Factor Analysis by Neural Network in China's A-share Market

Key Highlights

- Neural network model is widely used in area of information extraction, analysis and forecast. Its power has been verified in multiple aspects. This research is based on fully connected neural network and total 44 factors, implementing yearly rolling window to divide in-sample data into 8 groups and forecast future return by ReLU function and RMSProp optimizer.
- The results show that fully connected neural network possesses relative good ability for return forecast and the construction of long and short hedge portfolio can finally achieve 17.26% annual return. This strategy can be verified in majority sectors and industries.
- This research proves that the factors can provide extra information for stock return, which means the unstructured new factors generated by the neural network successfully account for part of the return which cannot be interpreted by traditional factors. New aspects and toolboxes can be implemented in market analysis and portfolio management.

Contents

Introduction

Artificial Neural Network (ANN) is pioneered out of a simulation of biological neural networks. The transmission of information from one layer to another is carried out by an adaptive nonlinear dynamic network system constructed by interconnecting a large number of neurons. Based on different connection types, layers and quantities, different ANN models are used by different industries.

The ANN model has been greatly enriched. Study of Hinton and Salakhutdinov (2006) has shown the feature learning ability of ANN and triggered a boom of ANN deep learning in the past decade.

The neural network simulates the transmission and processing of information between different neurons by organisms. It is a process of dimensionality reduction, information extraction and data integration. Therefore, the ANN model derived from image processing has been widely used in meteorology, medicine, transportation and other fields. Meanwhile, the high degree of freedom of the approximation function and the strong self-learning advantages also make relevant ANN model a hotspot for data mining and analysis in financial area.

This research generates a fully-connected neural network model by extracting multiple factors from market information and then tests the effect of the factors. The following part briefly introduces the methodology of fully connected neural network factor model.

Fully-Connected Neural Network and Factor Calculation

1. Fully-Connected Neural Network

This part will be discussed based on the research by LeCun et al. (2015). As mentioned above, the neural network model simulates the information processing mechanism, processes the multidimensional information (such as sensation) into signals through the interconnection of neurons at each layer and nodes. Finally, the network passes the information from the source to the processor. This workflow can be abstracted as an end-to-end information processing structure.

Let us simply observe the transit structure of a neural network:

[Figure 1: Neural Network Simple Structure]

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \\ \alpha \end{bmatrix} \dashrightarrow F_x \dashrightarrow Y = h(F_x)$$

Where x_n indicates the input information matrix, α is the offset (intercept in regression model). After weighting procedure with $f(\cdot)$, a new information matrix F_x is generated (matrix dimension depends on the weighting function), and the dimension of this matrix should be significantly lower than the dimension of original information matrix.

The new information matrix needs to be activated by the activation function h(·) and converted into information that can be processed, as the input of next layer or the output of final layer. This step simulates the process of activating neurons in the biological nervous system, that is, by setting a certain threshold to activate the nerve meta, which passes information to the next layer.

We can gradually reduce the dimension of data by designing multiple nodes and repeating the process of weighting. It can be shown as follow:

$$X^L = X^{L-1}W^{L-1} + A^{L-1}$$

Where L represents the Lth layer, X^L is the information extracted in Lth layer, W^L is the weights information in Lth layer and A^L is the offset in Lth layer. When the information of this layer is weighted based on all the information of the previous layer, it is a so-called fully-connected neural network. In addition, the foregoing process uses a linear transformation, but usually the relationship between the information is nonlinear, so the correlation transformation function tends to adopt a nonlinear form.

Here we use the backward propagation to minimize the loss function. In the process of backward propagation, it is essential to adjust the weights that were previously set unreasonably. First, since our goal is to minimize the loss function (EE(·)), we first use the loss function to derive the weight to determine the marginal influence of the weight change on the error (ie the magnitude of the loss), and then use the margin as weights. For example, the adjusted weight in layer 4 should be:

$$\mathrm{w}_4' = \mathrm{w}_4 - \beta \frac{\partial E(\cdot)}{\partial w_4} = \mathrm{w}_4 - \beta \frac{\partial E(\cdot)}{\partial h(F_{x_4})} \cdot \frac{\partial h(F_{x_4})}{\partial x_4} \cdot \frac{\partial x_4}{\partial w_4}$$

Where β is the learning rate.

