# Reflections on Asset Pricing Factors: A Machine Learning-Based Perspective

Ziding Yuan *

School of Finance, Nankai University, Tianjin, China

* Corresponding Author Email: 2113136@mail.nankai.edu.cn

**Abstract.** In recent years, scholars have explored hundreds of asset pricing factors built upon the foundation of the Fama five-factor model, sparking widespread discussion in the academic community. Simultaneously, the advancement of machine learning techniques has brought innovation to asset pricing factors. This article provides an overview of typical asset pricing factors and the application of machine learning in pricing models. It begins by discussing the construction of new asset pricing factors and then delves into the innovations brought about by machine learning in asset pricing models. The diversity of factors increases the fit of asset pricing models but also presents corresponding challenges. The application of machine learning techniques addresses issues such as overfitting in pricing models, further enhancing model effectiveness. The primary value of this article lies in summarizing various perspectives on asset pricing factors in recent years, exploring the significant applications of machine learning in asset pricing models, and providing a forward-looking view on the development of asset pricing models.

**Keywords:** Asset pricing models; Machine learning; Factor zoo.

## 1. Introduction

In recent years, there has been a more comprehensive understanding of the shortcomings of the CAPM model. As a result, in-depth research has been conducted on the use of asset factors to construct models, with a focus on Fama's three-factor model and five-factor model. Dirkx and Peter conducted model testing in the German market [1]; Ragab et al. compared the Fama three-factor model with the Fama five-factor model [2]; Benali et al. analyzed the pros and cons of the Carhart four-factor model compared to the Fama three-factor model [3]. Scholars from various countries have conducted in-depth research on these pricing models.

However, scholars have not been satisfied with the five factors proposed by Fama. Through continuous research and discussion, hundreds of pricing factors have emerged. Zaremba et al. summarized liquidity factors that measure market breadth [4]; Sun and GAO proposed liquidity factors that measure depth [5]; Asness et al. constructed quality factors [6]. Meanwhile, with the emergence and gradual popularization of machine learning methods, many scholars have come up with new ideas for constructing asset pricing factors. Chen et al. reevaluated asset pricing factors using deep learning methods [7]; Stefano et al. investigated recent methodological contributions in asset pricing using factor models and machine learning [8].

This article aims to provide a literature review on asset pricing factors and machine learning. First, the article summarizes the construction and improvements of asset pricing factors and reflects on the representational significance of these factors. Next, it analyzes the interpretability and effectiveness of these factors based on the development of machine learning technology. Finally, the article discusses the feasibility and applications of such pricing models in other markets, summarizes the model's shortcomings, and provides recommendations for future improvements to the models.

# 2. Representative Factors in Asset Pricing

## 2.1. Value Factor

### 2.1.1. Explanation of the value factor

The explanations for the value factor can be divided into two main categories: systematic risk compensation and investor behavioral biases. In terms of systematic risk compensation, research suggests that financial distress risk may be an important explanation. Griffin and Lemmon discovered that, compared to other stock portfolios, the return difference between high and low book-to-market (BM) stock portfolios is twice as great in businesses with a high risk of financial trouble, suggesting a link between the value element and that risk. [9]. In other words, high BM may reflect higher financial distress risk. Peterkort and Nielsen's research also supports this view; they found that in companies without debt, future stock returns are not much impacted by BM, but they are adversely impacted in businesses with negative net assets. [10]. On the other hand, some scholars propose explanations based on investor behavioral biases. Zhang argues that companies find it difficult to reduce fixed assets during economic downturns, so high BM companies have relatively higher risk and require higher returns [11]. Some also believe that BM reflects risks related to business cycles: high BM companies are more sensitive to changes in term spreads, so higher returns can be expected. Additionally, Lewellen decomposes the predictability of BM into two parts, one related to risk and one unrelated to risk, by linearly relating BM to factor exposures in the Fama-French three-factor model, emphasizing the important role of risk in explaining BM variation [12].

From a behavioral economics perspective, people tend to assess a company's prospects by simply extending past performance. This trend leads to an overly pessimistic attitude towards companies with poor historical earnings, creating the so-called value effect. Because people tend to focus more on concrete and visible information and pay less attention to abstract information, there is a strong negative correlation between a stock's future returns and its intangible earnings. BM happens to be a good predictor of a company's intangible earnings, giving it the ability to significantly forecast a stock's future returns.

