

# ***ESG Scoring System Construction: Portfolio Investment Based on Machine Learning***

**Yulin Dong<sup>1,a,\*</sup>**

<sup>1</sup> *Shanghai University of Finance and Economics, Shanghai 200433, China*

*a. EvelinaDong@gmail.com*

*\*corresponding author*

**Abstract:** The majority of existing ESG rating systems in the Chinese market are based on categorical classification ratings, and as a result of the voluntary disclosure system, rating data provided by rating organizations is occasionally absent or delayed. This article employs natural language processing (NLP) to extract keywords such as green, clean, renewable, poverty alleviation, and moral from the financial reports of CSI 300 constituent companies, and then counts their corresponding frequencies in order to construct percentage ESG ratings that address the discontinuity, imprecision, and time lag inherent in the original ratings. This article employs a self-normalized neural network (SNN) to develop a multi factor model based on the suggested ESG ratings and then conducts sector neutral hierarchical back-testing to compare the proposed rating to the traditional ratings. The results indicate that the model generated using the ESG ratings developed in this research yields a higher rate of return than the model built using traditional ESG ratings, and the model constructed without an ESG factor. This may be because deriving ESG ratings directly from financial statements eliminates the risk of corporate falsification or whitewashing of accounts. This work adds to the body of knowledge by proposing a novel approach to constructing an ESG scoring system and incorporating it into portfolio investments to maximize returns.

**Keywords:** ESG Investment, Natural Language Processing (NLP), Neural Network, Multi Factor Investment.

## **1. Introduction**

Along with the opening of the Chinese financial market, ESG has garnered increasing attention from Chinese investors as an emerging investment concept. ESG stands for Environmental (E), Social (S), and Governance (G). These three non-economic factors are used to assess a company's performance in terms of social responsibility, establishing a sustainable development pattern, and implementing a modern management style. They serve as a benchmark for investors seeking to make socially responsible investments. A scientific and objective ESG rating not only assists investors in forecasting a company's future financial performance but also in assessing the company's sustainable development status.

ESG ratings are a set of indicators developed by professional rating agencies that are critical for investors to make investment choices. However, in the Chinese market, this is not a mandatory requirement, and regulators and industry associations only encourage companies to disclose pertinent information voluntarily. As a result, the Chinese market's existing ESG ratings are characterized by

an opaque evaluation process and a small number of covered companies.

At the moment, research on ESG in Chinese markets is scarce, and the majority of the literature focuses on its definition and strategy application, rather than on the quantitative impact of ESG factors on the stock investment process and portfolio selection. Additionally, in terms of ESG rating, the indicators published by existing rating agencies, such as Huazheng and Wind, are general categorical gradings and usually have a time lag which is intolerable in making real time transaction decisions. In this paper, I will propose a new ESG rating system that is based on a numerical grading mechanism to cope with these two problems and then backtest its effect using a self-normalized neural network. With its percentage grading scale, this numerical rating system not only compensates for the shortcomings of overly broad traditional rating systems. Additionally, it could avoid the constraints associated with companies withholding ESG information since the rating is based on ESG keywords in financial reports.

## **2. Literature Review and Hypotheses**

### **2.1. Impact of ESG on Companies' Financial Performance**

In terms of the relationship between ESG performance and revenue, Rockness [1] was the first to conduct serious research on the topic and determined that there was no correlation existed at the time. Several further researchers concurred that there is a minimal association between the two: Jaggi and Freedman [2] conducted a study in the United States and concluded that the environmental and financial performance of a business is not associated in the short term; Andreas [3] studies the influence of European firms' sustainability performance on their stock price and concludes that a corporation's social performance within a specific industry has no discernible effect on stock returns.

On the contrary, other research indicates that whether or not a company complies with its social obligation does have an effect on the financial success of a corporation. According to several publications, strong environmental performance is a possible expense to the business and will have a detrimental effect on their income level (Walley and Whitehead [4]; Gray and Shabegian [5]). These adverse effects of ESG may result in a fall in the market price impairing their anticipated future earnings. However, other studies assert that improving environmental protection and reducing pollution benefits business financial performance (Hart and Ahuja [6]; Telle [7]). In general, positive correlations exist between corporate performance and public relations in terms of shareholder value, which represents a company's profitability and competitiveness.

