniques

In this section, we introduce three types of machine learning algorithms related to our study: linear regression with penalty term, regression trees, and neural networks.

Traditional linear regression often yields unreliable estimates when dealing with a large number of covariates. This unreliability can be attributed to high correlation or redundancy among the covariates, leading to issues such as multi-collinearity and loss of efficiency. In the context of high-dimensional regression, one widely used machine learning technique is linear regression with penalty terms [37]. This approach aims to identify valid predictors by introducing penalty terms into the loss function. In this study, we employ the elastic net (Enet) method to address the problem of overfitting. The combination of penalty terms and the Huber function [38] allows us to identify the most relevant predictors while considering potential outliers and maintaining robustness. The objective function for Enet is defined as follows:

$$L_H^{Enet} = \frac{1}{NT} \sum_{s=1}^{N} \sum_{t=1}^{T} H(R_{s,t} - f^*(z_{s,t-1}; \theta) R_{M,t}; M) + (1 - \rho)\lambda \sum_{j=1}^{P} |\theta_j| + \frac{1}{2}\rho\lambda \sum_{j=1}^{P} \theta_j^2 \quad (1)$$

$$H(x; M) = \begin{cases} x^2, & \text{if } |x| \le M \\ 2M|x| - M^2, & \text{if } |x| > M \end{cases} \quad (2)$$

where λ and ρ are hyperparameters that control the size of the penalty; the key tuning parameters are $\lambda \in (0,1)$ and $\rho \in (0,1)$. Additionally, $R_{s,t}$ is the excess return of stock, $R_{m,t}$ is the excess return of the market portfolio. Finally, $H(x; M)$ is the Huber loss function; the threshold is determined by the tuning parameter M, and θ represents the vector of coefficients.

Next, we explore regression tree models, including random forests (RF) and gradient-boosted regression trees (GBRT), which have gained prominence in machine learning for their ability to handle classification and regression tasks flexibly [35]. These methods rely on the construction of multiple individual trees to make predictions. Mathematically, a basic regression tree with K leaves and depth L can be represented as follows:

$$f^*(z_{s,t-1}; \theta, K, L) = \sum_{k=1}^{K} \theta_k \mathbf{1}_{\{z_{s,t-1} \in C_K(L)\}} \quad (3)$$

where $C_K(L)$ is k-th division of data, and L represents the maximum number of nodes in a complete branch. The stock s with characteristics $z_{s,t-1}$ is clustered into the k-th leaf, and the basic tree predicts the systematic risk as θ_k.

To predict beta, we aggregate forecasts from regression trees into a single one, using ensemble methods, including bagging and boosting. Bagging methods combine the results of multiple parallel models through voting, while boosting obtains the final result by summing the predictions of multiple series models. In this study, GBRT follows the boosting approach, which combines multiple shallow trees to create a single strong learner that surpasses the performance of a deep tree. GBRT iteratively improves the model by sequentially fitting new trees to the residuals of previous iterations, reducing the overall prediction error. On the other hand, RF employs a bagging approach, where bootstrap samples are drawn from the original dataset and each sample is used to train an independent decision tree, the predictions are then averaged to create a strong learner. The construction of RF is illustrated in Figure 1.
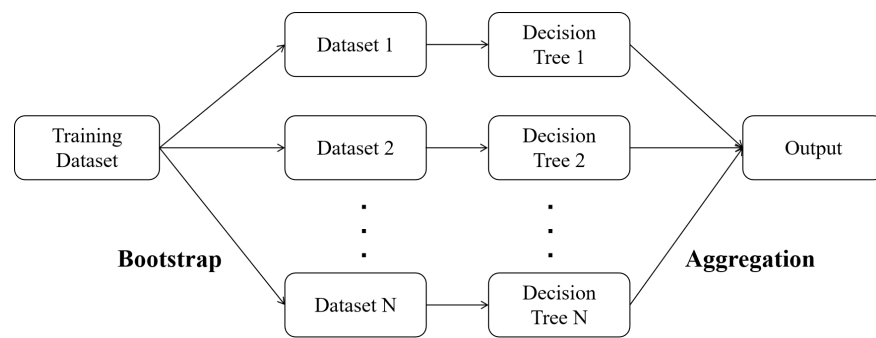
Figure 1. The construction of the random forest.

The steps involved in the RF model are as follows [39]. Firstly, N training datasets are generated using the bagging method by sampling from the original training dataset. Secondly, N decision trees are trained independently based on these N training datasets. Thirdly, the random forest is composed of these N decision trees. In classification problems, the final classification result is determined by aggregating the predictions of the N decision tree classifiers. For regression problems, the average of the predicted values from the N decision trees determines the final prediction result. In summary, random forests are formed by integrating a large number of decision trees as the basic unit.

Moving on, we delve into neural networks (NN), which are widely employed in stock return forecasting due to their ability to construct nonlinear models effectively [40]. Similar to the human brain, neural networks consist of interconnected computational units called neurons. While individual neurons may provide limited predictive power, the collective power of a network composed of multiple neurons is substantial. With the highly parameterized and fault-tolerant nature, they are well-suited for solving complex problems [41]. However, it is important to note that the interpretability of neural network algorithms is limited, and their predictions may not always be easily explained or understood.

In our analysis, we utilized feed-forward neural networks. The model included input layers to capture lagged stock-level features, hidden layers to capture interactions between input predictors, and output layers to generate predictive outputs (realized betas). Each node in the network is connected to all nodes in the preceding layer, following a one-way direction. The structure of the neural network is illustrated in Figure 2. We chose the number of neurons in the hidden layer according to the geometric pyramid rule [42], referring to the practice of Gu et al. [10] and Leppiod et al. [9].
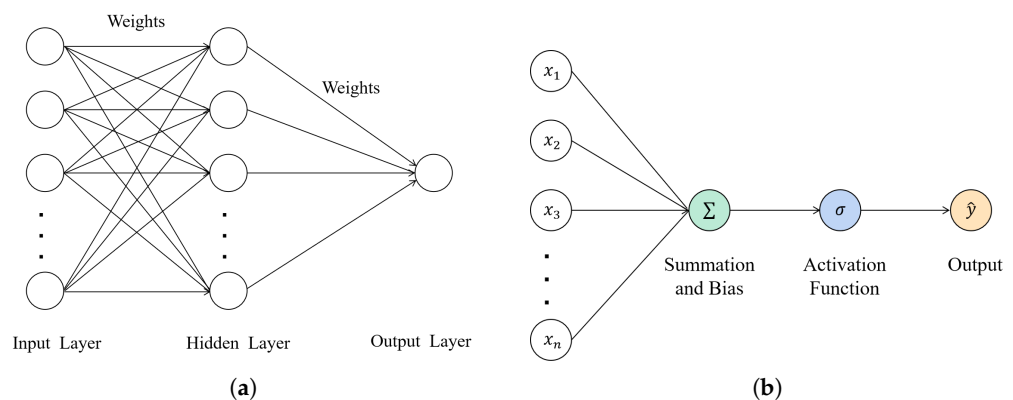

Figure 2. The construction of neural networks. (**a**) Two-layer fully connected neural network. (**b**) Activation function.

We considered neural networks with one to five hidden layers. Each layer consists of a certain number of neurons, built using commonly-used activation functions. For instance, the predicted systematic risk under the NN2 model can be represented as follows:

$$\beta_{s,t} = \alpha_1 + W_1\sigma(\alpha_2 + W_2\sigma(\alpha_3 + W_3 z_{s,t-1})) + \varepsilon_{s,t} \tag{4}$$

where $\sigma()$ represents activation function, $\{a_1, a_2, a_3\}$ is the bias set, $\{W_1, W_2, W_3\}$ is the predicted weight matrics, and $\varepsilon_{s,t}$ is the residual term. The specific parameter configurations for NN1 to NN5 are presented in Appendix A—Table A1.

3. Methodology

3.1. Data

In this study, we collected individual stock return data and financial data for all A-share companies listed on the Shanghai and Shenzhen stock exchanges from China Stock Market and Accounting Research (CSMAR) and Wind, which are the two most influential financial databases in China. The corresponding monthly macro data were obtained from the CSMAR and National Bureau of Statistics websites. We calculated risk-free rates using monthly one-year government bond yields, and market portfolio returns were derived by taking a weighted average of the market value of all equity returns. The complete sample period spanned from January 2002 to December 2020, covering a total of 19 years. To mitigate the effects of outliers generated by special stocks, we excluded stocks that were classified as special treatment, delisted, or listed for less than one year. As a result, the final sample consisted of 3619 stocks.