### 2.1.2. Failure and improvement of value factor

In the U.S. stock market, the performance of the value factor has been quite poor since the global financial crisis of 2008. Compared to the trend of stock indices continuously reaching new all-time highs during the same period, the performance of the value factor has lagged behind significantly, leading to widespread questioning. Even in the Chinese stock market, strategies representing the value factor, centered on the book-to-market ratio (BM), have faced considerable challenges since 2018.

Given the relative decline in the premium of the value factor over the past two decades, Fama and French conducted a detailed study of the U.S. stock market, dividing the period from 1963 to 2019 into two empirical intervals: 1963 to 1991 and 1991 to 2019. The aim of the study was to examine whether the premiums of the value factor were equal in these two sub-periods. Empirical results show that despite the significantly lower value factor premium in the latter period (roughly corresponding to the empirical interval when the HML factor was introduced), the pronounced volatility in the monthly excess returns of the value factor makes it impossible to reject the null hypothesis that the value factor premiums are equal in both periods [13].

On this topic, Asness et al. provided some valuable empirical findings that have important implications for factor investing [14]. Their research indicates that constructing composite value factors using multiple variables can lead to more robust results compared to relying solely on the book-to-market ratio (BM) to distinguish between undervalued and overvalued stocks. Using variables such as BM, earnings-to-price ratio (EP), cash flow-to-price ratio (CF/P), dividend-to-price ratio (D/P), and the past 5-year cumulative returns as examples for the U.S. stock market, the study examined the value factor. Across different historical empirical intervals, the performance of the composite value factor was more stable, highlighting the advantages of mitigating noise from a single variable by combining multiple variables. This approach contributes to enhancing the robustness of factor investing strategies.

## 2.2. Volatility Factor

### 2.2.1. Definition of volatility factor

In modern portfolio theory, Harry Markowitz was the first to propose using the standard deviation of asset return sequences to measure their price volatility. Since then, the volatility of returns has become one of the most commonly used methods to measure risk. In traditional asset pricing theory, asset volatility is one of the key parameters. When the volatility of an asset increases, the uncertainty of expected returns also rises, leading to an increase in the fair price of the asset. This provides risk compensation for investors who bear the uncertainty of returns.

Researchers generally believe that there is a negative correlation between the volatility of individual stocks and their expected future returns. This implies that investors tend to prefer holding low-volatility stocks, while high-volatility stocks do not necessarily offer corresponding risk compensation over the long term. Table 1 displays commonly used volatility factors in recent research.

**Table 1.** Common volatility factors.

| | Specific factors | Factor description |
|---|---|---|
| Volatility factor | STD | Standard deviation of daily return (total risk of individual stocks = standard deviation of yield over the last 30 days) |
| | VOLBT | Beta of trading volume (regression coefficient of the change rate of daily trading volume of individual stocks in the last 120 days and the change rate of trading volume of the Shanghai Composite Index) |
| | IVCAPM | CAPM-based trait volatility (standard deviation of the residual term of the daily return of an individual stock over the past 30 days and the regression of market factors) |
| | IVFF | Fama-French-based volatility (daily returns and market factors for individual stocks over the past 30 days, market capitalization factors, standard deviations of residual terms for valuation factor regression) |
| | IVFF-DOWN | Fama-French-based downside volatility (daily return and market factors for individual stocks over the past 30 days, market capitalization factors, less than zero standard deviation of residual items from valuation factor regression) |
| | IVFF-UP | Fama-French-based upside volatility (daily return and market factors for individual stocks over the past 30 days, market capitalization factors, greater than zero standard deviation of residual items resulting from valuation factor regression) |
| | IVCARHART | Carhart-based idiosyncratic volatility (standard deviation of the residual term from the last 30 days of individual stocks' daily returns and market factors, market capitalization factors, valuation factors, momentum factor regression) |

### 2.2.2. Test for volatility factors

Li et al. conducted Information Coefficient (IC) analysis and return analysis on volatility factors [15]. For the STD factor, the research indicates that it has a relatively low level of responsiveness to the market, hence its effectiveness is not very high. The VOLBT factor is characterized by high volatility and significant influence from trading volume, making the factor relatively unstable. However, during stock market crashes, the VOLBT factor exhibits some predictive capability.

The performance of idiosyncratic volatility factors is similar, with IVFF and IVCARHART factors showing strong trends. In the IVCAPM factor, portfolios with low exposure have relatively higher cumulative returns. In the IVFF-DOWN and IVCARHART factors, the difference in returns between long and short strategies is quite significant.