H1: Companies' ESG performance and compliance have a significant impact on their stock returns.

### **2.2. Criteria Selection of ESG Rating Index**

In terms of constructing performance valuation indicators for firms' environmental, social, and corporate governance performance in the Chinese market, a review of previous studies by scholars reveals that the indicators selected mostly fall into two categories. The first category is qualitative indicators. Wen et al. [8] assess the ESG performance from a stakeholder's perspective, taking into account factors such as the company's obligation to its employees, creditors, and government; Lv et al. [9] selected whether the firm had received environmental penalties as a metric for ESG rating; Wang et al. [10] selected ISO certification status, the penalty history, and the environmental expenditure rate as indicators.

The second category is quantitative indicators. Hu [11] utilized the emission fee as a proxy for environmental performance; Chen [12] evaluated corporate governance effectiveness using ten variables which include equity concentration, the number of outstanding A shares, and the proportion of independent directors; Yu et al. [13] evaluated the performance of the firm ESG on six parameters, including environmental compliance, environmental operational efficiency, and environmental

expenditures; Wang et al. [14] quantified ESG performance using the ratio of environmental investment to operational revenue. Therefore, this paper proposes the second hypothesis:

H2: Companies' ESG performance could be quantitatively measured by relevant metrics.

Due to the late start of ESG research in China, Chinese scholars' current research focuses primarily on the development of ESG indicator systems and the theoretical discussion of related disclosure, while relevant research on specific ESG evaluation systems and Chinese enterprises' ESG disclosure and financial performance remains relatively scarce. Additionally, the majority of the available research discusses just a subset of elements such as equity structure, environmental governance spending, and so forth. The proposed rating system in this paper presents a complete set of ESG rating criteria from a novel viewpoint, based on the frequency of ESG-related terms in financial reports of publicly traded corporations.

### **3. Research Design**

#### **3.1. Overview**

To begin, this article employs natural language processing to extract keyword frequencies such as green, clean, renewable, poverty alleviation, and ethical from the financial reports of CSI 300 constituent companies and then converts them to percentage ESG ratings in order to address the discontinuity, inaccuracy, and time lag associated with traditional ESG ratings. The study then uses a self-normalized neural network to construct a multifactor model based on the constructed ESG scores. Following that, a sector-neutral hierarchical backtest is conducted to compare the suggested ratings' efficacy to that of the standard ratings.

#### **3.2. Data Selection**

The analyses in this paper are carried out on portfolios chosen from CSI 300, a weighted index that consists of 300 Ashare stocks listed on the Shanghai or Shenzhen Stock Exchanges. Additionally, the time span covered by this article is 2010/12/31 to 2020/12/31, and the sample includes 29 distinct industries, including renewable energy, communication, iron, real estate, agriculture, and others. For establishing the ESG rating, this article chose the annual financial reports of CSI 300 businesses and extract the associated keywords from these PDF documents.

#### **3.3. Methodology**

##### **3.3.1. Developing ESG Rating System**

Nature Language Processing (NLP) is a powerful technique for mining documents and papers to get insights, make conclusions, and create extra approaches to enhance financial knowledge and make insightful analyses (Fisher et al. [15]). This paper employs the NLP technique to extract ESG keywords from companies' annual financial reports. These ESG related words could be divided into three parts: a) Environmental (E): Green, Climate, Environmental protection, Energy saving, Drainage, Sewage, Renewable, Clean, Carbon sink, Environment, and Environmental friendly; b) Social (S): Hours of work, Female, Gender, Poverty alleviation, Industrial revitalization, Community communication, Supply, Management, Data protection, and Employee benefits; c) Governance (G): Ethics, Code of Conduct, Independent Directors, Anti Unfair Competition, Forwards, Forward Development, Swaps, Option Incentives, Long Term Equity and etc.

Based on China ESG Development White Paper 2021[16], this paper then assigned different weights to these words according to their relative importance and managed to develop a scoring system using keyword frequencies and their corresponding weights.

### 3.3.2. Constructing Multi-Factor Model

Based on the factor model built by Nalini [17], a total of 89 factors, including financial quality, leverage, momentum, volatility, investor sentiment, and technical fundamentals, are selected in this study. In this approach, this study is trying to integrate various elements that may affect the return into the model and boost the explanatory power of this multi-factor model on the fluctuations of the market return.

