Predicting Direction and Magnitude of Credit Spread by Machine Learning[1]

Abstract: Machine learning is now widely used in quantitative investments. We selected the investment-grade bond and national debt credit spread from 1997 to 2018 as the research object, and used the ridge regression, Lasso regression, Polynomial regression, and decision tree regression to train the data to predict the spread of the next period. We use the decision tree to train the data and predict the direction of the next credit spread. Based on the results, we found that using the decision tree to predict the accuracy of the ups and downs is as high as 60% and using decision tree regression to get the best fit and prediction results, the training set and the test set Goodness is as high as 0.95.

Keywords: Machine learning; Credit spread; Decision tree

CONTENTS

1. Introduction 1
2. Machine Learning Model 2
2.1 Ridge Regression 2
2.2 Lasso Regression 3
2.3 SVM Regression 4
2.4 Decision Tree 4
3. Indicator Selection and Data Resources 4
3.1 Variable Selection 5
3.2 Data Resources 6
3.3 Data Description 6
4. Output 7
4.1 Liner Regression 7
4.1.1 Simple Liner Regression 7
4.1.2 Ridge Regression 7
4.1.3 LASSO Regression 8
4.1.4 Polynomial regression 8
4.2 SVM Regression 8
4.3 Decision Tree 8
4.3.1 Decision Tree Regression 8
4.3.2 Decision Tree Classfication 9
5. Conclusion 9
Reference 9

1. Introduction

Credit spreads represent the difference between the yields on risky bonds and risk-free securities. They reflect not only the uncertainty at the macro and corporate levels, but also the supply and demand dynamics for bonds in the market. We can benefit by predicting changes in credit spreads. More specifically, we can buy corporate bonds and sell treasuries when we have confidence in the macro economy and expect spreads to narrow in the future. The opposite is true when we expect future spreads to widen.

Based on a series of related literatures, we established the potential influencing factors to credit spread. They can be divided into two categories. The first includes changes in macroeconomic fundamentals, such as the non-agricultural employment and the CPI index. The second is market supply and demand indicators, such as bond supply and VIX index. However, in the subsequent empirical test, we found that fundamentals rather than supply and demand factors have a stronger ability to explain and predict credit spreads in the long run, and the addition of short-term emotional factors actually reduces the stability of the model. In order to ensure a sustainable return in the market, we

[1] XueLongFan, maste of UIBE, JiaWeiZhu,XiaoLiSun,LinYiJiang

selected the most valuable factors to build our model, which is mainly composed of macroeconomic indicators.

In addition, considering that our main study object is the investment grade corporate bonds with high credit level,, and considering that the firm-level risk can also be eliminated by fully diversified individual risk, so unlike many related literature, company-specific business situation is not in the scope of this paper. Instead we focus on the systemic factors.

The variables we chose mainly include some of the macroeconomic leading indicators, such as bank credit growth, return on the S&P 500 index, the M2 year-on-year growth rate, which precede the real economy. We also chose some factors which can indicate the boom and bust of economic activities, such as the year-on-year growth rate of CPI, the Industrial Production Index, the non-agricultural employment growth rate, etc. And we also added to the model the dummy variables and time trend, to join the influence of time trend and the kind of bonds. The specific reasons and references for the selection of variables will be detailed in the third part.

On the selection of the model, we begin with the most ordinary linear regression, followed by different linear regressions such as the Ridge Regression, LASSO Regression, Multiple Regression, SVM Regression and decision tree. Most of the regressions have achieved ideal result, which proves that in this paper, the model of credit spread has higher ability to predict indeed.

2. Machine Learning Model

During this modeling process, we mainly use the following machine learning models: Ridge regression, LASSO regression, SVM regression and Decision tree learning. We will give a general introduction to these machine learning models.

2.1 Ridge Regression

Ridge regression, in machine learning it is known as weight decay, and with multiple independent discoveries, it is also variously known as the Tikhonov–Miller method, the Phillips–Twomey method, the constrained linear inversion method, and the method of linear regularization. It is related to the Levenberg–Marquardt algorithm for non-linear least-squares problems.

Ridge regression is a biased estimation regression method dedicated to collinear data analysis. It is essentially an improved least squares estimation method. By abandoning the unbiasedness of least squares method, ridge regression is at the expense of losing part of information and reducing precision to Obtain a regression coefficient that is more realistic and more reliable, and the regression effect with ill-conditioned data is better than the least squares method.