2. Neural Network to Factors

Various information of stocks (fundamental, macro, price, etc.) is reflected as factors and their exposure values. The information of a large number of factors is purified, extracted and synthesized, which is actually a dimension reduction process.

Taking Figure 2 as an example, X is the factor value of a stock at time t. If we consider the case of n periods, we actually construct a matrix of factor values (ie, a panel data form). This matrix shows in this case, the stock has multiple factors in multiple periods. In theory, if we choose a long enough period and enough factors, then the information in this matrix should show the full aspects of this stock. This is similar to the concept in image processing, in which case this matrix is also a "factor photo" of the stock.

[Figure 2: Factor Photo Matrix]

$\mathbf{A_t}$	$\mathbf{A_{t-1}}$	…	$\mathbf{A_{t-n}}$
$\mathbf{B_t}$	$\mathbf{B_{t-1}}$	…	$\mathbf{B_{t-n}}$
…	…	…	…
$\mathbf{X_t}$	$\mathbf{X_{t-1}}$	…	$\mathbf{X_{t-n}}$

When the above matrix is constructed for all stocks of A-shares, the "factor photos" of more than 3,000 stocks can be obtained, and as the initial input of each stock, a fully-connected neural network system can be generated by referring to the foregoing method.

In the generation of factors, the target stocks must first be screened. We labeled stocks (which are not suspended in following two months) based on their return rate. The top 50% of the stocks are positive (y = 1) and the last 50% are negative (y = 0).

The label on the one hand contributes to the reasonable settings of the training set and the associated loss function, and on the other hand to the easy calculation of the final factor. Based on this label form, the factor values obtained by the model actually represents the forecast of stock returns in the next month. The higher the factor value, the higher the stock will earn in the next month.

Subsequently, the factors that need to be input are determined, which means that a factor pool for generating a fully-connected neural network factor is determined.

In this paper, a total of 44 factors of 11 categories were selected as the pool of candidate factors for fully-connected neural network (see Figure 3). The classification system of mainstream factors was basically covered, and the various aspects of stocks data were comprehensively investigated. The sufficient content of the information improves the probability of successful predict in the stock return ranking next month. In the meantime, we gathered every single stock’s factor exposures of current month and previous 4 months (total 5 months) as a 44×5 two-dimensional matrix.

[Figure 3: Factor Pool]

Factor Type	Factor Name	Factor Type	Factor Name
Value	BP Ratio	Analytic Forecast	180 Days Earning Organization Numbers
	Quarterly BP Ratio		180 Days Rating Organization Numbers
	PS Ratio		3-Month EPS Dispersion Index
	Quarterly PS Ratio		3-Month EPS Change
	EP Ratio	Volatility	30-Day Index China beta
	4-Quarters EP		60-Day Index China beta
	EBITDA/Price		90-Day Index China beta
	Enterprise Multiple		250-Day Index China beta
	Quarterly Enterprise Multiple		30-Day Index China Residual volatility
Growth	4-Quarter Income Growth		60-Day Index China Residual volatility
	Annual Earnings Growth		90-Day Index China Residual volatility
	4-Quarter Earnings Growth		250-Day Index China Residual volatility
Earnings	Quarterly Gross Profit Margin		30-Day Volatility
	Total Asset Growth		60-Day Volatility
	Net Asset Growth		90-Day Volatility
	Gross Profit Margin		250-Day Volatility
Debt	Quarterly Liquidity Ratio	Technical Indicators	1-Month Reversion
	Quarterly Debt to Equity Ratio		14-Day RSI
Momentum	3-Month Adjusted Momentum	Market Value	A-Share Total Value
	6-Month Adjusted Momentum	Turnover	30-Day Turnover
	9-Month Adjusted Momentum		60-Day Turnover
	12-Month Adjusted Momentum		90-Day Turnover

When calculating the factors by scrolling window, the in-sample data and out-sample data are divided into eight groups according to the timeline. For example, when predicting the data for the period from 2012.05 to 2013.04, a total of 72 months of data from 2006.05 to 2011.04 will be combined as an in-sample training data set.

After the above factors are selected into the factor pool, the factor exposure values need to be pretreated. First, the MAD method is used to process the extreme values. We replaced the missing values with the cross-sectional mean of the factor in the Shenwan first-level industry classification.