3.2. Machine Learning-Based Dynamic CAPM

Based on the literature related to the cross-section of stock returns and studies on market-specific factors in China, we collected 70 variables documented in Green et al. [43] and Ma et al. [44] to build a large-scale micro-firm factor database. These variables were categorized into 9 groups: beta, valuation (bpr), earnings (ey), growth, leverage (lever), liquidity (liq), momentum (mom), size, and volatility (vol). Additionally, we included 4 binary variables representing the ownership type of listed companies for subsample analysis. To ensure the accuracy of stock return forecasting, we excluded factors that were updated semi-annually and annually. Regarding data frequency, 17 firm-level characteristics were updated monthly, while 53 were updated quarterly. To mitigate the impact of outliers on prediction results, we followed the approach of Gu et al. [10] and standardized all company characteristics by mapping their values to the $[-1,1]$ interval. Appendix A—Table A3 provides detailed information on characteristics and categories.

Furthermore, we constructed 10 macroeconomic predictors commonly used in stock market analysis. These predictors fell into two categories based on the definitions provided by Welch [45]: national economic indicators, such as consumer confidence index (cei), macroeconomic prosperity index (hj), inflation (infl), and M2 growth rate (m2gr); and stock market monitoring indicators, such as book-to-market ratio (bm), dividend–price ratio (dp), earnings–price ratio (ep), volatility (lvol), stock variance (svar), and turnover rate (to). These indicators have been previously validated as effective in relevant studies and are summarized in Appendix A—Table A4.

A fundamental assumption in the CAPM is that the beta used to measure asset risk is constant. However, subsequent research has shown that beta coefficients exhibit time-varying properties. Existing studies on time-varying betas often rely on small datasets with irregular sample selection, failing to leverage the valuable information contained in macro- and micro-influences within big data. Moreover, these studies often focus solely on traditional linear regression methods to estimate betas, disregarding techniques such as factor reduction, variable screening, and nonlinear models, which have the potential to enhance predictive accuracy. Given the limitations in the current literature, this paper

aimed to address these gaps by constructing a dynamic CAPM using machine learning algorithms:

$$R_{s,t} = \alpha_s + \beta_{s,t} R_{m,t} + \mu_{s,t} \tag{5}$$

where $\beta_{s,t}$ is the time-varying beta for systematic market risk measure, α_s is the excess return of stock s under CAPM, and $\mu_{s,t}$ is the residual term.

Machine learning-based approaches offer a distinct and rigorous methodology for capturing cross-sectional variations in future beta. These techniques focus explicitly on predicting beta with multiple sources of information as predictors. For example, the achieved beta is included as a dependent variable in the regression, helping the recall of the forecast target and potentially maximizing the predictive power. In this study, we employed a generalized additive prediction error model, as outlined in Gu et al. [10], to characterize stock beta and its influential predictors:

$$\beta_{s,t} = E_t[\beta_{s,t}] + \epsilon_{s,t} \tag{6}$$

We further assumed that, given the available information in period $t - 1$, the conditional expectation of beta can be expressed as a functional form of a set of predictors:

$$E_t[\beta_{s,t}] = f^*(z_{s,t-1}) \tag{7}$$

where the specific functional $f^*()$ remains unspecified, and $z_{s,t-1}$ represents a set of p-dimensional predictors. Despite belonging to different families, the machine learning-based models employed in this study were all designed to optimize prediction performance.

To examine the combined influence of macroeconomic and micro-firm characteristics on time-varying beta, we defined the stock-level covariates $z_{s,t-1}$ as cross-product terms:

$$z_{s,t-1} = x_{t-1} \otimes c_{s,t-1} \tag{8}$$

where $c_{s,t-1}$ is a $P_C \times 1$ matrix of the firm's micro-feature matrix, x_{t-1} is a $P_X \times 1$ matrix of the macro-variables, and $z_{s,t-1}$ represents a $P_C P_X \times 1$ matrix of the combined macro–micro dataset. The total number of covariates was $70 \times (10 + 1) = 770$. It is worth noting that $z_{s,t-1}$ contains the interactions between stock-level characteristics and macroeconomic state variables, thereby providing deeper insights into stock return forecasts.

In total, we employed nine machine learning methods to construct the dynamic CAPM: (1) linear regressions, including partial least squares (PLS) and Enet; (2) regression trees, including GBRT and RF; (3) neural networks, ranging from one to five layers. Both the Enet and GBRT models were equipped with Huber loss functions to mitigate potential interference from extreme values [38]. This paper primarily utilized the Sklearn library in Python to construct a dynamic CAPM, with comprehensive details regarding the computational framework and hyperparameters provided in Appendix A—Table A2.

Considering the time-series continuity, we divided the data into three periods: the training sample (2002–2007), the validation sample (2008–2010), and the test sample (2011–2020). The training sample was used to estimate the model parameters based on pre-specified hyperparameters, while the validation sample was utilized to optimize the hyperparameters of models. The test sample contained data for the subsequent 12 months after validation to assess the models' predictive performances.

Given the computational complexity associated with machine learning, we adopted the sample splitting approach introduced by Gu et al. [10] to annually update the prediction model. When modifying the model, we expanded the size of the training sample by 1 year, while the validation period and the 1-year test period shifted forward to include the most recent 12 months. The data spanning from January 2002 to December 2020 were divided into 10 periods, as illustrated in Figure 3.
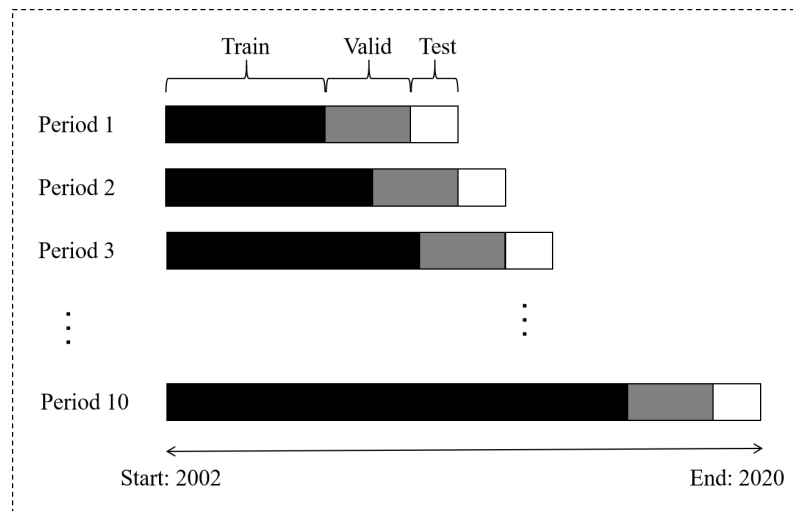
Figure 3. Rolling window of sample division.

3.3. Performance Evaluation

In evaluating the performance of our stock return forecasts, we followed the standard approach commonly employed in the literature [9,34]. To assess the predictive accuracy of individual excess stock return forecasts, we calculated the non-demeaned out-of-sample R^2 using the following formula:

$$R^2_{\text{oos}} = 1 - \frac{\sum_{(s,t)\in T_3}\left(R_{s,t} - \widehat{R}_{s,t}\right)^2}{\sum_{(s,t)\in T_3} R^2_{s,t}} \tag{9}$$

where $R_{s,t}$ denotes the actual return of stock s at time t, and $\widehat{R}_{s,t}$ denotes the predicted return rate of stock s at time t. T_3 indicates that the calculation is restricted to the testing sample; in other words, the data never enter into model estimation or tuning.

Throughout the above analysis, we constructed a modelling flow for the dynamic CAPM based on various machine learning algorithms, as shown in Figure 4.
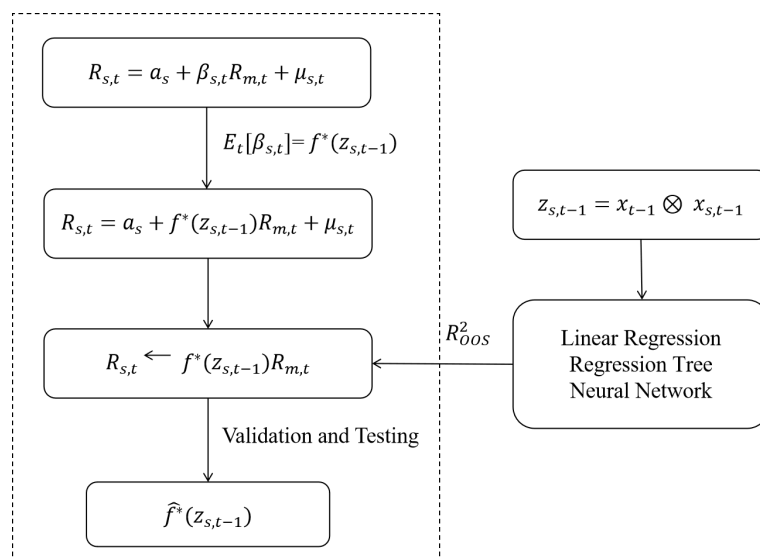

Figure 4. Dynamic CAPM modeling process based on machine learning.