Then, three multi-factor models were built: a) the model with only the basic factor pool, b) the model with the basic factor pool plus traditional ESG rating published by Wind Database, and c) the model with the basic factor pool and the ESG rating proposed above.

### 3.3.3. Training the Model

#### (1) Self-Normalizing Neural Network

Deep learning could be deemed as a multi-layered and more sophisticated version of machine learning. This paper applies self-normalizing neural networks (SNN) proposed by Klambauer [18] to calculate the optimal parameters for these models. The term "normalization" refers to the process of converting inputs to zero mean and unit variance; this is often performed as a preprocessing step. It accelerates learning and increases accuracy since it allows for comparison between the values of different characteristics. By employing SNN to construct neural networks, it could guarantee the data distribution across each layer stays stable, allowing practitioners to use neural net-based techniques for non-perception tasks such as predicting the stock returns in financial markets.

Table 1: Basic Settings of Self-Normalizing Neural Network.

SNN Settings	
Hyperparameter	Default Value
Kernel function	LeCun initialization
Activation function	SELU
Normalization	Self-normalization
Optimizer	Momentum optimization

#### (2) Time Series Cross-Validation

After deciding on a model and training strategy, the next step is to choose hyperparameters for training the SNN. One of the most effective approaches to evaluating and choosing hyperparameters is cross-validation. Cross-validation, some-times referred to as out-of-sample testing, is a resampling technique used to assess machine learning models on limited training data. Traditionally, cross-validation is performed by dividing the original data set into k subsets and then using one dataset as the test set and the remaining data as training sets. Then, by training the model for k times on a periodic basis, k MSEs may be generated. The ultimate metric CV, as shown below, equals the average of these k MSEs.

$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^k MSE_i \quad (1)$$

#### (3) Multi-Period Rolling Approach

As mentioned above, this study builds the model over a ten-year period (2010/12/31-2020/12/31) and calculates the parameters using a rolling forecast. To be more precise, this study develops the forecast by training a model across a 13-month sliding window (12 months of training data + 1 month of testing data), predicting over the following 1 month, moving 1 month, retraining the model, and

forecasting over the next 1 month, and so on. Mathematically, for the training matrix  $X$  and targets  $y$ , we have:

Epoch One

$$\begin{aligned} X_{train}^{(1)} &= \{x_{T-12}, x_{T-11}, \dots, x_{T-1}\}, y_{train}^{(1)} = \{y_{T-12}, y_{T-11}, \dots, y_{T-1}\} \\ X_{test}^{(1)} &= \{x_T\}, y_{test}^{(1)} = \{y_T\} \end{aligned} \quad (2)$$

Epoch Two

$$\begin{aligned} X_{train}^{(2)} &= \{x_{T-11}, x_{T-10}, \dots, x_T\}, y_{train}^{(2)} = \{y_{T-11}, y_{T-10}, \dots, y_T\} \\ X_{test}^{(2)} &= \{x_{T+1}\}, y_{test}^{(2)} = \{y_{T+1}\} \end{aligned} \quad (3)$$

#### 4. Results and Discussion

This section aims to provide an assessment of the three models. For back-testing, this paper stratified the data into five layers based on return forecasts for  $T+1$ . Then, it conducted back-testing and analyzed the models' effectiveness from 8 aspects - annualized return, volatility, Sharpe ratio, maximum drawdown, excess return, tracking error, information ratio (IR), and success ratio. A specific stratification involved labeling all stocks by sector, sorting them by forecasted return, and dividing each sector into five stratified portfolios. As a result, five sector-neutral stratified portfolios were established using this strategy.

##### 4.1. The Model with Basic Factor Pool

Figure 1 gives an overview of the cumulative returns of each of the five portfolios as well as the benchmark CIS 300 for the model built with only the basic factor pool. Table 2 compares the benchmark CSI 300 Index to the five portfolios plus an additional Long-Short Portfolio constructed by long Portfolio 1 and short Portfolio 5. It could be observed that Portfolio 1 yields an annualized return of 0.0869, and has the highest Sharp Ratio of 0.01489, which is consistent with the inherent nature of the stratification method this paper employed.