Generally, the R-squared value of the ridge regression equation is slightly lower than the ordinary regression analysis, but the significance of the regression coefficient is often significantly higher than that of the ordinary regression. It has great practical value in the research of collinearity problems and ill-conditioned data.

Suppose that for a known matrix A and vector b, we wish to find a vector x such that:

$$\mathbf{Ax} = \mathbf{b}$$

The standard approach is ordinary least squares linear regression. However, if no x satisfies the equation or more than one x does—that is, the solution is not unique—the problem is said to be ill posed. In such cases, ordinary least squares estimation leads to an overdetermined (over-fitted), or more often an underdetermined (under-fitted) system of equations. Most real-world phenomena have the effect of low-pass filters in the forward direction where A maps x to b. Therefore, in solving the inverse-problem, the inverse mapping operates as a high-pass filter that has the undesirable tendency of amplifying noise (eigenvalues / singular values are largest in the reverse mapping where they were smallest in the forward mapping). In addition, ordinary least squares implicitly nullifies every element of the reconstructed version of x that is in the null-space of A, rather than allowing for a model to be used as a prior for x. Ordinary least squares seeks to minimize the sum of squared residuals, which can be compactly written as:

$$\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2$$

where $\|\bullet\|_2$ is the Euclidean norm.

In order to give preference to a particular solution with desirable properties, a regularization term can be included in this minimization

$$\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2+\|\boldsymbol{\Gamma x}\|_2^2$$

for some suitably chosen Tikhonov matrix $\boldsymbol{\Gamma}$. In many cases, this matrix is chosen as a multiple of the identity matrix $\boldsymbol{\Gamma}=\boldsymbol{\alpha I}$, giving preference to solutions with smaller norms; this is known as $\boldsymbol{L_2}$ regularization. In other cases, high-pass operators (e.g., a difference operator or a weighted Fourier operator) may be used to enforce smoothness if the underlying vector is believed to be mostly continuous. This regularization improves the conditioning of the problem, thus enabling a direct numerical solution. An explicit solution, denoted by $\hat{\boldsymbol{x}}$, is given by:

$$\hat{\boldsymbol{x}}=(\boldsymbol{A^T A}+\boldsymbol{\Gamma^T\Gamma})^{-1}\boldsymbol{A^T b}$$

The effect of regularization may be varied by the scale of matrix $\boldsymbol{\Gamma}$. For $\boldsymbol{\Gamma}=\boldsymbol{0}$ this reduces to the unregularized least-squares solution, provided that$(\boldsymbol{A^T A})^{-1}$ exists.

$\boldsymbol{L_2}$ regularization is used in many contexts aside from linear regression, such as classification with logistic regression or support vector machines, and matrix factorization.

2.2 Lasso Regression

In statistics and machine learning, lasso (least absolute shrinkage and selection operator; also Lasso or LASSO) is a regression analysis method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the statistical model it produces.

Lasso was originally formulated for least squares models and this simple case reveals a substantial amount about the behavior of the estimator, including its relationship to ridge regression and best subset selection and the connections between lasso coefficient estimates and so-called soft thresholding. It also reveals that (like standard linear regression) the coefficient estimates need not be unique if covariates are collinear.

Though originally defined for least squares, lasso regularization is easily extended to a wide variety of statistical models including generalized linear models, generalized estimating equations, proportional hazards models, and M-estimators, in a straightforward fashion. Lasso's ability to perform subset selection relies on the form of the constraint and has a variety of interpretations including in terms of geometry, Bayesian statistics, and convex analysis. The LASSO is closely related to basis pursuit denoising.

Lasso was originally introduced in the context of least squares, and it can be instructive to consider this case first, since it illustrates many of lasso's properties in a straightforward setting.

Consider a sample consisting of N cases, each of which consists of p covariates and a single outcome. Let y_i be the outcome and $\chi_i=\left(x_1,x_{2,}\dots,x_P\right)^T$ be the covariate vector for the I^{th} case. Then the objective of lasso is to solve:

$$min_{\beta_0,\beta}\{\frac{1}{N}\sum_{i=1}^{N}(y_i-\beta_0-x_i^T\beta)^2\}$$

subject to $\sum_{j=1}^{P}|\beta_j|\le t.$

Here t is a prespecified free parameter that determines the amount of regularisation. Letting X be the covariate matrix, so that $X_{ij}=(x_i)_j$ and x_i^T is the I^{th} row of X, the expression can be written more compactly as:

$$min_{\beta_0,\beta}\{\frac{1}{N}\|y-\beta_0 1_N-\chi\beta\|_2^2\}$$

subject to $\|\beta\|_1 \leq t$.