Then regress the exposure values against the industry dummy variable (Shenwan first-level industry) and the logarithm of market value, and regard the residual value as the market’s industry neutral factor value. In order to ensure the comparability between different factors, the factors are also standardized. For details of the above steps, please refer to Appendix A.

The factor construction can be carried out according to the aforementioned fully-connected neural network. In this paper, three fully-connected layers are set up in the model, with respectively 128, 64 and 2 neurons. This paper selects the ReLU function as activation function and classification cross entropy function (Categorical Cross-entropy) as loss function. (More details are Appendix B) The learning rate is set to be 0.00001, and the RMSProp optimizer is implemented to prevent gradient summation from being too large.

The model follows these steps when adjusting parameters:

- First, randomly choose 10% of the sample as validation set and observe the performance of the convolutional neural network on the validation set.
- Then stop training when the accuracy stops improving in the next three rounds of training and extract the optimized parameters.
- Afterwards, we can forecast next year stock return by neural network.

Factor Performance Analysis

1. Factor Data Descriptive statistics

This section gives the descriptive statistics of the fully-connected neural network factors. In Figure 4, it can be found that the factors distributions differ in different industries. Factor values in financial and heavy industry are smaller than the average value, while light industry, technology enterprises tend to have higher factor values. That is, based on the results of a fully connected neural network, it is more likely that stocks in sectors such as finance and heavy industry are less likely to rank higher earnings in the next month. Of course, in general, the difference of factor values between industries is very small, and as a consequence, the analysis based on the industry may not be very significant. Therefore, the latter part will be based on the individual stock factor exposure for the back test.

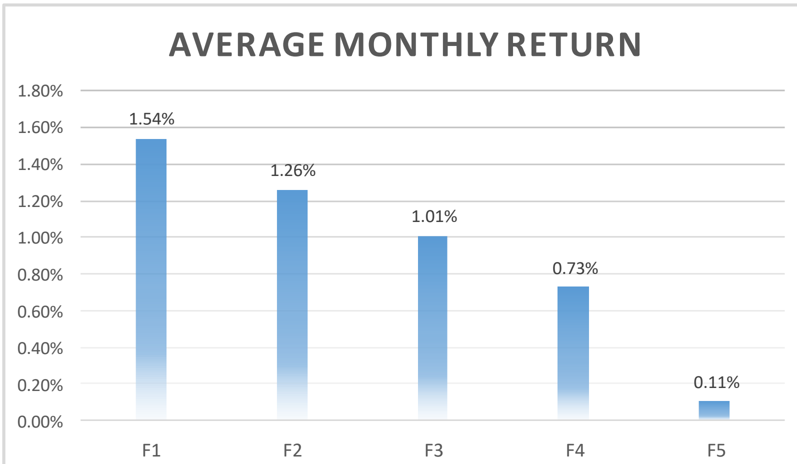
[Figure 4: Factor Descriptive Statistics]

Industry	Mean	Min	Median	Max
Overall	0.4965	0.0526	0.5007	0.8526
Banking	0.4655	0.0526	0.4736	0.7756
Finance (Excluding Banking)	0.4719	0.0854	0.4749	0.812
Mining and Quarrying	0.4754	0.0884	0.4791	0.8005
Steel	0.4844	0.1561	0.4845	0.7977
Services	0.4922	0.1861	0.497	0.7719
Commercial Trades	0.4935	0.1269	0.4984	0.7815
Transportation	0.4936	0.0721	0.4977	0.8329
Household Appliance	0.4938	0.1333	0.4972	0.8271
Architectural decoration	0.4939	0.1752	0.4983	0.8422
Automotive	0.4941	0.1247	0.4998	0.8121
Construction Materials	0.4941	0.1365	0.4998	0.7771
Food and Beverage	0.4951	0.1824	0.4963	0.8523
Real Estate	0.4959	0.097	0.4999	0.8408
Chemical Industry	0.496	0.1476	0.501	0.8471
Comprehensive	0.496	0.1951	0.4961	0.8105
public utilities	0.4964	0.115	0.4999	0.8276
Defence Force Workers	0.4964	0.1582	0.4997	0.808
Nonferrous Metals	0.4977	0.1224	0.5001	0.8412
Agriculture, Forestry and Fishing	0.4981	0.1485	0.5036	0.7982
Light Manufacturing	0.499	0.1022	0.5049	0.823
Communications	0.4992	0.0876	0.5029	0.8375
Electrical Equipment	0.5	0.0971	0.5056	0.8111
Pharmaceuticas	0.5001	0.1143	0.5036	0.8038
Computer	0.5005	0.0709	0.5037	0.8526
Machinery and Equipment	0.5006	0.1032	0.5062	0.8197
Media	0.5007	0.0966	0.5034	0.818
Electronics	0.5011	0.1489	0.5057	0.844
Textiles and Clothing	0.5012	0.1255	0.5072	0.7649