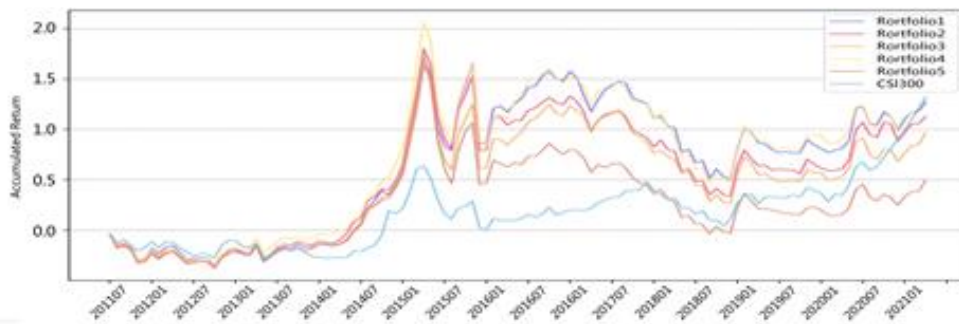


Figure 1: Accumulative Portfolios Return of Model 1 (Without ESG).

Table 2: Back-Testing results of Model 1.

Portfolios	Annualized Return	Annualized Vol	Sharp Ratio	Max Drawdown	Annualized Excess R	Ann Tracking Error	1R	Success Ratio
Portfolio1	0.086959	0.008590	0.014882	0.447126	-0.002209	0.005595	0.084791	0.542373
Portfolio2	0.079805	0.008348	0.005675	0.522000	-0.009363	0.005333	0.331389	0.525424
Portfolio3	0.071411	0.008574	0.002606	0.539106	-0.017757	0.005216	0.083513	0.508475
Portfolio4	0.089944	0.008252	0.012746	0.522697	0.000776	0.005236	0.178966	0.516949
Portfolio5	0.041696	0.008732	-0.024044	0.635360	-0.047472	0.005579	0.064199	0.466102
CS300	0.089168	0.033971	0.000000	0.389131	0.000000	0.000000		
L&S-Portfolio	0.060150	0.002101	-0.283513	0.165537	-0.029017	0.002087		

## 4.2. The Model with Wind ESG Rating

As with the previous section, Figure 2 illustrates the cumulative return for each portfolio, and Table 3 exhibits the back-testing result for the model created using features from the basic factor pool plus Wind ESG rating, the most widely used ESG rating in the Chinese market. As can be seen from these graphs, all the metrics in Table 2 have increased in comparison to Model 1, suggesting that portfolios created using ESG are indeed capable of providing better outcomes. The graph also shows that the portfolio selected following the method described above can yield a return that is up 28% higher than the benchmark CSI 300 Index.

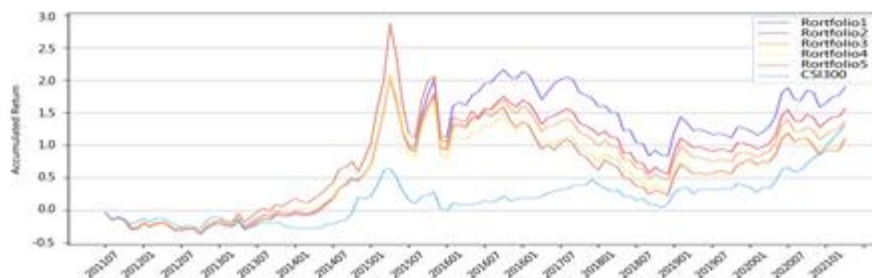


Figure 2: Accumulative Portfolios Return of Model 2 (With Wind ESG).



Table 3: Back-Testing results of Model 2(With Wind ESG).