Where $\|\beta\|_p = \left(\sum_{i=1}^{N}|\beta_i|^P\right)^{\frac{1}{P}}$ is the standard l^Pnorm, and $1_N$is the $N \times 1$vector of ones.

Denoting the scalar mean of the data points X_i by $\bar{x}$, and the mean of the response variables y_i by $\bar{y}$, the resulting estimate for β_0 will end up being $\widehat{\beta_0} = \bar{y} - \bar{x}^T\beta$, so that:

$$y_i - \widehat{\beta_0} - \bar{x}^T\beta = y_i - (\bar{y} - \bar{x}^T\beta) - x_i^T\beta = (y_i - \bar{y}) - (x_i - \bar{x})^T\beta$$

and therefore it is standard to work with variables that have been centered (made zero-mean). Additionally, the covariates are typically standardized $\left(\sum_{i=1}^{N}x_i^2 = 1\right)$ so that the solution does not depend on the measurement scale.

2.3 SVM Regression

In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

2.4 Decision Tree

Decision tree learning uses a decision tree (as a predictive model) to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). It is one of the predictive modeling approaches used in statistics, data mining and machine learning. Tree models where the target variable can take a discrete set of values are called classification trees; in these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. Decision trees where the target variable can take continuous values (typically real numbers) are called regression trees.

Decision tree learning is the construction of a decision tree from class-labeled training tuples. A decision tree is a flow-chart-like structure, where each internal (non-leaf) node denotes a test on an attribute, each branch represents the outcome of a test, and each leaf (or terminal) node holds a class label. The topmost node in a tree is the root node.

There are many specific decision-tree algorithms. Notable ones include: ID3 (Iterative Dichotomiser 3)、C4.5 (successor of ID3)、CART (Classification And Regression Tree)、CHAID (CHi-squared Automatic Interaction Detector)、MARS.

3. Indicator Selection and Data Resources

We focus on AAA and AA corporate bonds yield data, trying to use a set of variables to model the history credit spreads of high grade bonds. Credit spreads are defined as the yield corporate bonds less the 10 year yield on US treasury bonds. We use the ICE BofAML US Corporate AAA and AA effective yield as the representative of corresponding grade of corporate bonds. Effective yield excludes the impact of taxation.

Considering that the risk of individual default of high-rated bonds is quite low, investment risks on these bonds mainly come from inflation risk, liquidity risk and systemic default risks. The three are mainly affected by changes in the macroeconomic environment. Meanwhile, the change of US treasury bonds yield is also mainly affected by macro factors. So our model considers several major macroeconomic indicators, include bank credit scales, overall equity market return, currency supply, industrial production, non-agricultural employment and inflation. The variables in the model use the year-on-year growth rate of the above indicators except stock returns. The ICE BofAML bonds data have an 22 year data history, which limits the length of time series to some extent. Considering the availability of data, the time window is from January 1997 to December 2018. And the data frequency is monthly.

3.1 Variable Selection

1.Bank credit scales. Bank credit scale includes the scale of monthly bank credit increments of all commercial banks in the United States, expressed as bank credit in the model. Bank credit scale can reflect the amount of funds obtained by real economy through major indirect financing channels, which is the leading indicator of macroeconomic changes. An increase of bank loans indicate that the investment demand of enterprises has increased and the credit environment is relatively loose, which often means the economic prosperity will increase, and the supply of corporate bonds may also increase.

2. Overall equity market return. The overall equity market is reflected by S&P500 index monthly yield, expressed as SPX. The yield of the stock market is generally considered to be the leading indicator of the macro economy. In addition, the stock return rate can be used as a proxy variable for the company's operating environment according to Merton's structured model (1974). When the equity market increase, assets value correspondingly rise and leverage fall, then firms become less likely to default. Thus the credit spread tighten as the equity market increase.