2. Factor Back Testing

In this part, according to the factor exposure values, the A-shares are divided into 5 groups, F1 group is the highest 20% factor exposure, correspondingly F5 group is the smallest 20% factor exposure. In back-testing, the portfolio position was changed at the end of each month, and the individual stocks in each portfolio were formed with equal weights. The back testing time period is from June 2011 to present. The average monthly returns for each group are shown in Figure 5:

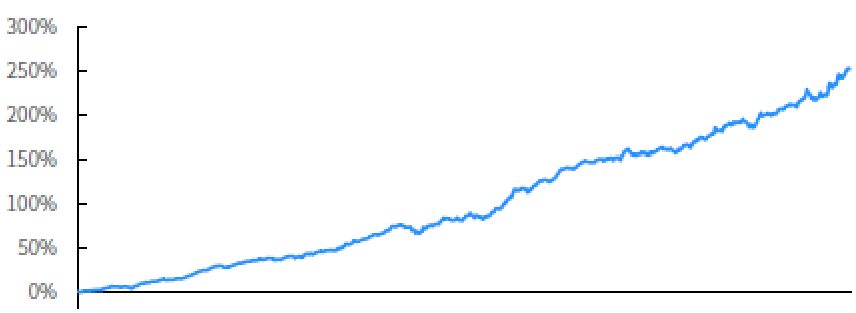
[Figure 5: Average Monthly Return of Factor Groups]



From the results, the factor presents an obvious monotonousness, that is, the higher the factor value, the higher the future stock return, which confirms that the method of this paper is relatively accurate in prediction. In addition, from the situation of Figure 6, the portfolio constructed by taking the long F1 and shorting F5 method can achieve an annualized yield of 17.26%, and from a graphical point of view, there is no significant sharp drawdown during the period (maximum drawdown of 6.02%), which means that the factor can bring relatively stable returns through this period.

[Figure 6: F1-F5 Hedging Portfolio Accumulative Return]

Time Line: 2011.06.01-2017.12.26



3. Industry-level Back Testing

In the back testing in the previous part, it can be found that the fully-connected neural network factors show good stock selection ability in the whole A-share pool. Although we have eliminated the possible impacts of industry in calculation, the stock selection capacity may still differ between industries. This article based on Shenwan industry classification, retest the stock selection capacity among industries. The back test method is consistent with previous A-share individual factor back test. The results are shown in Figure 7.

[Figure 7: Group Back Test of Industries]

Industry	F1-Fn	F1	F2	F3	F4	F5
Computer	24.96%	24.73%	25.10%	29.31%	27.69%	-0.23%
Construction Materials	16.15%	18.69%	8.83%	5.45%	-3.17%	2.53%
Steel	14.92%	17.92%	3.12%	12.22%	1.05%	3.00%
Transportation	14.85%	19.53%	17.41%	12.37%	10.80%	4.67%
Electrical Equipment	12.84%	22.86%	24.07%	16.12%	14.17%	10.03%
Nonferrous Metals	11.94%	14.55%	5.82%	1.49%	4.31%	2.60%
Machinery and Equipment	11.87%	21.12%	13.80%	9.61%	10.90%	9.25%
Chemical Industry	10.35%	17.32%	16.49%	14.13%	11.55%	6.97%
Banking	8.87%	10.00%	9.78%	4.35%	4.58%	1.13%
Pharmaceuticas	8.50%	18.47%	16.11%	16.78%	16.11%	9.98%
Food and Beverage	8.41%	12.75%	5.40%	9.39%	4.61%	4.34%
Light Manufacturing	6.26%	5.25%	4.14%	12.58%	13.19%	-1.01%
Commercial Trades	4.16%	11.03%	12.51%	7.47%	9.54%	6.87%
Finance (Excluding Banking)	3.77%	8.67%	6.30%	3.25%	3.62%	4.90%
Real Estate	2.88%	17.06%	15.98%	12.37%	11.24%	14.17%
Mining and Quarrying	2.77%	9.09%	-3.77%	-2.17%	6.50%	6.31%
Textiles and Clothing	2.33%	12.32%	8.55%	5.85%	7.53%	10.00%
Electronics	1.65%	15.61%	16.35%	18.06%	21.97%	13.96%
Architectural decoration	1.26%	9.71%	15.40%	10.71%	8.23%	8.45%
Services	0.85%	9.21%	12.15%	-2.80%	13.51%	8.36%
public utilities	0.10%	13.45%	11.86%	8.62%	13.11%	13.36%
Media	-1.95%	6.68%	22.41%	14.40%	3.80%	8.63%
Comprehensive	-2.36%	-1.74%	3.37%	2.05%	-0.99%	0.62%
Household Appliance	-3.00%	17.23%	20.94%	18.95%	14.41%	20.23%
Communications	-4.24%	15.64%	3.12%	24.31%	17.06%	19.88%
Defence Force Workers	-5.31%	15.61%	17.63%	9.21%	19.63%	20.92%
Automotive	-5.39%	1.06%	2.14%	-4.95%	7.31%	6.45%
Agriculture, Forestry and Fishing	-12.13%	-0.60%	13.03%	9.84%	-4.09%	11.53%