4. Experimental Results

4.1. Low-Risk Pricing Anomaly in China

At the beginning of each January, the beta value for each stock was calculated using the daily stock returns over the previous 12 months, following the approach of Hong and Sraer [23]:

$$r_{s,t} = a_s + \beta_{s,0}r_{m,t} + \beta_{s,1}r_{m,t-1} + \beta_{s,2}r_{m,t-2} + \beta_{s,3}r_{m,t-3} + \beta_{s,4}r_{m,t-4} + \beta_{s,5}r_{m,t-5} + \varepsilon_{s,t} \quad (10)$$

where $r_{s,t}$ is the excess return of stock s at period t, $r_{m,t}$ is the excess return of the market portfolio at period t, $r_{s,t-1}, \ldots, r_{s,t-5}$ are lagged excess returns of the market portfolio, and $\varepsilon_{s,t}$ is the regression residual.

To account for variations in the dissemination of market information, the model incorporated five lagged excess returns, considering factors such as illiquidity and asynchrony. The value of predicted beta was the sum of parameter estimates:

$$\widehat{\beta}_{s,t} = \widehat{\beta}_{s,0} + \widehat{\beta}_{s,1} + \widehat{\beta}_{s,2} + \widehat{\beta}_{s,3} + \widehat{\beta}_{s,4} + \widehat{\beta}_{s,5} \quad (11)$$

The stocks were categorized into 10 groups based on their beta values, arranged in descending order. Portfolios were constructed for each group, and they were held throughout the year until December. The total monthly returns for 10 portfolios were calculated over a 10-year period, spanning from January 2011 to December 2020. The results are presented in Table 1.

Table 1. The mean and variance of portfolio return.

Portfolio	Beta	Mean	Variance	Portfolio	Beta	Mean	Variance
1	0.664	1.007	0.013	6	1.219	1.228	0.021
2	0.907	0.915	0.017	7	1.283	1.112	0.016
3	1.014	0.739	0.017	8	1.353	0.744	0.012
4	1.091	1.319	0.018	9	1.442	0.581	0.015
5	1.157	1.262	0.014	10	1.626	0.432	0.017

Table 1 displays the mean of and variance in monthly returns for the 10 portfolios, where *Beta* represents the market capitalization-weighted beta of stocks in the portfolio; *Mean* represents the mean of the market capitalization-weighted monthly returns of all stocks in the portfolio in %; and *Variance* represents the variance in monthly returns of all stocks in the portfolio. Notably, the portfolios with lower beta values (1–5) generally exhibited higher returns compared to the portfolios with higher beta values (6–10), indicating a decreasing trend in returns as beta increased. These results suggest that the static beta from traditional CAPM cannot adequately explain the pattern of stock returns. In other words, higher risk does not necessarily correspond to higher returns. Therefore, there exists a low-risk pricing anomaly in the Chinese equity market.

To further analyze the anomaly under the static CAPM, the time series of Portfolio 1 (L) and Portfolio 10 (H) were selected, which represented the lowest and highest beta portfolios:

$$r_{s,t} = a_s + \beta_s r_{m,t} + \varepsilon_{s,t} \quad (12)$$

The regression results are presented in Table 2. The alpha was 0.29% for Portfolio L and -0.53% for Portfolio H over the sample period. By investing long on Portfolio H and short on Portfolio L, the alpha of the long–short portfolio H–L was -0.82%, which was statistically significant at the 5% level. This finding indicates that the static CAPM fails to address the low-risk pricing anomaly.

Table 2. The regression results of portfolios under the static CAPM.

Portfolio	CAPM_Alpha	CAPM_Beta	R^2
L	0.29 (1.53)	0.77 (25.18)	0.84
H	−0.53 (−1.41)	1.50 (25.26)	0.84
H–L	−0.82 (−1.73)		

4.2. Dynamic CAPM with Time-Varying Beta

This section examines the performance of machine learning-based dynamic CAPM in predicting future stock returns, with a focus on addressing the low-risk anomaly. We constructed time-varying betas using different families of models: linear regression, regression trees, and neural networks. These are all representative mainstream machine learning algorithms in the financial field.

We utilized the return time-series data of Portfolio H and Portfolio L, as constructed in Section 4.1. The training period spanned six years, from January 2002 to December 2007, while the verification period covered three years, from January 2008 to December 2010. Given the computational intensity and time-consuming nature of machine learning, we adopted the sample splitting method introduced by Gu et al. [10] to refit the models annually. In our experiment, the first monthly time-varying beta to be predicted was for January 2011. The final prediction interval extended from the beginning to December 2020, including a total of 120 monthly return observations.

Table 3 presents the regression results of the dynamic CAPM based on different machine learning techniques. To demonstrate the efficacy of our proposed dynamic CAPM model in addressing the low-risk pricing anomaly, we utilized the excess return (a%) of the long-short portfolio H–L as the validating performance metric. This metric provided a measure of the improvement in returns achieved by our model. Furthermore, we ensured the accuracy and reliability of our research results by assessing the regression values and statistical t-values associated with various machine learning algorithms. For the regularized models, including PLS and Enet, the monthly excess returns for the H–L portfolio were −0.45% and −0.32%, respectively, with corresponding t-values of −0.73 and −0.52. Compared to the static CAPM, these improved models provided a more than 50% increase in excess returns, indicating the effectiveness of dimensionality reduction when dealing with a large number of covariates. This suggests that, when employing factor analysis to predict stock market returns, certain micro-level stock characteristics may be redundant, supporting the conclusions of Ma et al. [44] regarding the existence of the factor zoo phenomenon in the Chinese market.

Additionally, in terms of the predictive ability of time-varying beta, the linear regression models performed better on high-beta portfolios compared to the low-beta group. For PLS and Enet, the excess return for portfolio H increased from −0.53% to 0.21% and 0.43%, respectively. This improvement can be attributed to the higher return volatility in the high-beta portfolio, wherein dimensionality reduction techniques and penalty terms effectively enhanced the stability of the time-series data.

Among all the models, the RF and NN models significantly addressed the low-risk pricing anomalies. Particularly, the dynamic model based on NN4 yielded the best performance, with an excess return of 0.36%, which was 118% higher than that of static CAPM. This result emphasizes the superiority of machine learning in capturing complex interactions among predictors, which are often overlooked by traditional linear models. It is worth noting that, although the complexity of an NN model increases by adding hidden layers, the excess returns of the long–short portfolio H–L initially rose and then declined; eventually, the five-layer model failed to improve over NN4. This observation indicates that, while machine learning improves excess returns, its benefits are limited in mitigating the pricing anomaly in the Chinese stock market.

Table 3. The regression results of dynamic CAPM based on machine learning.

Portfolio	Indicator	CAPM	PLS	Enet	GBRT	RF	NN1	NN2	NN3	NN4	NN5
L	a (%)	0.29 [1]	0.66	0.75	0.66	0.52	0.56	0.55	0.54	0.36	0.43
		(1.53)	(1.43)	(1.61)	(0.74)	(0.61)	(0.69)	(0.66)	(0.67)	(0.49)	(0.51)
	β_s	0.77 [2]	0.10	0.15	0.24	0.29	0.28	0.28	0.27	0.25	0.24
		(25.18)	(1.37)	(2.07)	(1.73)	(2.16)	(2.20)	(2.21)	(2.17)	(1.70)	(1.82)
	$\beta_{s,t}$		−0.36 [3]	0.37	1.74	1.91	2.38	2.10	2.00	1.22	1.56
			(−3.27)	(3.32)	(4.05)	(5.04)	(7.01)	(6.33)	(6.41)	(2.65)	(5.67)
H	a (%)	−0.53	0.21	0.43	0.88	0.78	0.82	0.80	0.80	0.72	0.77
		(−1.41)	(0.33)	(0.48)	(1.91)	(1.72)	(1.89)	(1.83)	(1.84)	(1.52)	(1.76)
	β_s	1.50	0.18	0.29	0.13	0.15	0.14	0.15	0.14	0.13	0.13
		(25.26)	(1.26)	(2.10)	(1.80)	(2.10)	(2.15)	(2.14)	(2.13)	(1.77)	(1.87)
	$\beta_{s,t}$		−0.79	0.89	0.83	0.79	1.05	0.91	0.88	0.54	0.74
			(−3.66)	(4.18)	(3.76)	(3.93)	(5.74)	(5.12)	(5.26)	(2.28)	(5.18)
H–L	a (%)	−0.82	−0.45	−0.32	0.22	0.26	0.25	0.26	0.27	0.36	0.34
		(−1.73)	(−0.73)	(−0.52)	(0.35)	(0.44)	(0.45)	(0.44)	(0.46)	(0.58)	(0.56)

[1] a (%) represents the excess return, and t-value is in parentheses; [2] β_s corresponds to the static beta coefficient obtained for the $R_{m,t}$ item regression; [3] $\beta_{s,t}$ corresponds to the dynamic beta coefficient obtained for the $f^*(z_{s,t-1})R_{m,t}$ item regression.