3.Money supply. Money supply is defined as the year-on-year growth rate of M2 and expressed as M2 in the model. Money supply indicates the easing degree of monetary policy, while reflecting the actual and potential purchasing power. It is a leading indicator of economic cycle. In a loose monetary environment, the liquidity of bond market will increase. In addition, In addition, M2 affects risk-free interest rate levels and inflation levels to some extent, which will further affects credit spreads.

4.Industrial production, non-agricultural employment. The current economic conditions are represented by the year-on-year growth rate of non-agricultural employment and industrial production, expressed by NonA and INDPRO respectively. The prosperity of the former economy will affect the default probability of corporate bonds. Similar to other variables above, The idea here is to measure the overall degree of corporation leverage and potential recovery rate in case of default. In addition, the credit spread will be reflected to the expected level. The above three macroeconomic leading indicators and the proxy variables of the two simultaneous indicators can reflect investors' expectations of the future economic situation and the trend of the credit spread.

5.Inflation.We use the YOY growth rate of consumer price index (expressed as CPI) as a proxy variable of inflation level. The level of inflation is a lagging indicator of the economic cycle. A reasonably stable inflation level is a sign of healthy economic performance, and dramatic changes in inflation levels often indicate a deterioration in the economy. The rise in inflation levels often leads to an increase in bond yields and a worsening of bond performance, which affects the liquidity of the bond market and further affects the credit spreads.

We add proxy variable for time trends in the model, expressed as time. time is defined as the number of month from January 1, 1997. And we use dummy variable level to distinguish AAA and AA grade of bonds. When the credit spread is AAA corporate bonds less the 10 year yield on US treasury bonds, level equals to 0, and when the credit spread is AA corporate bonds less the 10 year yield on US treasury bonds, level equals to 1.

3.2 Data Resources

The sample in this paper includes monthly data for a total of 22 years from January 1997 to December 2018. Both the AAA and AA corporate bonds effective yields and 10-year treasury constant maturity rate come from FRED. Apart from the monthly return of S&P 500 Index, the macro data are all from the Wind database, including the growth rate of bank credit scale, M2 year-on-year growth rate, CPI year-on-year growth rate, INDPRO index growth rate, non-agricultural employment growth rate, etc. Monthly returns on the S&P 500 are based on the resset database.

3.3 Data Description

Figure 1 and 2 show the yields on AAA-grade corporate bonds, AA-grade corporate bonds and their credit spreads over the 10-year Treasury rate, respectively. As can be seen from the figure, there is a strong correlation between credit spreads and corporate bond yields, but a weak correlation between credit spreads and risk-free yields.

We calculated the correlation coefficient between the corporate bond yield and the Treasury bond yield and the credit spread, and the result confirmed our observation. The correlation coefficient between the AAA corporate bond and the its credit spread to risk-free rate is 0.5023, and that of the AA corporate bond is 0.5421, while the correlation coefficient between the 10-year Treasury bond and the two credit spreads was only 0.1567 and 0.1165. We speculate that this may be due to that when economic uncertainty is rising, people tend to sell corporate bonds and buy treasuries or hold cash as a hedge, pushing up corporate bond rates and pushing down the

risk-free rate. In that case, credit spreads and corporate bond rates move in the same direction, as opposed to risk-free treasury rates.