From the results of the industry back testing, in 29 industries, 22 industries have shown the high factor exposure value portfolio return will be higher than the low factor exposure value portfolio return. In majority industries, the factor could still show an apparent monotonousness. On the other hand, it can be found that the industry average factor exposure value is not closely related to its factor monotonousness. For example, the media and communication industry with higher average factor exposure, in fact, perform relatively poor among the all industries; while industries with low average factor exposure values, such as steel and banking, have a very significant advantage.

The Robustness Test of Factors

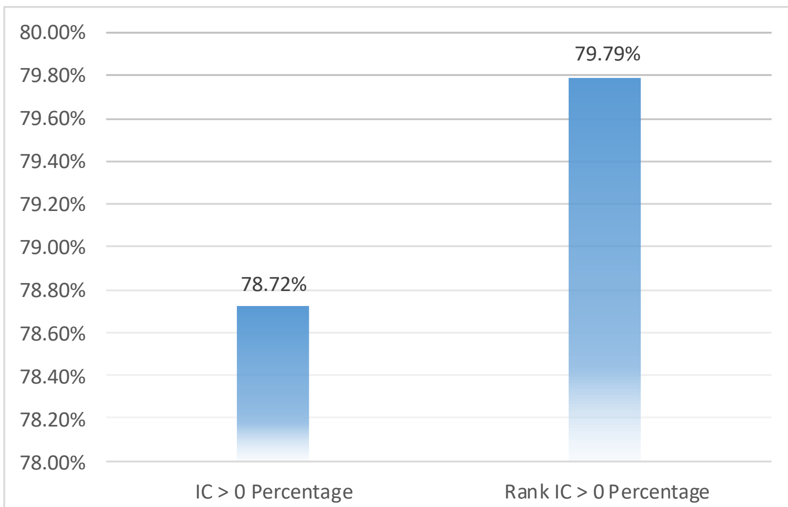
In the previous analysis, this paper has found that the factors constructed by the fully-connected neural network have great stock-picking ability. In the following part we will analyze the robustness of the factor stock-picking ability from the statistical method.

1. Factor IC Test

The IC test of the factor examines the coefficient of the factor exposure value and the next period return. The higher the absolute value of the IC value means the greater correlation between the factor exposure and the future benefit. The detailed explanation can be found in Appendix C. The results of the IC test are shown in Figure 8:

[Figure 8: Factor IC Table]

IC Mean	IC Mean/IC Std	Rank IC Mean	Rank IC Mean/Rank IC Std
0.0364	0.7527	0.0438	0.8521



From the results, the correlation coefficient between factor exposure and future earnings is relatively high. Both IC or Rank IC have nearly 80% mean value which is greater than 0. It can be concluded that from the prediction ability, the fully connected neural network factor has a relatively good effect.

2. Factor Regression Test

This part uses the individual stock return to regress on the fully connected neural network factors. The regression coefficient is actually the rate of return of the factor ---- and we need to examine whether the factor yield is significantly different from 0. If it can be investigated by significance, it means that statistically, the factor can indeed provide a credible return. Generally speaking, the method of judging the significance of the regression coefficient is the t-test.