### 2.3. Liquidity Factor

#### 2.3.1. Liquidity factors that measure breadth

Díaz and Escribano's review indicates that market breadth primarily reflects trading volume and price changes, as well as the relationship between the two [16]. For instance, Zaremba et al. defined market breadth (MBR) as the ratio of the number of advancing (RS) and declining (FS) stocks in a portfolio [4]:

$$MBR_{i,t} = \frac{RS_{i,t-1} - FS_{i,t-1}}{RS_{i,t-1} + FS_{i,t-1}} \tag{1}$$

Meanwhile, Zaremba introduced results from alternative cross-sectional regression-based measures to assess market breadth, known as MBRA:

$$MBRA_{i,t} = \frac{RS_{i,t-1}}{\overline{RS}_{i,t-24:t-1}} - \frac{FS_{i,t-1}}{\overline{FS}_{i,t-24:t-1}} \tag{2}$$

$RS_{i,t}$ Means the number of advancing stocks in $t$ portfolio at time$i$, $FS_{i,t-1}$ means the number of declining stocks in $t$ portfolio at time$i$.

#### 2.3.2. Liquidity factors that measure depth

Sun and GAO conducted a dynamic analysis of various market liquidity factors before and after intraday price jumps in the CSI 300 Index futures market [5]. In this study, they focused on liquidity factors representing market depth, including bid depth, ask depth, depth, and depth imbalance (DI).

Bid depth and ask depth are measured by the number of orders at the best bid (ask) price. A higher number of orders indicates better market liquidity. In this study, the authors used the market average bid (ask) depth during the ith time period as a factor to measure market liquidity, where Ni represents the number of trades in the market during that time period:

$$BidDepth_i = \sum_{k=1}^{N_i} \frac{BidDepth_{i,k}}{N_i} \tag{3}$$

$$Ask\ Depth_i = \sum_{k=1}^{N_i} \frac{AskDepth_{i,k}}{N_i} \tag{4}$$

$$Depth_i = BidDepth_i + Ask\ Depth_i \tag{5}$$

The expression for the depth imbalance factor is:

$$DI_i = (Ask\ Depth_i - BidDepth_i)/Depth_i \tag{6}$$

By monitoring depth factors, investors can anticipate the occurrence of price jumps, thereby capturing more arbitrage opportunities. When negative price jumps occur (i.e., prices drop), depth imbalance (DI) will significantly increase, meaning that the increase in ask depth far exceeds the increase in bid depth. This phenomenon may present investors with a situation where they can trade based on depth imbalance when prices are declining.

#### 2.3.3. Liquidity factors that measure resilience

Kim et al. proposed a new method for calculating market elasticity. The author uses Beveridge Nielsen decomposition and frequency domain spectrum analysis to define market elasticity as the mean regression speed of short-term prices [17]:

$$Resiliency_{i,t} = \frac{1}{\left[\frac{D_{i,t}}{2}\right]} \sum_{k=1}^{\left[\frac{D_{i,t}}{2}\right]} \frac{2\left|\tilde{Z}_{k,i,t}\right|}{T_{k,i,t}} = \frac{1}{\left[\frac{D_{i,t}}{2}\right]} \sum_{k=1}^{\left[\frac{D_{i,t}}{2}\right]} 2\left|\tilde{Z}_{k,i,t}\right| \cdot f_{k,i,t} \tag{7}$$

Where $D$ refers to the total trading days, $i$ represents imaginary units, and $f_{k,i,t}$ represents the reciprocal of the scaled frequency component, $\tilde{Z}_{k,i,t}$ represents the distance between the peak of short-term prices within a cycle and the fundamental value at different frequencies.

## 2.4. Quality Factor

Asness et al. created a portfolio site for each nation/region and then weighted the global portfolio by dividing the portfolio for each nation by the worldwide (lag) market value [6]. He divided the equity in each nation into ten portfolios that had varying quality. The QMJ factors are formed by intersecting six value-weighted investment portfolios that are built based on size and quality, or portfolios. Asness et al. divided equities into two investment portfolios and arranged them by size based on market value at the end of each month. The QMJ factor return is the difference between the average return of two portfolios of high-quality investments and the average return of two portfolios of low-quality (junk) investments [6]:

$$QMJ = \frac{1}{2}(Small\ Quality + Big\ Quality) - \frac{1}{2}(Small\ Junk + Big\ Junk) \tag{8}$$

$$QMJ = \frac{1}{2}(Small\ Quality - Small\ Junk) + \frac{1}{2}(Big\ Quality - Big\ Junk) \tag{9}$$