Portfolios	Annualized Return	Annualized Vol	Sharp Ratio	Max Drawdown	Annualized Excess R	Ann Tracking Error	1R	Success Ratio
Portfolio1	0.114 291	0.008 506	0.038 937	0.420 292	0.025 387	0.005 609	0.109 008	0.559 332
Portfolio2	0.100 340	0.008 431	0.027 310	0.488 188	0.011 436	0.005 524	0.247 875	0.516 949
Portfolio3	0.092 026	0.008 584	0.021 307	0.523 374	0.003 121	0.005 700	0.198 943	0.5
Portfolio4	0.080 560	0.008 749	0.010 416	0.578 460	-0.008 344	0.006 027	0.329 93	0.508 475
Portfolio5	0.077 970	0.009 595	0.015 657	0.681 626	-0.010 935	0.007 285	0.276 806	0.466 102
CS300	0.088 904	0.003 775	0.000 000	0.393 569	0.000 000	0.000 000		
L&S- Portfolio	0.061 959	0.003 140	- 0.239 702	0.841 744	-0.026 946	0.003 137		

#### 4.3. The Model with Constructed ESG Rating

Graphs below illustrate the return and other financial indicators of the model built using proposed ESG rating data. As demonstrated by the following results, the model trained using the proposed ESG criterion exceeds the two preceding models across all relevant metrics, including the return, Sharpe ratio, and success ratio. Among these, portfolio1's Sharpe ratio is 32% greater than model2's and 235.9% greater than model1, indicating that the continuous and non-defective ESG scoring criteria developed in this research outperform traditional ESG scoring criteria in the real world.

This could be because the continuous ESG rating, which is calculated on a percentage basis, provides a more accurate depiction of a company's ESG performance. Additionally, deriving ESG ratings directly from financial statements eliminates the risk of corporate falsification or whitewashing of accounts, which might be a significant disadvantage for the existing ESG rating system, as corporations are only required to report ESG data voluntarily.

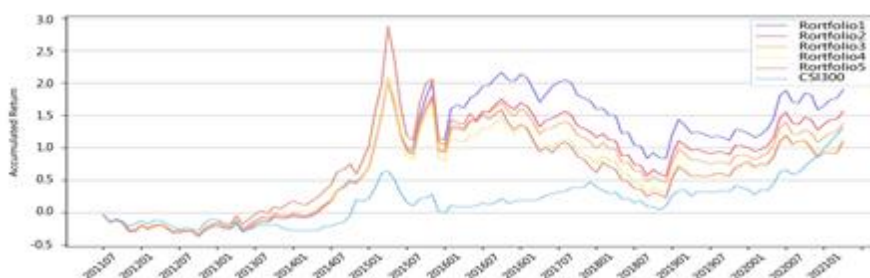


Figure 3: Accumulative Portfolios Return of Model 3 (With Constructed ESG).

Table 4: Back-Testing results of Model 3.

Portfolios	Annualized Return	Annualized Vol	Sharp Ratio	Max Drawdown	Annualized Excess R	Ann Tracking Error	1R	Success Ratio
Portfolio1	0.126 290	0.008 556	0.050 198	0.421 147	0.037 788	0.005 643	0.254 13	0.542 373
Portfolio2	0.099 614	0.008 520	0.026 438	0.502 892	0.011 111	0.005 647	0.276 295	0.508 475
Portfolio3	0.095 847	0.008 430	0.023 480	0.501 782	0.007 344	0.005 585	0.208 598	0.516 949
Portfolio4	0.066 559	0.008 774	-0.001 309	0.612 784	-0.021 943	0.005 995	0.241 122	0.491 525
Portfolio5	0.076 153	0.009 570	0.013 718	0.660 990	-0.012 349	0.007 251	0.184 393	0.457 627
CS300	0.088 503	0.033 554	0.000 000	0.394 484	0.000 000	0.000 000		
L&S-Portfolio	0.008 585	0.002 968	-0.214 495	0.522 651	-0.006 918	0.002 964		

## 5. Conclusion

The crucial finding in this article is that constructing a new ESG scoring system using natural language processing and a self-normalized neural network can generate higher returns when conducting portfolio investing in the Chinese stock market.

This paper investigates the relationship between ESG ratings and portfolio returns by extracting ESG keywords using NLP, building multi-factor models with SNN, and then conducting sector-neutral back-testing to compare the results of these three models. Empirical research has established that the ESG factor generally improves the effectiveness of the model - models incorporating the ESG factor provide higher returns than the model that does not. What's more, the ESG factor established in this work dwarfs the typical ESG factor in terms of returns. With the rating system outlined in this paper, investors will have access to more accurate and timely ESG ratings for all publicly traded companies, enabling them to make sound investment decisions.