[Figure 9: Regression Test Statistics]

Factor Return Mean	t-Value Mean	|t| Mean	|t|>2 Percentage
0.0400	1.7754	2.4117	52.13%

From the results, the mean of t-value is at a significant level of 10% to reject the null hypothesis. The proportion of the absolute value of t is greater than 2 also reached 52.13%, which indicates that the forecast ability of factors can be verified.

3. The Test of Factor's Ability to Independently Explain the Benefit

This part of the test mainly examines whether the fully connected neural network factors provide additional information based on the Fama-French three-factor model, and explain the stock's return. If the fully connected neural network factor can pass the correlation test, it means in addition to the traditional pricing model, factors have a certain ability to independently explain stock returns and risks. The specific meaning of the relevant test is detailed in Appendix C. The specific statistical results are shown in Figure 10.

[Figure 10: Factor Independence Test Statistic]

Statistics	Fama-French Model	Fully-Connected Model
GRS	2.0254	2.0222
P-Value of GRS	0.0115	0.0119
$\|\alpha\|$ Mean	0.0046	0.0040
R^2 Mean	0.9519	0.9533

The results in the third column of Figure 10 show that after adding the fully-connected neural network factor, the relevant indicators of the model have been improved and the return that cannot be explained by Fama-French have declined. Simultaneously the fitness of the model and its estimation efficiency has been improved. This result suggests that the factors provide information on the interpretation of stock returns and this interpretation information is not covered by traditional mature models.

Conclusion

This research constructs the fully-connected neural network factor based on China A-shares. These factors are extracted from 44 traditional factors and predicts the future stock return. The back test results of the factors show that the fully-connected neural network factors exhibit significant monotonicity and the stocks with higher factor exposures have higher return in the next month, which may indeed be obtained in the future. In addition, these factors provide additional information based on a mature asset pricing model therefore provide new opportunities for investors to make profit.

Based on fully-connected neural network method and by constructing factors with non-traditional structure model, this paper provides us with a new way of thinking about mass data processing and data mining in the financial market. The performance of factors confirms that the new method does show investors a powerful tool and new perspective to recognize the A-share market opportunities. Our team will continue to launch research based on similar deep learning methods.

Reference

[1] Fama, E., French, K., “Common Risk Factors in the Returns on Stocks and Bonds”, Journal of Financial Economics, 1993, 33(1):3-56.

[2] Hinton, G., Salakhutdinov, R., “Reducing the Dimensionality of Data with Neural Networks”, Science, 2006, 313(5786): 504-507.

[3] LeCun, Y., Bengio, Y., and Hinton, G., “Deep Learning”, Nature, 2015, 521(7553): 436-444.

Appendix

1. Appendix A

Extreme Values

In this paper, the extreme values are processed by the MAD method (Median Absolute Deviation, i.e. absolute deviation method), which is mainly used to measure the robustness of the sample data and can effectively overcome the possible influence of sample dispersion.

For factor k, the factor exposure β_k in all stocks can be processed based on the following formula:

$$\text{MAD}(\beta_k)=meidan(\left|\beta_{i,k}-median(\beta_k)\right|)$$

Where $\beta_{i,k}$ represents the factor k exposure of stock i, $\mathrm{i}=1,2,3\ldots\ldots\mathrm{n}$, where n is the number of stocks in A-share.

After obtaining the MAD value, it is necessary to select the factor exposure based on MAD value to determine the normal exposure range of the factor. Here 5 times the MAD value is selected as the normal exposure upper and lower limit of the factor, that is, the value falling in the range of $[meidan(\beta_k)-5\times MAD(\beta_k), meidan(\beta_k)+5\times MAD(\beta_k)]$. Finally, replace the values preceding the limits with corresponding upper or lower bound.

Neutralization

The "neutralization" of the factor is the factor exposure value which excludes the influence of market and industry. This can be calculated as follow:

$$\beta_{i,k}^{\mathrm{M}}=\alpha+\beta_i^{\mathrm{size}}f_{size}+\sum_{s=1}^{S}\beta_i^s f_s+\epsilon_i$$

Where$\beta_{i,k}^{\mathrm{M}}$ represents the factor value after the MAD method in the previous step, β_i^{size} represents the market value factor exposure of the stock, f_{size} is the return of the stock i on the market value factor k, β_i^s represents the factor exposure value of the stock i on the industry s, and f_s represents the return of the stock i on the industry s, and ϵ_i is the residual. Among them, the industry factor is 0-1 dummy variable ---- when the stock is classified as Shenwan first-level industry classification of the s industry s, its factor exposure on the s industry is 1, otherwise it is 0.