In summary, the integration of machine learning into the dynamic CAPM effectively addresses the "low risk with high return" anomaly in China's A-share market. The impressive performance of neural networks highlights the benefits of incorporating potentially complex interactions among predictors.

4.3. Determining Which Predictors Are Important

Understanding the economic mechanisms behind machine learning models is crucial, as they often lack transparency in economic reasoning [46]. With a multitude of influencing factors, we aimed to identify the predictors that matter in constructing time-varying beta. We examined the variable importance of 70 micro-firm characteristics and 10 macroeconomic characteristics in each prediction model. Following Gu et al. [10], we calculated the decrease in R^2 when setting the given predictor to zero, and then averaged the results across training samples to obtain a single importance measure for each predictor. The variable importance was normalized to sum to one, allowing for a explanation of relative importance. In general, the greater the drop in R^2 after elimination, the more important the predictor.

Firstly, we investigated the impact of micro-characteristics on beta forecasts, and this impact varied depending on the specific prediction model employed. Figure 5 presents the overall influence of micro-features across the entire sample. The vertical axis represents each characteristic, with the feature rankings sorted from high to low. This ranking reflects the combined contribution of each feature across all models, where the most influential features are at the top and the least influential features are at the bottom. On the horizontal axis, each applied machine learning technique is represented. Additionally, the color gradient indicates the variable importance of the predictors in their respective models. A dark gradient signifies the most influential features, while a light gradient represents the least influential features.

We found that each individual model consistently selected the most decisive predictors, which could be categorized into three groups. Market liquidity features were most influential, including volatility of turnover (std_turn, std_rvol, and chato), illiquidity indicators (illiq), and number of zero trading days (zero). The second-most influential group comprised recent price trends, such as short-term momentum (mom1m), momentum change (chmom), and recent maximum return (maxret). The third group comprised fundamental features and valuation ratio, including market capitalization (mve), sales and inventory changes (si and sp), and book-to-market ratio (bm).
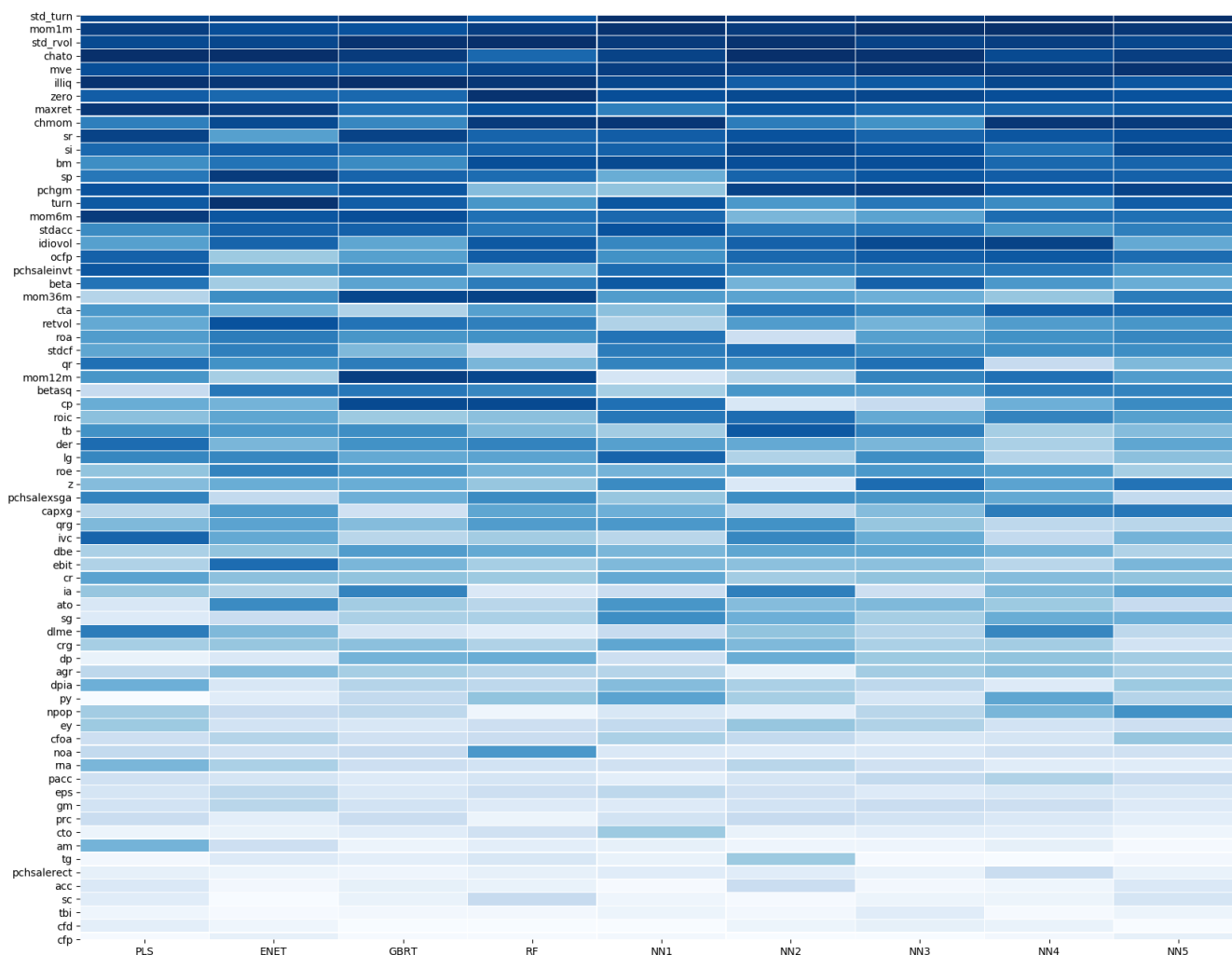
Figure 5. Variable importance of predictors for all models.

Our experimental results differed from the conclusions of previous research. Leippold et al. [9] suggested that, with the exception of recent maximum return, price trend features have minimal impact on the Chinese stock market. However, our study reveals that short-term momentum reversals are the second-most decisive predictor, and other momentum features also exhibit significant influence, aligning with previous studies on momentum and reversal effects [47]. On the other hand, Gu et al. [10] demonstrated that, in the US market, fundamental and valuation ratios have lower overall variable importance compared to risk measure features, such as total return volatility (retvol) and idiosyncratic return volatility (idiovol). However, our findings present a contrasting perspective. We find that, apart from market capitalization, fundamental signals play a crucial role in predicting stock returns. Predictors such as si, bm, and sp exhibit notable influences, consistent with previous research in the Chinese stock market [48].

Machine learning algorithms are often referred to as "black boxes" due to their complexity and opacity; thus, we attempted to explain influential predictors from the underlying economic mechanism. Firstly, the turnover indicator is commonly used to reflect the liquidity of the stock market, where higher turnover indicates active trading and strong liquidity. Frequently fluctuating turnover (std_turn, std_rvol) is often associated with high volatility risk and speculation risk, and stocks with speculative tendency always carry high risk expectations. Therefore, liquidity features form the first influential group in time-varying beta forecasts. Secondly, the momentum effect, a well-known market anomaly, has been observed across various stock markets. The existing literature suggests that, in the

Chinese market, the short-term momentum effect and long-term reversal effect are related to the unique phenomenon of a large proportion of short-dated investors, such as retail investors and lobbyists [49]. Among different momentum features, short-term momentum (mom1) demonstrates superior explanatory power and forecasting ability, as a sustained high-return trend in the short term implies a high-risk state. Thirdly, fundamental and valuation features play a crucial role in evaluating performance, estimating value, and monitoring operations. Companies with strong turnover capacity, high-quality assets, and a positive growth trend tend to have relatively stable share prices and are less exposed to systemic risk.