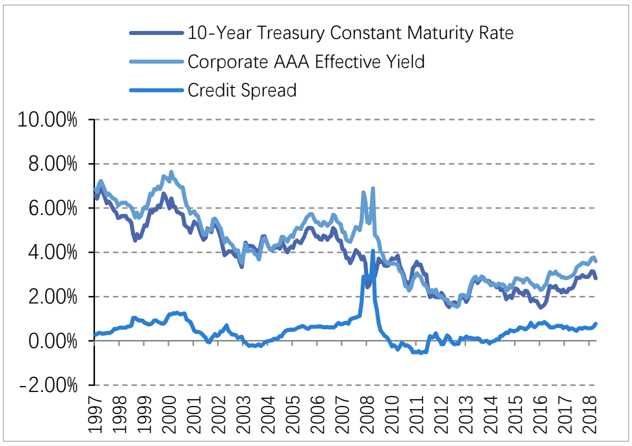

Figure 1 .Plot of AAA Bond Yield, 10-year Treasury and Credit Spread

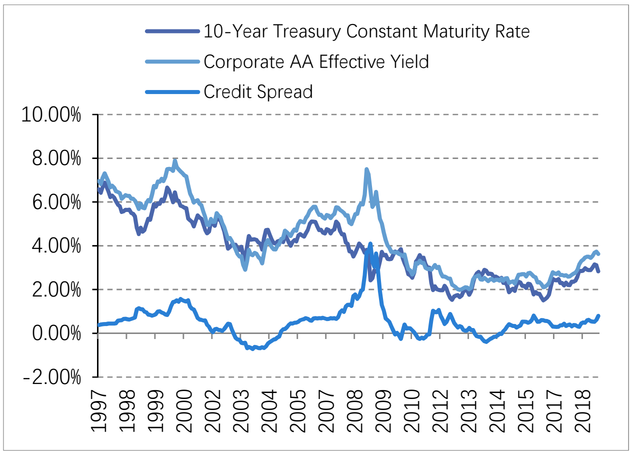

Figure 2 .Plot of AAA Bond Yield, 10-year Treasury and Credit Spread

Table 1 shows the descriptive statistical results of selected macroeconomic data. We see that the year-on-year growth rate of bank credit is the most volatile, with a standard deviation of 14.40 percent, while the volatility of S&P 500 index's monthly return is similar to that of the industrial production index.

Table 1 Summary Statistics For All Variables

Variable	Obs	Mean	Std.Dev.	Min	Max
Bankcredit	264	6.89%	14.40%	-24.50%	39.80%
SPX	264	0.61%	4.28%	-16.90%	10.80%
M2	264	6.25%	1.80%	1.62%	10.60%
CPI	264	2.18%	1.20%	-2.10%	5.60%
INDPRO	264	1.60%	4.25%	-15.30%	8.56%
NonA	264	1.01%	1.46%	-4.35%	4.00%

4. Output

The followings are the outputs of machine learning methods.

4.1 Liner Regression

Building a linear regression model:

$$spread_{i,t} = C + \alpha_1 * spread_{i,t-1} + \alpha_1 * Time_{i,t-1} + \alpha_2 * Bankcredit_{i,t-1} + \alpha_3 * SPX_{i,t-1} + \alpha_4 * M2_{i,t-1} + \alpha_5 * CPI_{i,t-1} + \alpha_6 * INDPRO_{i,t-1} + \alpha_7 * NonA_{i,t-1} + \alpha_8 * Level_{i,t-1}$$

Among them, Spread is the difference of interest rate betweent investment-grade bond and the yield of government bond, Time is the time trend item, representing the number of months from January 1, 997, Bankcredit is the year-on-year growth rate of bank credit scale, SPX is the S&P 500 monthly yield rate, M2 is M2 year-on-year growth rate, CPI is CPI year-on-year growth rate, INDPRO is INDPRO index monthly growth rate, NonA is non-agricultural growth rate, level is a dummy variable, when Spread is derived from AAA bond yield minus government bond yield, it takes 0, otherwise takes 1.

4.1.1 Simple Liner Regression

The train_test_split function is used to divide the data set into 80% of the data as the training set, 20% of the data as the test set, and the random number seed selection (random_state) 10. Train the training set using the LinearRegression.fit function.

In order to reasonably determine whether it is appropriate to divide the training set and the test set into this ratio, we tested the error obtained by linear regression machine learning under different training set sizes.

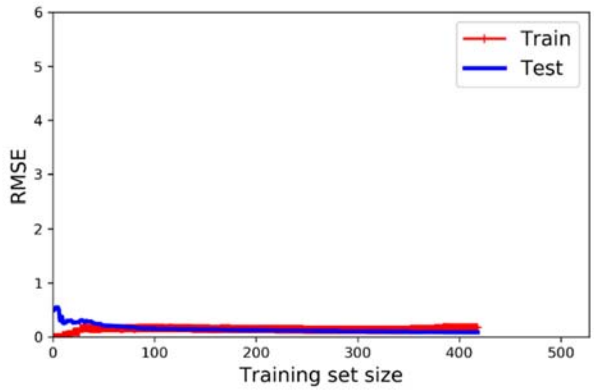

Figure 3 .Linear regression RMSE

Where Train is the bias of the training set, and Test is the bias of the test set. As can be seen from the graph(Figure 3), when the test set contains more than 200 samples, the error of the test set and the training set has been stabilized, so the training set and it is appropriate to divide the test set into 4:1.

The results obtained after training are as follows:

Training score:0.886849

Test score:0.965359.