Scholars from various countries have conducted empirical analyses on quality factors. The empirical results of Li and Feng show that when analyzing small market value factor (SMB) portfolio returns, the introduction of quality investment factor (QMJ) resulted in a significant improvement in risk adjusted returns [18]. At the same time, the fitting degree of the asset pricing model has also significantly improved. Moreover, when analyzing the returns of other factor combinations, incorporating the Quality Investment Factor (QMJ) also has a certain degree of improvement effect on the pricing ability of the model.

## 2.5. Green Factor

### 2.5.1. ESG premium

Pastor et al. constructed a theoretical model to describe environmental, social, and governance (ESG) pricing under market equilibrium. Specifically, they assume that investors have a preference for ESG: that is, investors gain positive utility from holding environmental companies; correspondingly, holding non environmental companies will reduce the utility of investors. In the context of this assumption, investors need to balance the expected wealth growth with the utility (negative utility) of holding green firms (brown firms) when selecting asset portfolios [19].

From this, they came to a series of conclusions. Firstly, in an equilibrium state, green firms has a negative CAPM alpha, which means that investors holding brown firms need additional compensation. In other words, on average, the profit premium of environmental companies is negative. Secondly, green firms can provide a hedging mechanism for climate and environmental risks. Therefore, when investors' concerns about environmental issues increase, the performance of Green Firms will be better. In addition, further consideration suggests that concerns about environmental issues may come from two sources, one is the investor channel, and the other is the consumer channel. The investor pathway refers to a time period when the risk of environmental issues increases beyond investors' expectations, and at this time point, green firms can provide effective risk avoidance, thus performing better. The consumer approach refers to the fact that end customers' concerns about environmental issues will encourage them to purchase more Green Firms products, thereby improving the fundamental indicators of these companies such as revenue and profits, and driving up stock prices, thereby making Green Firms perform better. Based on the above analysis, asset pricing can be carried out using a two factor model that includes market factors and ESG factors.

### 2.5.2. Setting and Derivative Issues of Green Factors

In order to measure the green degree of a company, Pastor et al. used MSCI's ESG rating data, and his definition of a company's green degree is very simple and intuitive [20]:

$$G_{i,t-1} = -\left(10 - E_{score_{i,t-1}}\right) \times \frac{E_{weight_{i,t-1}}}{100} \tag{10}$$

Among them, $E_{score_{i,t-1}}$ is the latest environmental pillar score available to company $i$ before month$t$, while $E_{weight_{i,t-1}}$ is the latest environmental pillar weight of company $i$ before month$t$, with a score between 0 and 10. The reason for having weights is because MSCI is a database that comprehensively examines the three dimensions of ESG.

Therefore, $\left(10 - E_{score_{i,t-1}}\right) \times \frac{E_{weight_{i,t-1}}}{100}$ measures how brown company $i$ is; after taking the opposite number again, it measures how far company i is from a company with a high degree of brown, that is, it measures the green degree of the enterprise. Finally, Pastor performed cross-sectional centralization on the indicators [20]:

$$g_{i,t} = G_{i,t} - \bar{G}_t \tag{11}$$

Among them, $\bar{G}_t$ is weighted average of market value for$G_{i,t}$ This means that the average market value of different companies with green degrees is 0, indicating that the market itself has no green exposure.

Finally, after Pastor ruled out the possibility that long and short factors could expose other pricing factors, he estimated the green factor using cross-sectional regression method. The green factor premium formula for period t is as follows:

$$\hat{f}_{gt} = \frac{g'_{t-1}\tilde{r}_t^e}{g'_{t-1}g_{t-1}} \tag{12}$$

Among them, $\tilde{r}_t^e = r^e - \beta_{m,t-1}r_{m,t}$ is the CAPM alpha of company $t$ ($\beta_{m,t-1}$ estimated based on data from the past 60 months).

In the past decade, green factors have performed well, and the Sharpe ratio of long short strategies is even on par with the overall stock market during historical bull markets. However, does this mean that people can reasonably expect green factors to continue to perform well in the future? Pastor et al. pointed out in their 2022 study that the high level of green premium in the past was mainly due to increasing market concerns about environmental issues [20]. However, it should be noted that investors often tend to overestimate current trends and extrapolate them to future expectations. However, the fact is often the opposite: past excellent performance is more likely to weaken the future, implying that future expected returns may be lower. In addition, the increasing trend of investors' concerns about environmental issues is also worth noting. Moreover, the theoretical model of PST 2021 suggests that the green premium should have been negative. If the future trend is reversed, the green premium may be severely impacted [19].