Furthermore, we investigated the differences between various machine learning methods in the variable importance of predictors. The results revealed that regression tree models place more emphasis on predictors such as 12-month momentum (mom12), 36-month momentum (mom36), and cash productivity (cp). The reason may be that regression trees randomly select a subset from all features and then split the nodes in each decision tree. Unlike the bagging method, which selects the optimal attribute from all sets, the subset-based selection of splitting attributes enhances training efficiency. Consequently, these predictors carry more weight in certain decision trees, making them more relevant to the overall models. Liquidity and momentum features are the two most important for tree models, although their rankings slightly differ from other models, particularly concerning the variable importance of medium-term and long-term momentum. In addition, neural network models demonstrate a preference for liquidity, fundamental, and valuation features, with predictors such as mve, sp, and si standing out. This finding can be attributed to the variation in variable importance with the time dimension in neural networks. By employing appropriate training algorithms and hyperparameter adjustments, these models can accurately forecast stock returns, reflecting the flexibility and high fault tolerance.

We also explored the impact of macro-characteristics on time-varying beta. The variable importance measures for 10 macroeconomic predictors, normalized to sum to one within each model, are presented in Figure 6.
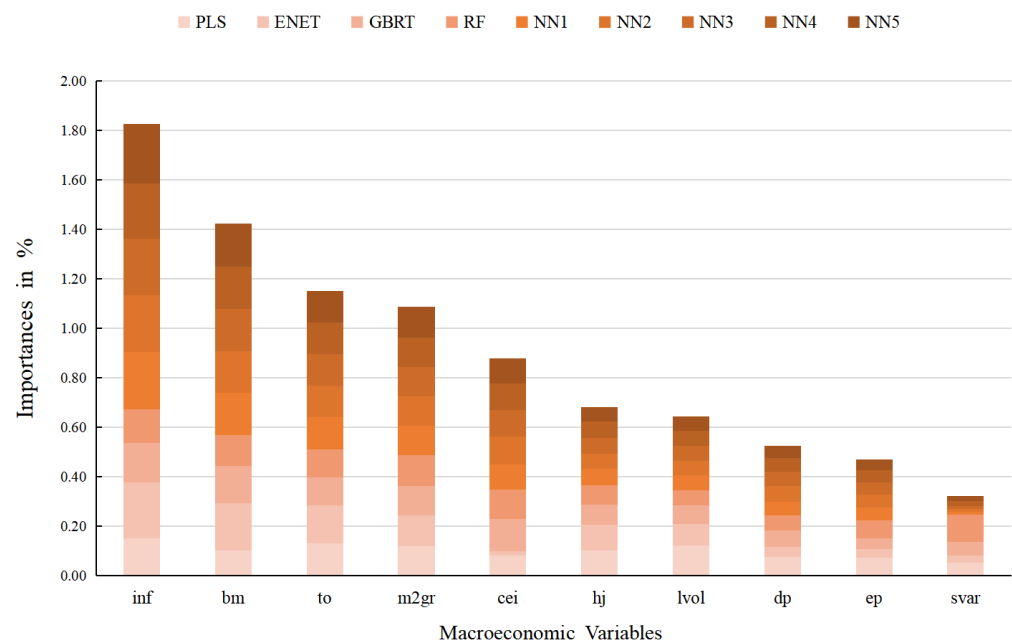

Figure 6. Variable importance of macro-predictors for all models.

The results indicated that inflation (inf) and book-to-market ratio (bm) are crucial macro-predictors in beta forecasting for all models, particularly in neural networks. Conversely, the effects of dividend–price ratio (dp), earnings–price ratio (ep), and stock variance (svar) were deemed negligible, except for a certain proportion observed in the PLS model.

5. Subsample Analysis

The Chinese stock market possesses a range of unique characteristics, including a significant presence of SOEs and stringent regulations on short selling. Given the prominence and distinct nature of SOEs, it is crucial to examine their impact separately. In this section, we explore the impact of dynamic CAPM on SOEs and non-SOEs, and reveal the deep reasons for the differences through factor analysis.

5.1. SOEs vs. Non-SOEs

By the end of 2022, the number of SOEs in China's A-share market accounted for 27.8%, and their total market value reached 49.86%, indicating a significant difference from other equity markets. SOEs have large market capitalization, as they often represent leading companies in industries such as military, chemicals, and utilities. Therefore, it was essential to examine whether the high proportion of SOEs affects stock return forecasts. We divide stocks into subsamples comprising SOEs and non-SOEs. Out of the 3619 stocks in our dataset, 1090 were classified as SOEs, while the remaining 2529 were non-SOEs. The excess returns for each subsample are reported in Table 4.

Table 4. The excess returns of SOEs and non-SOEs.

Sample	Portfolio	CAPM	PLS	ENET	GBRT	RF	NN1	NN2	NN3	NN4	NN5
SOE	L	0.40 (1.70)	1.05 (2.16)	1.01 (1.86)	1.15 (1.08)	1.25 (1.18)	0.29 (0.46)	0.24 (0.38)	0.22 (0.32)	0.08 (0.42)	0.20 (0.30)
	H	−0.14 (−0.39)	1.02 (1.08)	0.94 (0.91)	1.29 (2.45)	1.35 (2.58)	0.68 (1.96)	0.65 (1.93)	0.66 (1.72)	0.57 (1.61)	0.65 (1.71)
	H–L	−0.54 (−1.21)	−0.14 (−0.19)	−0.07 (−0.10)	0.14 (0.18)	0.10 (0.13)	0.39 (0.61)	0.41 (0.65)	0.44 (0.69)	0.49 (0.77)	0.45 (0.70)
non-SOE	L	1.54 (3.69)	1.51 (3.87)	1.55 (3.67)	1.60 (3.79)	1.57 (3.71)	1.60 (3.80)	1.53 (3.46)	1.51 (3.54)	1.56 (3.64)	1.54 (3.61)
	H	2.21 (1.30)	2.14 (1.27)	2.55 (1.49)	3.00 (1.84)	3.10 (1.99)	2.87 (1.75)	2.87 (1.60)	3.30 (2.09)	4.06 (2.90)	4.07 (2.99)
	H–L	0.67 (0.44)	0.64 (0.42)	0.99 (0.65)	1.40 (0.96)	1.53 (1.12)	1.27 (0.86)	1.34 (0.83)	1.80 (1.31)	2.52 (2.13)	2.53 (2.25)

The regression results demonstrated that the NN-based dynamic CAPM, particularly the 3–5 layer models, exhibited a positive and robust impact on pricing anomaly for both SOEs and non-SOEs. In addition, the performance of regression trees were mixed. RF and GBRT performed exceptionally well on non-SOE stocks, surpassing the linear regression models and even outperforming the 1–2 layer neural networks. However, their performance on the SOEs subsample was far from impressive.

Notably, two findings regarding SOEs stand out. Firstly, the neural networks consistently outperformed all other models for SOEs, while, for non-SOEs, achieving comparable performance was feasible with other methods. Secondly, as the number of layers increased, the excess return initially increased and then decreased, reaching a peak at NN4, which aligned with the results from the full-sample analysis. Based on these observation, we argue that the beta forecasts of SOEs require adaptable and flexible algorithms to capture the nonlinear effects among the predictors. This could be attributed to factors such as opaque information disclosure and specific social utility, which undermine the corporate performance of SOEs.

Based on the subsample analysis, machine learning approaches—particularly, multi-layer neural networks—demonstrate their ability to address low-risk pricing anomaly in SOEs. Additionally, our proposed dynamic CAPMs showed improvement in excess returns for stock forecasting in non-SOEs, with almost all methods performing well.

5.2. The Predictability of Neutral Network

Neural networks consistently outperformed other models in terms of predictability performance, both in full samples and subsamples. However, their black-box nature made interpretation challenging, and it is difficult to find corresponding support in economic theory [50]. To gain intuitive insights into predictors that contribute to stock predictability, we focused on the NN4 model, which performed best in improving the low-risk pricing anomaly. This section aims to compare the differences in variable importance between SOE and non-SOE stocks under the NN4 model.