2. Appendix B

ReLU Activation Function

The ReLu (Rectified Linear Unit) function is also known as a correction linear unit and can be expressed as

$$\mathrm{f(x)}=\max(0,\mathrm{x})$$

This function means that when the value of the node x is less than 0, the output of the activation function is 0, otherwise the output value is x. This method can be interpreted as threshold of 0 as a demarcation of neuron sleep and excitability, where the intensity of neuron output is proportional to the intensity of input stimulation. Some of the output values under the ReLU function is 0, which makes the network sparse, helping reduce the model overfitting. In addition, the ReLU function has a faster convergence speed and is suitable for a network model with more layers.

Categorical Cross-entropy

Categorical cross entropy function is suitable for model loss measurements under multiple classifications, assessing the difference between the probability distribution and the real distribution obtained by the current training. It can be written as follow:

$$\mathrm{E_{loss}} = -\sum_{i=1}^{n} \hat{y}_{i1} log y_{i1} + \cdots + \hat{y}_{im} log y_{im}$$

$$\mathrm{s.t.\, y_{i1} = h(z)}$$

$$\mathrm{z} = \sum W_j X_j + b$$

where n is the sample size, m is the number of categories, $\hat{y}_{im}$ is the expected output of m category and y_{im} is the actual output, h(z) denotes the activation process.

Since the loss function is made up of multiple categories, the calculation of loss is also multiple, and the formula is:

$$\frac{\partial \mathrm{E_{loss}}}{\partial \mathrm{w}} = \sum (h(z) - y_{i1})$$

The loss depends on the difference between the expected output and the actual output. The larger the actual output is than expected, the greater the loss. When the difference between the two is smaller, the loss can be approximated to 0.

3. Appendix C

IC Test

The IC value of the factor refers to the correlation coefficient between the exposure in t period and the stock return in t+1 period. The factor IC value measures the linear correlation of the next period return of individual stocks and the exposure of the current factor, reflecting the robustness of the forecast ability. Single factor IC testing usually includes calculating the mean of IC values, the mean of absolute values of IC values, the standard deviation of IC values, and the ratio of the mean of IC values to the standard deviation. The higher the above indicators (the smaller the standard deviation of the IC value), the more stable the factor's prediction ability is.

Similarly, Rank IC refers to the cross-sectional correlation coefficient between the factor's ranking of all stock exposures and its next-period return ranking.

GRS Statistics

This statistic actually belongs to F-statistics, examining the effectiveness of the model on the stock return estimate. The main idea is that if the asset pricing model is effective, then the factor used by the model should be able to cover all changes in the stock yield, then the regression intercept should be 0 (i.e. pricing error is 0)." The GRS statistics are constructed as follows:

$$\mathrm{GRS} = \frac{T}{N} \times \frac{T-N-K}{T-K-1} \times \frac{\alpha' \sum^{-1} \alpha}{1+\bar{\mu}\Omega^{-1}\bar{\mu}}$$

Among them, T is the sample period, N is the number of portfolios, K is the number of model factors, and α is N x 1 vector, where the elements are the intercept terms of regressions. $\boldsymbol{\Sigma} = \frac{\boldsymbol{\epsilon'\epsilon}}{\boldsymbol{T-K-1}}$ is the unbiased estimation of residual covariance matrix, where ϵ is the residual matrix that is obtained after the regression. $\bar{\mu}$ is a $\mathrm{K} \times 1$ vector, the elements are respectively the mean value the factors; and Ω is the factor the variance-covariance matrix

Comparing the Fama-French factor model with the fully connected neural network factor model, the GRS in back test should be reduced if the new factor is valid and contributes to the forecast accuracy.

Alpha Value

Alpha is the intercept term of the traditional Fama-French three-factor model or the fully connected neural network factor model, which indicates the stock return that cannot be explained by the model. Therefore, if the alpha of the model decreases after the new factor is added, it can be concluded that the factor can independently explain the stock return source outside the three-factor model. Here the analysis is based on the mean of the absolute value of alpha.