In addition, green factors not only help to understand the causes of green premiums, but also help explain the downturn in value factors over the past period of time: this is because value stocks contain a large number of brown firms. The excellent performance of green factors has led to relatively poor performance of these companies, thereby causing a drag on value factors. Once the influence of green factors is removed, although the excess return of the value factor relative to the market is still negative, it is no longer statistically or economically significant. For deep value investors, this is also a positive signal.

# 3. Thinking on the Establishment of Asset Pricing Model based on Machine Learning Theory

## 3.1. Contribution of Machine Learning Technology to Asset Pricing Model

In recent years, the application of machine learning method is often seen in the field of asset pricing. For example, rapach et al. used lasso method in their research to predict the return of the global stock market by using the lag returns of all countries [21].

In addition, some papers also apply the neural network method to the prediction of derivatives prices. For example, butaru et al. uses the regression tree model to predict the default and default of consumer credit cards [22]. On the other hand, Heaton et al. developed a neural network method for portfolio selection [23].

Although it appears less frequently in asset pricing literature, machine learning methods show potential in predicting different aspects of the financial market. They provide new tools and perspectives for exploring market behavior and trends from different perspectives.

Gu et al. developed a new set of standards to evaluate the machine learning method's accuracy in predicting the risk premium of the overall market and specific companies. This accuracy assessment can be considered from two main perspectives. First of all, they observed that their out of sample prediction $R^2$ was higher than that in the past literature, which showed robustness in different machine learning models [24]. Secondly, more importantly, they have shown investors huge economic benefits through the prediction of machine learning. They used the neural network model to make the out of sample Sharpe ratio of the annualized sample reach 1.35, which exceeded the performance of the regression based strategy in the previous literature.

Their research provides a new standard for the application of machine learning method in predicting market risk premium and individual stock risk premium. These methods not only perform well in prediction accuracy, but also bring significant economic benefits in actual investment.

Chen et al. introduced a non-parametric adversary estimation method in the financial field, which is interpreted as a data-based method for constructing information inspection assets [7]. The most important moment conditions may be chosen repeatedly using this approach from an endless candidate moment set. They also had a pioneering role in the development of the long-term and short-term memory network (LSTM), which can record the dynamic evolution of several macroeconomic time series in a limited number of economic states. They horizontally group a few time series into a big dimension panel, and then they use these panels to derive a nonlinear time series model. In the event of a constrained economic situation, this strategy enables them to more effectively comprehend and model the changing trend of a huge number of macroeconomic data.

## 3.2. Some Specific Application Methods of Machine Learning Technology

### 3.2.1. FFN, RNN & GAN models

Gu et al. used only feedforward networks to predict returns and labeled them FFN. It estimates the conditional mean by minimizing the sum of the mean squares of the prediction errors [24]:

$$\mu_{i,t} = \mu\left(I_t, I_{t,i}\right) \tag{13}$$

$$\hat{\mu} = min_{\mu} \frac{1}{T} \sum\nolimits_{t=1}^{T} \frac{1}{N_t} \sum\nolimits_{t=1}^{N_t} (R_{t+1,i}^{e} - \mu\left(I_t, I_{t,i}\right))^2 \tag{14}$$

Chen et al. made the following estimation in the construction of the asset pricing model [7]: first, assume the optimal weight y of the Gan network $y = \omega$, Then, assuming the optimal tool $y = g$ under the current conditions in the Gan network, the conditional average return $y = E_t\left[R_{t+1,i}^{e}\right]$ is obtained, finally, the second moment $y = E_t\left[R_{t+1,i}^{e} F_{t+1}\right]$ is estimated.