Figure 7 illustrates the fluctuations in variable importance among the 20 most influential predictors in the analysis of the full sample. The red gradient denotes an increase in variable importance, the green gradient denotes an decrease, and the white denotes stability. The rankings of the three most important predictors remained largely unchanged when transitioning from SOEs to non-SOEs. These predictors were as follows: (1) mom1m, 1-month momentum, which represents the cumulative daily return from the end of month $t-1$ and serves as a short-term momentum indicator; (2) std_turn, the monthly standard deviation of daily share turnover, which proxies for liquidity; and (3) mve, market capitalization, which is used to measure firm size. Given that SOEs often represent leading companies in various industries with large market capitalization, there exists a strong correlation between company size and the classification of SOE or non-SOE stocks.
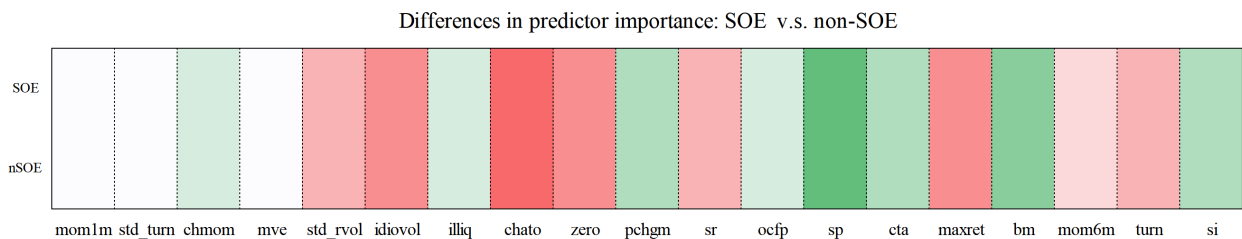
Differences in predictor importance: SOE v.s. non-SOE

Figure 7. Relative variable importance of predictors.

While the variable importance of the top-ranked indicators remained largely consistent, in other predictors, it changed noticeably. We observed that liquidity features such as std_rvol, chato, and zero gained more weight for non-SOEs, aligning with previous research indicating that investors pay more attention to market liquidity for small-cap stocks [20]. In contrast, fundamental signals such as ocfp, sp, cta, and bm received less weight when transitioning from SOEs to non-SOEs. In other words, under the NN4 model, valuation, earnings, and growth features exert greater influences on the predictability of SOEs.

Furthermore, changes in volatility features, such as idivol and maxret, are also of interest. Firstly, the variable importance of idivol increases when shifting from SOE to non-SOE stocks. Since non-SOEs generally have smaller market capitalization and idivol is an influential predictor for small-cap stocks, our findings support the limited arbitrage theory, which suggests that the predictability increases when the anomaly of stocks with high idiosyncratic risk strengthens [51]. Secondly, the variable importance of maxret increased for non-SOE stocks, which further corresponds to the unique characteristics of the Chinese market. Retail investors have a particular preference for small-cap stocks with significant upside potential; thus, extreme positive returns display considerable predictability in asset pricing.

Next, we focused on predictor categorization to delve deeper into the differences in stock return forecasts between SOE and non-SOE. Following the classification methods employed by Gu et al. [10], we categorized all predictors into nine distinct groups: beta (C_beta), valuation (C_bpr), earnings (C_ey), growth (C_growth), leverage (C_lever), liquidity (C_liq), momentum (C_mom), size (C_size), and volatility (C_vol). Specific classification details are presented in Appendix A—Table A3.

Figure 8 illustrates the disparities in group importance between SOEs and non-SOEs. We observed a slight increase in the variable importance of liquidity and momentum,

which were considered the two most critical groups when transitioning from SOEs to non-SOEs. Notably, the differences in pricing anomalies between subsamples were primarily influenced by fundamental and volatility features. In the NN4 model, the beta of SOEs assigned higher weights to fundamental signals (C_size, C_ey) and growth features (C_bpr, C_growth). This can be attributed to the concentration of SOEs in the infrastructure industry, where stable operating capital flows make it challenging for investors to pursue short-term excessive profits. Therefore, fundamental indicators play a decisive role in stock forecasts for SOEs. On the other hand, non-SOEs place greater emphasis on predictors related to volatility and corporate leverage. This finding aligns with previous research suggesting that private companies are less resilient to risk, even with high growth rates and profitability [52].
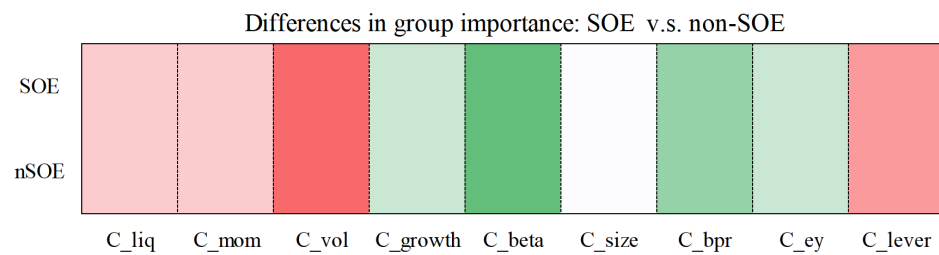

Figure 8. Relative importance of variable categories.

Overall, the variable importance of micro-firm characteristics and their associated categories align with previous studies in the Chinese stock market. Our analysis reveals that predictors such as short-term momentum, market liquidity, and market capitalization significantly influence the behavior of the dynamic CAPM model for each subsample. When transitioning from SOEs to non-SOEs, we observe a gradual shift in relative importance, with fundamental signals and valuation indicators being gradually replaced by liquidity and momentum. Additionally, the volatility category appears to play a more substantial role in smaller firms.

6. Conclusions and Future Work

6.1. Conclusions

Based on the analysis of a large database comprising 70 micro-firm characteristics and 10 macroeconomic indicators, our research highlights the effectiveness of machine learning-based dynamic CAPM in mitigating the low-risk pricing anomaly in the Chinese market. Our findings demonstrate the superiority of nonlinear models, such as RF and NN, over linear regression models in stock return forecasts. Among the neural network models, NN4 stands out as the best performer, leveraging its ability to capture complex interactions among predictors. Additionally, we have identified liquidity features as the most critical predictors influencing time-varying beta, with momentum as the second-most important, followed by fundamental signals and valuation ratios.

Moreover, our investigation into the subsample analysis of SOEs and non-SOEs revealed the significant capabilities of machine learning-based dynamic CAPM. Multilayer neural networks display substantial capability in addressing the beta anomaly prevalent in SOEs. When transitioning from SOEs to non-SOEs, we observe that fundamental and valuation features gradually give way to liquidity and momentum in relative variable importance. Notably, volatility features play a more influential role in the stocks of smaller capitalization firms.

The implications of our research are twofold and contribute significantly to asset pricing research. Firstly, our study highlights the value of employing machine learning techniques to capture systematic risk, providing insights into the underlying economic explanations of the low-risk pricing anomaly. By demonstrating the superiority of nonlinear models, we offer compelling evidence for the adoption of machine learning in asset pricing research, which can lead to more accurate and robust forecasts.

Secondly, considering the differences in market structure between the Chinese market and mature capital markets, our findings open up exciting avenues for future research to compare systemic risk factors across different regions and countries. Understanding these variations can enhance our understanding of market dynamics and contribute to the broader field of finance and investment research.

Overall, our study successfully showcases the practical application of machine learning techniques in asset pricing models within the Chinese markets. By providing a deeper understanding of the factors influencing pricing anomalies, we contribute to the Fintech field and inspire further exploration of novel methodologies and cross-market comparisons. The insights gained from our research have implications for practitioners, regulators, and academics, offering new perspectives on risk assessment in dynamic market conditions.

6.2. Future Work

The proposed approach in this research has theoretical and practical limitations, which may impact the generalizability and applicability of the findings. Firstly, while machine learning techniques offer improved performance in capturing complex patterns and interactions, the interpretability of the models might be compromised. The black-box nature of these algorithms makes it challenging to provide clear economic interpretations for the relationships between predictors and time-varying beta. These limitations restrict the extent to which the findings can be generalized and applied in practical settings that require transparent decision-making processes.

Secondly, the findings of this study are based on the Chinese stock market, which has unique characteristics and institutional structures. These features, such as the dominance of retail investors and the presence of SOEs, might limit the generalizability of the research findings to other markets or regions. The dynamics of risk factors and the impact of predictor variables on time-varying beta could differ significantly in different markets. Therefore, caution should be exercised when directly applying the results to other stock markets, particularly those with distinct market structures or regulatory environments.