4.1.2 Ridge Regression

Determine the penalty factor size from 0.1 to 1 for each increment from 0.1 to 1 for each increment from 1.(Figure 4)

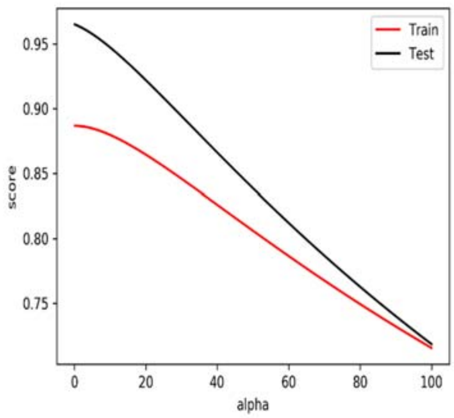

Figure 4 .Ridge regression score

Where Train is the score of the training set, Test is the score of the test set, and when the penalty coefficient alpha is 0.1, the prediction set has the highest score.

Using the Ridge function, alpha takes 0.1 pairs of data to train and get the result.

Train set score:0.886848

Test set score:0.965247

4.1.3 LASSO Regression

In order to determine the size of the penalty coefficient, take 1 from the negative 15th power of 10, and increase it by 1 time each time.(Figure 5)

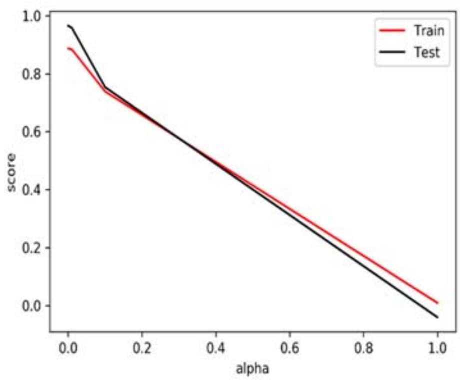

Figure 5 .Lasso regression score

As can be seen from the figure, the smaller the alpha, the higher the model score. Using the Lasso function, the alpha takes 0 and the result is obtained.

Train set score:0.886849

Test set score:0.965359

4.1.4 Polynomial regression

Use PolynomialFeatures to set polynomial parameters to get 52 independent variables in Table 1.

Table 2 independent variables generated by PolynomialFeatures

S(−1)	S(−1) ∗ CPI	Bank ∗ SPX	M2 ∗ CPI
Time	S(−1) ∗ INDPRO	Bank ∗ M2	M2 ∗ INDPRO
Bank	S(−1) ∗ NonA	Bank ∗ CPI	M2 ∗ NonA
SPX	S(−1) ∗ Level	Bank ∗ INDPRO	M2 ∗ Level
M2	Time^2	Bank ∗ NonA	CPI^2
INDPRO	Time ∗ Bank	Bank ∗ Level	CPI ∗ INDPRO
NonA	Time ∗ SPX	SPX^2	CPI ∗ NonA
Level	Time ∗ M2	SPX ∗ M2	CPI ∗ Level
$S(-1)^2$	Time ∗ CPI	SPX ∗ CPI	INDPRO^2
S(−1) ∗ Time	Time ∗ INDPRO	SPX ∗ INDPRO	INDPRO ∗ NonA
S(−1) ∗ Bank	Time ∗ NonA	SPX ∗ NonA	INDPRO ∗ Level
S(−1) ∗ SPX	Time ∗ Level	SPX ∗ Level	NonA^2
S(−1) ∗ M2	Bank^2	M2^2	NonA ∗ Level

Note: S(-1) means the last spread, Bank means Bankcredit

The results obtained after training are as follows

Train set score:0.935728

Test set score:0.954971

4.2 SVM Regression

In order to determine the size of the coefficient C, take from 1 to 100, increase by 1 each time.(Figure 6)

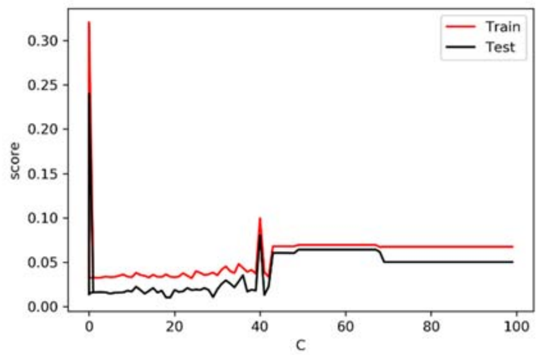

Figure 6 .SVM regression score

According to the output result, it is found that C is optimal at 1, and the result is trained using the SVR function with C equal to 1.