Next, they used recurrent neural network based on LSTM, and RNN has long-term and short-term memory, which can estimate hidden macroeconomic state variables. Macroeconomic time series variables usually show obvious cross-sectional dependence. It implies that various factor models can be used to capture some common information. The number of time series observations in the macroeconomic panel data that academics have investigated is roughly the same as the number of cross-section observations. This requires scholars to take some measures to reduce the dimension of cross-section data, which is also one of the important ideas of machine learning. RNN calculates the nonlinear time series correlation of vector-valued sequences in a recursive manner. Common RNN model takes the current input variable $x_t$ and previous hidden state $h_{t-1}^{RNN}$

The current state h is obtained by nonlinear transformation$h_t^{RNN}$:

$$h_t^{RNN} = h^{RNN}(x_{0,\ ...,\ } x_t) = \sigma(W_h h_{t-1}^{RNN} + W_x x_t + \omega_0) \tag{15}$$

In short, RNN first converts a large vector $x_t$ is linearly combined into a low dimensional vector to summarize the cross-section information, and then a nonlinear generalization of the autoregressive process is carried out, in which the lag variable is the transformation of the lag observation variable.

Inspired by the generation of confrontation network (GAN), Chen et al. selected the time conditions under which the two networks compete with each other and lead to the maximum pricing difference. In order to maximize loss when the SDF is provided, the conditional network's parameters are tuned to reduce unavoidable losses.

Last but not least, the criteria functions of the two lstms based on the two networks, namely $h_t$ is the hidden state that may reduce the pricing mistake, and $h_t^g$ produce test assets with the greatest erroneous pricing:

$$\{\hat{\omega}, \hat{h}_t, \hat{g}, \hat{h}_t^g\} = min_{\omega,h_t} max_{g,h_t^g} L(\omega|\hat{g}, h_t^g, h_t, I_{t,i}) \tag{16}$$

### 3.2.2. Ensemble learning

On account of the high dimension and nonlinear characteristics, how to train deep neural network is a complex task. Chen et al. used dropout technology to prevent over fitting of the model and handle with amount of parameters [7]. Dropout is a regularized form, which is usually better than the regularization effect of the traditional $l_1/l_2$. Chen et al. also adopted the adaptive learning rate strategy to realize the gradient based optimization method, so as to effectively optimize the objective function [7].

By averaging the results of multiple fitting, Chen obtained robust and reliable fitting results. It should be noted that a significant feature of neural network is that its results may be affected by the initial value. The standard method adopted by Gu et al. is to select different initial values from the optimal distribution, and then train the model separately. By averaging the results of multiple fitting, two goals are achieved [24]. The first is to reduce the impact of local suboptimal fitting on the results, and the second is to reduce model’s variance.

Chen's neural network, including the prediction model, is based on the average result of fitting more than nine different initial values. Specifically, for the $j-th$ model fitting, they calculated the optimal combination weight SDF load respectively, using $\hat{w}^{(j)}$ and$\hat{\beta}^{(j)}$. The final integrated model is the average output of the model optimized based on different initial values under the same architecture:

$$\hat{\omega} = \frac{1}{9}\sum_{j=1}^{9} \hat{w}^{(j)} \quad , \qquad \hat{\beta} = \frac{1}{9}\sum_{j=1}^{9} \hat{\beta}^{(j)} \tag{17}$$

Chen et al. divided the data into three sets, which are named training, validation and test [7]. Superparameters such as the network depth (number of layers), the number of basic functions per layer (nodes), and the number of macroeconomic States, the number of adjustment tools, and the modification of the network structure are all adjusted using the validation set. By optimizing the Sharpe ratio of SDF (risk adjustment of total return) on the validation data, researchers aim to choose the optimum configuration from all feasible superparameter combinations. With the help of this technique, the asset pricing model is guaranteed to perform better with unknowable data.

### 3.2.3. Penalty linearity

In the face of many predictive factors, simple linear models will probably fail, while complex models are prone to over fitting. The key to avoid over fitting is to reduce the number of estimated parameters. The most common machine method is to use the objective function with additional penalty to estimate more concise parameters. The most popular one is the popular "elastic net", which is in the form of:

$$\varphi(\theta;\lambda,\rho)=\lambda(1-\rho)\sum_{j=1}^{P}\left|\theta_j\right|+\frac{1}{2}\lambda\rho\sum_{j=1}^{P}\theta_j^2 \tag{18}$$

Elastic networks contain two nonnegative hyperparameters $\lambda$ and $\rho$ , two well-known regularization terms are included as special cases. $\rho=0$ Corresponds to lasso, using absolute value or $l_1$ Parameter penalty; $\rho=1$ corresponds to ridge regression, using $l_2$ parameter penalty; for intermediate values $\rho$, Elastic mesh uses a simple model through shrinkage and selection.