There are several promising avenues for further research that have emerged. In order to gain a comprehensive understanding of the effectiveness and robustness of dynamic CAPM, future research should focus on conducting rigorous cross-market comparisons. Comparative analyses across diverse markets, encompassing both emerging and mature economies, would provide valuable insights into how this approach performs under various market conditions. By examining the role of market structure and investor behavior in shaping asset pricing anomalies, such studies can offer essential implications for global investment strategies and portfolio management practices.

Furthermore, the development of methodologies to enhance the interpretability of machine learning models needs to attract enough attention. Efforts should be directed towards devising transparent and interpretable techniques that offer meaningful economic insights. By fostering a clearer understanding of the underlying drivers of asset pricing anomalies, such research can lead to more informed decision-making and enhance the acceptance of advanced machine learning techniques in practical financial scenarios.

Appendix A

Table A1. Number of neurons for all neural networks.

Model	Hidden Layer
NN1	32
NN2	32, 16
NN3	32, 16, 8
NN4	32, 16, 8, 4
NN5	32, 16, 8, 4, 2

Table A2. Hyperparameters for all prediction models and corresponding specifications.

PLS	Enet	GBRT	RF	NN

K	$\rho = 0.5$
	 $\lambda \in \left(10^{-4}, 10^{-1}\right)$
		#Depth
		 L = 1~3
		 #Trees
		 B = 1~1000
		 #Learning Rate
		 $LR \in \{0.01, 0.1\}$
			#Depth
			 L = 1~7
			 #Trees
			 B = 100~300
			 #Features
			 f = 3~50
				#L1 penalty
				 $\lambda \in \left(10^{-5}, 10^{-2}\right)$
				 #Learning Rate
				 $LR \in \{10^{-4}, 10^{-2}\}$
				 #Batch Size
				 $B \in \{64, 512, 2048, 10{,}000\}$
				 Epochs = 100
				 Patience = 5
				 Adam Para. = Default
				 Ensemble = 10

Table A3. Details on micro-firm characteristics.

No.	Acronym	Stock Characteristics	Frequency	Category
1	acc	accruals	Quarterly	ey
2	agr	asset growth	Quarterly	growth
3	am	assets-to-market	Quarterly	bpr
4	ato	asset turnover	Quarterly	ey
5	beta	market beta	Monthly	beta
6	betasq	beta squared	Monthly	beta
7	bm	book-to-market equity	Quarterly	bpr
8	capxg	capital expenditure growth	Quarterly	growth
9	cfd	cash flow-to-debt	Quarterly	lever
10	cfoa	cash flow over assets	Quarterly	ey
11	cfp	cash flow-to-price	Quarterly	bpr
12	chato	change in asset turnover	Quarterly	liq
13	chmom	change in 6-month momentum	Monthly	mom
14	cp	cash productivity	Quarterly	bpr
15	cr	current ratio	Quarterly	lever
16	crg	current ratio growth	Quarterly	lever
17	cta	cash-to-assets	Quarterly	ey
18	cto	capital turnover	Quarterly	ey
19	dbe	change in shareholders equity	Quarterly	ey
20	der	debt-to-equity ratio	Quarterly	lever
21	dlme	long term debt-to-market equity	Quarterly	lever
22	dp	dividend-to-price ratio	Quarterly	ey
23	dpia	changes in PPE and inventory-to-assets	Quarterly	bpr
24	ebit	earnings before interests and taxes	Quarterly	ey
25	eps	earning per share	Quarterly	bpr
26	ey	earnings yield	Quarterly	ey
27	gm	gross margins	Quarterly	ey
28	ia	investment-to-assets	Quarterly	ey
29	idiovol	idiosyncratic return volatility	Monthly	vol

Table A3. *Cont.*

No.	Acronym	Stock Characteristics	Frequency	Category
30	illiq	illiquidity	Monthly	liq
31	ivc	inventory change	Quarterly	size
32	lg	liability growth	Quarterly	lever
33	maxret	maximum daily return	Monthly	mom
34	mom1m	1-month momentum	Monthly	mom
35	mom6m	6-month momentum	Monthly	mom
36	mom12m	12-month momentum	Monthly	mom
37	mom36m	36-month momentum	Monthly	mom
38	mve	size	Monthly	size
39	noa	net operating assets	Quarterly	ey
40	npop	net payout over profits	Quarterly	ey
41	ocfp	operating cash flow-to-price	Quarterly	bpr
42	pacc	percent accruals	Quarterly	bpr
43	pchgm	change in gross margin—change in sales	Quarterly	growth
44	pchsaleinvt	change in sales—change in inventory	Quarterly	growth
45	pchsalerect	change in sales—change in A/R	Quarterly	ey
46	pchsalexsga	change in sales—change in SG&A	Quarterly	groth
47	prc	price	Monthly	liq
48	py	payout yield	Quarterly	bpr
49	qr	quick ratio	Quarterly	lever
50	qrg	quick ratio growth	Quarterly	lever
51	retvol	return volatility	Monthly	vol
52	rna	return on net operating assets	Quarterly	ey
53	roa	return on assets	Quarterly	ey
54	roe	return on equity	Quarterly	ey
55	roic	return on invested capital	Quarterly	ey
56	sc	sales-to-cash	Quarterly	ey
57	sg	sustainable growth	Quarterly	growth
58	si	sales-to-inventory	Quarterly	bpr
59	sp	sales-to-price	Quarterly	bpr
60	sr	sales growth	Quarterly	growth
61	std_rvol	volatility of RMB trading volume	Monthly	liq
62	std_turn	volatility of turnover	Monthly	liq
63	stdacc	accrual volatility	Quarterly	ey
64	stdcf	cash flow volatility	Quarterly	ey
65	tb	debt capacity/firm tangibility	Quarterly	lever
66	tbi	taxable income-to-book income	Quarterly	ey
67	tg	tax growth	Quarterly	bpr
68	turn	share turnover	Monthly	liq
69	z	z-score	Quarterly	ey
70	zero	zero trading days	Monthly	liq

Table A4. Details on macroeconomic variables.

No.	Acronym	Macroeconomic Variable	Frequency
1	bm	book-to-market ratio	Monthly
2	cei	consumer expectation index	Monthly
3	dy	dividend–price ratio	Monthly
4	ep	earnings–price ratio	Monthly
5	hj	economic climate index	Monthly
6	inf	inflation	Monthly
7	lvol	volatility	Monthly
8	m2gr	m2 growth rate	Monthly
9	svar	stock variance	Monthly
10	to	turnover	Monthly