## References

- [1] Rockness J, Schlachter P, Rockness H: Hazardous waste disposal, corporate disclosure, and financial performance in the chemical industry. *Advances In Public Interest Accounting*, 1.1986(1041-7060), 167-191(1986).
- [2] Jaggi, B., & Freedman, M.: An examination of the impact of pollution performance on economic and market performance: pulp and paper firms. *Journal of business finance & accounting*, 19(5), 697-713 (1992).
- [3] Ziegler, A., Schröder, M., & Rennings, K.: The effect of environmental and social performance on the stock performance of European corporations. *Environmental and Resource Economics*, 37(4), 661-680 (2007).
- [4] Walley, N., & Whitehead, B.: It's not easy being green. *Reader in Business and the Environment*, 36(81), 4 (1994).
- [5] Gray, R., Kouhy, R., & Lavers, S.: Corporate social and environmental reporting: a review of the literature and a longitudinal study of UK disclosure. *Accounting, Auditing & Accountability Journal* (1995).
- [6] Hart, S. L., & Ahuja, G.: Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. *Business strategy and the Environment*, 5(1), 30-37 (1996).
- [7] Telle, K.: "It pays to be green"—a premature conclusion. *Environmental and Resource Economics*, 35(3), 195-220 (2006).
- [8] Wen, S.B., & Fang, W.: Qiyie Shehui zeren yu Caiwu jixiao guanxi de shizheng yanjiu: Liyi xiangguanzhe shijiao de mianban shuju fenxi [An empirical study of the relationship between corporate social responsibility and financial performance: a panel data analysis from stakeholder perspectives]. *Zhongguo gongye jingji*, 10, 150-160 (2008).
- [9] Lv, J., & Jiao, S.Y.: Huanjing pilu, Huanjing jixiao he caiwu jixiao guanxi de shizheng yanjiu [An empirical study



*of the relationship between environmental disclosure, performance, and financial performance]. Shanxi caijing daxue xuebao, 1, 109-116 (2011).*

- [10] Wang, B., & Zhao, Y.P.: *Qiye huanjing jixiao yu caiwu jixiao xiangguanxing shizheng yanjiu [An empirical study on the correlation between corporate environmental and financial performance: Based on panel data of listed companies from 2006 to 2010]. Caikuai tongxun, 36, 50-52 (2012).*
- [11] Hu, Q.Y.: *Shangshi gongsi huanjing jixiao yu caiwu jixiao de xiangguanxing yanjiu [Study on the correlation between environmental and financial performance of listed companies]. Zhongguo rengou ziyuan yu huanjing, 6, 23-32 (2012).*
- [12] Ye, C.G., & Qiu, L., & Zhang, L.J.: *Gongsi zhili jieou, neibu kongzhi zhiliang yu qiye caiwu jixiao [Corporate Governance Structure, Internal Control Quality and Corporate Financial Performance]. Shengji yanjiu, 2, 104-112 (2016).*
- [13] Yu, H.X., & Wang, Q.H.: *Shiyou qiye huanjing jixiao pingjia zhibiao tixi goujian [Construction of environmental performance evaluation index system for oil enterprises]. Shidai jing-mao, 27, 16-17 (2020).*
- [14] Wang, W.Y., & Jia, Z.X.: *Huanjing jixiao yu caiwu jixiao guanxi hangye bijiao yanjiu [Industry Study on the Relationship between Environmental and Financial Performance: Based on Data from Listed Companies in the Paper, Chemical, and Extractive Industries]. Kuaiji zhiyou, 11, 79-84 (2020).*
- [15] Fisher, I. E., Garnsey, M. R., & Hughes, M. E.: *Natural language processing in accounting, auditing and finance: A synthesis of the literature with a roadmap for future research. Intelligent Systems in Accounting, Finance and Management, 23(3), 157-214 (2016).*
- [16] China Alliance of Social Value Investment (CASVI), China Asset Management Co. (China AMC): *White Paper on ESG Development and Innovation in China 2021 (2021).*
- [17] Nalini, R.: *Optimal Portfolio construction using Sharpe's Single Index Model-A study of selected stocks from BSE. International Journal of Advanced Research in Management and Social Sciences, 3(12), 72-93 (2014).*
- [18] Klambauer, G., Unterthiner, T., Mayr, A., & Hochreiter, S.: *Self-normalizing neural networks. Advances in neural information processing systems, 30 (2017).*