## 3.3. Comparison of Regression Models before and after Using Machine Learning Technology

The previous part of the article introduced a large number of machine learning methods to optimize the regression model. In order to more comprehensively understand the changes brought by machine learning technology to the regression model, I will specifically compare the data and capabilities of the regression model before and after the use of machine learning technology.

Li et al. used 12 machine learning algorithms (MLG) to build stock return prediction model and portfolio [25]. The research shows that the MLG has excellent ability to identify the complex correlation pattern between abnormal factors and excess returns. The investment strategies based on these models show better investment performance than the traditional linear algorithm and single factor method. It can be seen that the performance of machine learning algorithm obviously exceeds that of single factor model and linear regression model, and the super recognition ability brought by deep learning algorithm can improve the performance of quantitative investment.

In addition to recognition ability, machine learning technology can also improve the prediction ability of regression model. Fang and Chen used CAPM, DNN, LSTM and other models to verify their prediction ability under different prediction windows (including fixed and rolling windows). The results reveal that the machine learning model is significantly better than the traditional linear model in terms of generalization ability [26]. After testing the prediction error of linear model and machine learning model, it can be found that there are obvious differences between two models in the prediction outside the sample. The out of sample prediction of machine learning model is obviously superior to the conventional linear model, which is also the aspect that machine learning technology can continue to be studied in depth in the future.

In addition, in terms of the most important pricing power, some studies have shown that the machine learning model is better. Jiang et al. used dimensionality reduction methods such as penalty linear regression and principal component analysis to construct the characteristic factors of the capital market of China's A-share listed companies based on financial big data [27]. They compared the performance of various reduction methods and pricing factors inside and outside the sample, and explained the contribution of each factor. They found that using the principal component factor as the pricing factor was better than using the characteristic factor, and the effect of elastic network was the best in the number of reduction factors. By comparing the advantages and disadvantages of many aspects, I found that the pricing ability of machine learning model is better than that of traditional model.

To sum up, the machine learning model is obviously superior to the traditional linear model in all aspects. They can capture more complex data relationships and patterns, and have higher generalization ability, so they perform better in prediction and classification tasks. In addition, the machine learning model can effectively deal with large-scale and high-dimensional data, which is suitable for a variety of applications. Compared with the traditional linear model, the machine learning model is more flexible and can automatically learn features and patterns, so as to produce more accurate results in practical problems. Therefore, machine learning model has become an indispensable tool in modern data science and decision analysis, providing us with more in-depth and efficient insights and solutions.

## 4. Conclusion

After years of research by scholars, the defects of Fama's asset pricing model have been gradually discovered. Scholars have constructed new asset pricing factors to improve the model. However, with the continuous enrichment of factor zoo, more and more problems are exposed. At the same time, the development of machine learning technology also provides new possibilities for the future of asset pricing model. First of all, this paper takes the value factor as the representative, analyzes the advanced ideas and defects of Fama model, and gives the possible improvement measures. Secondly, this paper selects the representative factors of factor zoo, and discusses the construction process, advantages and disadvantages of these factors. Finally, combined with the newly developed machine learning technology, this paper shows its contribution and specific application in asset pricing model.

After research, I have found the problems in the factors constructed by Fama, and according to the ideas provided by Asness, I have given the improvement ideas for the establishment of composite value factors. In the process of studying the volatility factor, I listed the common related factors in recent years, tested them respectively, and found the differences brought by each factor. In the study of the liquidity factor, the liquidity factor is divided into three types: breadth, depth and elasticity, and it is found that the three have defects. In the study of quality factor, referring to the empirical results of other scholars, it is concluded that quality factor does enhance the pricing power of the model. In the study of green factors, it is verified that the environmental problems that have attracted much attention are closely related to asset pricing. In the research of machine learning technology, I summarized the brief methods of generating confrontation network, integrated learning, regression tree and other methods, and found that these technologies can improve the accuracy of asset pricing model.

In fact, machine learning technology is not very perfect at present. Although it can improve the asset pricing model, it will also bring some disadvantages. For example, pls sacrifices the accuracy of the model. At the same time, the factors in the factor zoo are not all reasonable. In the future, scholars should screen reasonable asset pricing factors based on increasingly perfect machine learning technology, and improve the factors based on machine learning technology, so as to make the asset pricing model more perfect, and its development prospect is very good.

## References

[1] Dirkx P., Peter F.J. The Fama-French five-factor model plus momentum: Evidence for the German market. Schmalenbach Business Review 2020, 72: 661-684.