References

1. Rodríguez-Moreno, M.; Peña, J.I. Systemic risk measures: The simpler the better? *J. Bank. Financ.* **2013**, *37*, 1817–1831. [CrossRef]
2. He, D.; Ho, C.Y.; Xu, L. Risk and return of online channel adoption in the banking industry. *Pac.-Basin Financ. J.* **2020**, *60*, 101268. [CrossRef]
3. Fama, E.F.; French, K.R. The capital asset pricing model: Theory and evidence. *J. Econ. Perspect.* **2004**, *18*, 25–46. [CrossRef]
4. Ang, A.; Chen, J. CAPM over the long run: 1926–2001. *J. Empir. Financ.* **2007**, *14*, 1–40. [CrossRef]
5. Baker, M.; Bradley, B.; Taliaferro, R. The low-risk anomaly: A decomposition into micro and macro effects. *Financ. Anal. J.* **2014**, *70*, 43–58. [CrossRef]
6. Schneider, P.; Wagner, C.; Zechner, J. Low-Risk Anomalies? *J. Financ.* **2020**, *75*, 2673–2718. [CrossRef]
7. Boguth, O.; Carlson, M.; Fisher, A.; Simutin, M. Horizon effects in average returns: The role of slow information diffusion. *Rev. Financ. Stud.* **2016**, *29*, 2241–2281. [CrossRef]
8. Li, Y.; Li, W. Firm-specific investor sentiment for the Chinese stock market. *Econ. Model.* **2021**, *97*, 231–246. [CrossRef]
9. Leippold, M.; Wang, Q.; Zhou, W. Machine learning in the Chinese stock market. *J. Financ. Econ.* **2022**, *145*, 64–82. [CrossRef]
10. Gu, S.; Kelly, B.; Xiu, D. Empirical asset pricing via machine learning. *Rev. Financ. Stud.* **2020**, *33*, 2223–2273. [CrossRef]
11. Bianchi, D.; Büchner, M.; Tamoni, A. Bond risk premiums with machine learning. *Rev. Financ. Stud.* **2021**, *34*, 1046–1089. [CrossRef]
12. Barbaglia, L.; Manzan, S.; Tosetti, E. Forecasting loan default in Europe with machine learning. *J. Financ. Econom.* **2023**, *21*, 569–596. [CrossRef]
13. Bollerslev, T.; Engle, R.F.; Wooldridge, J.M. A capital asset pricing model with time-varying covariances. *J. Political Econ.* **1988**, *96*, 116–131. [CrossRef]
14. Hansen, L.P.; Richard, S.F. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econom. J. Econom. Soc.* **1987**, *55*, 587–613. [CrossRef]
15. Ferson, W.E.; Siegel, A.F. Testing portfolio efficiency with conditioning information. *Rev. Financ. Stud.* **2009**, *22*, 2735–2758. [CrossRef]
16. Cederburg, S.; O'DOHERTY, M.S. Does it pay to bet against beta? On the conditional performance of the beta anomaly. *J. Financ.* **2016**, *71*, 737–774. [CrossRef]
17. Cosemans, M.; Frehen, R.; Schotman, P.C.; Bauer, R. Estimating security betas using prior information based on firm fundamentals. *Rev. Financ. Stud.* **2016**, *29*, 1072–1112. [CrossRef]
18. Mazzola, P.; Gerace, D. A comparison between a dynamic and static approach to asset management using CAPM models on the Australian securities market. *Australas. Account. Bus. Financ. J.* **2015**, *9*, 43–58. [CrossRef]
19. Chen, J.; Tindall, M.L. Constructing Equity Market–Neutral VIX Portfolios with Dynamic CAPM. *J. Altern. Investments* **2016**, *19*, 70–87. [CrossRef]
20. Hollstein, F.; Prokopczuk, M.; Wese Simen, C. The conditional Capital Asset Pricing Model revisited: Evidence from high-frequency betas. *Manag. Sci.* **2020**, *66*, 2474–2494. [CrossRef]
21. Leal, D.; Jiménez, R.; Riquelme, M.; Leiva, V. Elliptical Capital Asset Pricing Models: Formulation, Diagnostics, Case Study with Chilean Data, and Economic Rationale. *Mathematics* **2023**, *11*, 1394. [CrossRef]
22. Black, F.; Jensen, M.C.; Scholes, M. The capital asset pricing model: Some empirical tests. In *Studies in the Theory of Capital Markets*; Jensen, M.C., Ed.; Praeger: New York, NY, USA, 1972.
23. Hong, H.; Sraer, D.A. Speculative betas. *J. Financ.* **2016**, *71*, 2095–2144. [CrossRef]
24. Bali, T.G.; Brown, S.J.; Murray, S.; Tang, Y. A lottery-demand-based explanation of the beta anomaly. *J. Financ. Quant. Anal.* **2017**, *52*, 2369–2397. [CrossRef]
25. Frazzini, A.; Pedersen, L.H. Betting against beta. *J. Financ. Econ.* **2014**, *111*, 1–25. [CrossRef]
26. Asness, C.; Frazzini, A.; Gormsen, N.J.; Pedersen, L.H. Betting against correlation: Testing theories of the low-risk effect. *J. Financ. Econ.* **2020**, *135*, 629–652. [CrossRef]
27. Mohammadi, M. Prediction of α-stable GARCH and ARMA-GARCH-M models. *J. Forecast.* **2017**, *36*, 859–866. [CrossRef]
28. Ismail, M.S.; Noorani, M.S.M.; Ismail, M.; Razak, F.A.; Alias, M.A. Predicting next day direction of stock price movement using machine learning methods with persistent homology: Evidence from Kuala Lumpur Stock Exchange. *Appl. Soft Comput.* **2020**, *93*, 106422. [CrossRef]
29. Nobre, J.; Neves, R.F. Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets. *Expert Syst. Appl.* **2019**, *125*, 181–194. [CrossRef]
30. Zhang, J.; Ye, L.; Lai, Y. Stock Price Prediction Using CNN-BiLSTM-Attention Model. *Mathematics* **2023**, *11*, 1985. [CrossRef]
31. Shah, V.H. Machine learning techniques for stock prediction. In *Foundations of Machine Learning*; Spring: Berlin/Heidelberg, Germany, 2007; Volume 1, pp. 6–12.
32. Hsu, M.W.; Lessmann, S.; Sung, M.C.; Ma, T.; Johnson, J.E. Bridging the divide in financial market forecasting: Machine learners vs. financial economists. *Expert Syst. Appl.* **2016**, *61*, 215–234. [CrossRef]
33. Zhu, C.; Yin, J.; Li, Q. A stock decision support system based on DBNs. *J. Comput. Inf. Syst.* **2014**, *10*, 883–893.
34. Drobetz, W.; Otto, T. Empirical asset pricing via machine learning: Evidence from the European stock market. *J. Asset Manag.* **2021**, *22*, 507–538. [CrossRef]

35. Yuan, X.; Yuan, J.; Jiang, T.; Ain, Q.U. Integrated long-term stock selection models based on feature selection and machine learning algorithms for China stock market. *IEEE Access* **2020**, *8*, 22672–22685. [CrossRef]

36. Yu, G.; Liu, Y.; Cheng, W.; Lee, C.T. Data analysis of ESG stocks in the Chinese Stock Market based on machine learning. In Proceedings of the 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 14–16 January 2022; pp. 486–493.

37. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [CrossRef]

38. Huber, P.J. Robust Statistics. In *International Encyclopedia of Statistical Science*; Lovric, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 1248–1251. [CrossRef]

39. Prinzie, A.; Van den Poel, D. Random forests for multiclass classification: Random multinomial logit. *Expert Syst. Appl.* **2008**, *34*, 1721–1732. [CrossRef]

40. Devadoss, A.V.; Ligori, T.A.A. Stock prediction using artificial neural networks. *Int. J. Data Min. Tech. Appl.* **2013**, *2*, 283–291.

41. Yu, J.; Wen, Y.; Yang, L.; Zhao, Z.; Guo, Y.; Guo, X. Monitoring on triboelectric nanogenerator and deep learning method. *Nano Energy* **2022**, *92*, 106698. [CrossRef]

42. Masters, T. *Practical Neural Network Recipes in C++*; Morgan Kaufmann: Burlington, MA, USA, 1993.

43. Green, J.; Hand, J.R.; Zhang, X.F. The characteristics that provide independent information about average US monthly stock returns. *Rev. Financ. Stud.* **2017**, *30*, 4389–4436. [CrossRef]

44. Ma, T.; Leong, W.J.; Jiang, F. A latent factor model for the Chinese stock market. *Int. Rev. Financ. Anal.* **2023**, *87*, 102555. [CrossRef]

45. Welch, I. *The Link between Fama-French Time-Series Tests and Fama-Macbeth Cross-Sectional Tests*; SSRN: Rochester, NY, USA, 2008.

46. Camacho-Urriolagoitia, O.; López-Yáñez, I.; Villuendas-Rey, Y.; Camacho-Nieto, O.; Yáñez-Márquez, C. Dynamic Nearest Neighbor: An Improved Machine Learning Classifier and Its Application in Finances. *Appl. Sci.* **2021**, *11*, 8884. [CrossRef]

47. Lou, D. A flow-based explanation for return predictability. *Rev. Financ. Stud.* **2012**, *25*, 3457–3489. [CrossRef]

48. Jiao, W.; Lilti, J.J. Whether profitability and investment factors have additional explanatory power comparing with Fama-French Three-Factor Model: Empirical evidence on Chinese A-share stock market. *China Financ. Econ. Rev.* **2017**, *5*, 7. [CrossRef]

49. Cao, B.; Zhao, J.; Lv, Z.; Gu, Y.; Yang, P.; Halgamuge, S.K. Multiobjective evolution of fuzzy rough neural network via distributed parallelism for stock prediction. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 939–952. [CrossRef]

50. Gao, J.; Guo, H.; Xu, X. Multifactor Stock Selection Strategy Based on Machine Learning: Evidence from China. *Complexity* **2022**, *2022*, 7447229. [CrossRef]

51. Pontiff, J. Costly arbitrage and the myth of idiosyncratic risk. *J. Account. Econ.* **2006**, *42*, 35–52. [CrossRef]

52. Habib, A.; Hasan, M.M.; Jiang, H. Stock price crash risk: Review of the empirical literature. *Account. Financ.* **2018**, *58*, 211–251. [CrossRef]