Train set score: 0.8801077948959344

Test set score: 0.9696801420535387

4.3 Decision Tree

4.3.1 Decision Tree Regression

In order to determine the size of the depth, take from 1 to 9, increase by 1 each time.

According to the output result(Figure 6), it is found that depth is optimal at 6.

Take depth equal to 6 and calculate the result under different min impurity decrease.

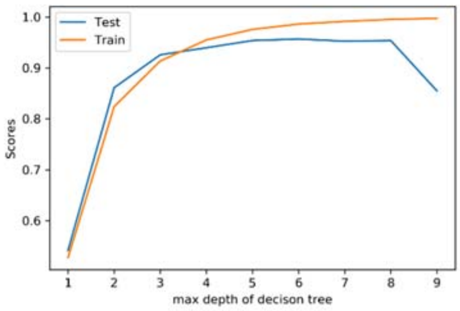

Figure 7. score of different max depth

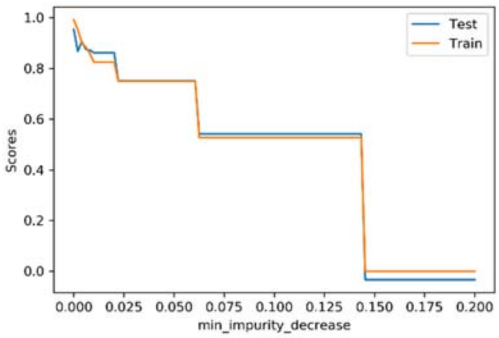

Figure8. score of different min impurity decrease

According to the output result, we decide that max depth takes 6 and min impurity decrease takes 0.002.(Figure 8) Using decision tree regression trains them and get the output:

Train set score: 9532637771370992

Test set score: 0.9531162483054741

4.3.2 Decision Tree Classfication

In order to perform classification training, we set a new dependent variable. When the spread of the t period is greater than the spread of the t-1 period, the dependent variable of the t period is taken as 'up', otherwise it is taken as 'down'. Use decision tree classification training to get results with max depth taking 6 and min impurity decrease taking 0.002

Train set score: 0.611904761904762

Test set score: 0.6037735849056604

5. Conclusion

According to the results of the regression, most of the credit spreads in the next period are explained by the current credit spreads. This result is also logical. In addition, there are still some differences that cannot be predicted by the current credit spread. According to our results, this difference can be better predicted by the current S&P monthly rate, M2 year-on-year growth rate, and CPI year-on-year growth rate

In the case of avoiding overfitting, the introduction of time trend items, bank credit size changes, INDPRO year-on-year growth rate, non-agricultural year-on-year growth rate and bond level also improved the validity of the model.

According to the results obtained by using different machine learning models, we find that the best results can be obtained by using the decision tree model. The goodness of fit of the training set reaches 0.953116, and the goodness of fit of the prediction set reaches 0.953263. There is no over-fitting. And under-fitting problems.

However, unfortunately, in the prediction direction, the prediction accuracy using the decision tree model is only about 0.6, in which the prediction accuracy of the training set is 0.611905, and the prediction accuracy of the test set is 0.603774, but this has exceeded the immediate prediction of 0.5. The probability level, so the decision tree model is effective for predicting the direction, and thanks to the high goodness of fit, the degree of variation can be well measured.

Reference

[1] Chong, T. T. L., He, Q., Ip, H. T. S., & Siu, J. T. (2017). Profitability of CAPM Momentum Strategies in the US Stock Market.

[2] Fan, M., Li, Y., & Liu, J. (2018). Risk adjusted momentum strategies: a comparison between constant and dynamic volatility

scaling approaches. Research in International Business and Finance.

[3] Siri, J. R., Serur, J. A., & Dapena, J. P. (2017). Testing momentum effectfor the US market: From equity to option strategies (No. 621). Universidad del CEMA.

[4] Du Plessis, J., & Hallerbach, W. G. (2016). Volatility Weighting Applied to Momentum Strategies. The Journal of Alternative Investments, 19(3), 40-58.