[2] Ragab N. S., Abdou R. K., Sakr A. M. A comparative study between the Fama and French three-factor model and the Fama and French five-factor model: Evidence from the Egyptian stock market. International Journal of Economics and Finance, 2020, 12: 52-69.

[3] Benali M.; Lahboub K.; El Bouhadi A. Pricing ability of Carhart four-factor and Fama–French three-factor models: Empirical evidence from Morocco. International Journal of Financial Studies, 2023, 11(1): 20.

[4] Zaremba A., Szyszka A., Karathanasopoulos A., Mikutowski M. Herding for profits: Market breadth and the cross-section of global equity returns. Economic Modelling, 2020, 97: 348-364.

[5] Sun B.X., GAO Y. Market liquidity and macro announcement around intraday jumps: Evidence from Chinese stock index futures markets. Physica A: Statistical Mechanics and its Applications, 2020, 541: 123308.

[6] Asness C.S., Frazzini A., Pedersen L.H. Quality minus junk. Review of Accounting Studies, 2019, 24: 34-112.

[7] Chen L.Y., Markus P., Zhu J. Deep learning in asset pricing. arXiv, April 2019.

[8] Giglio S., Kelly B., Xiu D.C. Factor models, machine learning, and asset pricing. Annual Review of Financial Economics, 2022, 14: 337-368.

[9] Griffin J. M., Lemmon M. L. Book-to-market equity, distress risk, and stock returns. Journal of Finance, 2022, 57(5): 2317-2336.

[10] Peterkort R.F., Nielsen J.F. Is the book-to-market ratio a measure of risk? Journal of Financial Research, 2005, 28: 487-502.

[11] Zhang L. The value premium. Journal of Finance, 2005, 60(1): 67-103.

[12] Lewellen J. The time-series relations among expected return, risk, and book-to-market. Journal of Financial Economics, 1999, 54(1): 5-43.

[13] Fama E. F., French K. R. The value premium. Review of Asset Pricing Studies, 2021, 11(1): 105-121.

[14] Asness C. S., Frazzini A., Israel R., Moskowitz T. Fact, fiction, and value investing. Journal of Portfolio Management, 2015, 42(1): 34-52.

[15] Li X., Hou K.Q. Zhang C., Intangible factor and idiosyncratic volatility puzzles. Finance Research Letters, 2020, 34: 101403.

[16] Díaz A., Escribano A. Measuring the multi-faceted dimension of liquidity in financial markets: A literature review. Research in International Business and Finance, 2020, 51: 101079.

[17] Kim J., Kim Y. Transitory prices, resiliency, and the cross-section of stock returns. International Review of Financial Analysis, 2019, 63: 243-256.

[18] Li B., Feng J.J. Quality effects: Evidence from Chinese stock market. Managemnt Review, 2019, 31(3): 14-26.

[19] Pástor Ľ., Stambaugh R. F., Taylor L. A. Sustainable investing in equilibrium. Journal of Financial Economics, 2021, 142(2): 550-571.

[20] Pástor Ľ., Stambaugh R. F., Taylor L. A. Dissecting green returns. Journal of Financial Economics, 2022, 146(2): 403-424.

[21] Rapach D. E., Strauss J. K., Zhou G. F. International stock return predictability: What is the role of the United States? The Journal of Finance, 2013, 68(4): 1633-1662.

[22] Butaru F., Chen Q.Q., Clark B., et al. Risk and risk management in the credit card industry. Journal of Banking & Finance, 2016, 72: 218-239.

[23] Heaton J. B., Polson N. G., Witte J. H. Deep learning for finance: deep portfolios. Applied Stochastic Models Business and Industry, 2017, 33: 3-12.

[24] Gu S.H., Kelly B., Xiu D. Empirical asset pricing via machine learning. The Review of Financial Studies, 2020, 33(5): 2223-2273.

[25] Li B., Shao X.Y., Li Y.Y. Machine learning-driven quantitative investment research on fundamentals. China Industrial Economics, 2019, 8: 61-79.

[26] Fang Y., Chen Y.Z. Comparison of machine learning and linear models in asset yield forecasting. Statistics & Decision, 2022, 38(16), 180-183.

[27] Jiang F.W., Xue H., Zhou M. Does big data improve the pricing power of multi-factor models?-An exploration of China's A-share market based on machine learning methods. Systems Engineering-Theory & Practice, 2022, 42(8): 2037-2048.