The Credit Risk Assessment for Chinese Companies Based on CAFÉ System by Using Bigdata Method

George X. Yuan, Chengxing Yan, Yunpeng Zhou and Haiyang Liu

May 31, 2022

Credit Risk Assessment for Chinese Companies based on CAFÉ System by using Bigdata Method[1]

George X. Yuan, Chengxing Yan, Yunpeng Zhou and Haiyang Liu.

1. Shanghai Hammer Digital Technology Co., Ltd, Shanghai 200093, China
2. Business School, Chengdu University, Chengdu 610106, China
3. School of Business, Sun Yat-sen University, Guangzhou 510275, China

Abstract

This paper examines how the CAFÉ System developed by using bigdata approach is able to resolve three problems for which the current credit rating market in China facing such as "the rating is falsely high; the differentiation of credit rating grades is insufficient; and the poor performance of predicting early warning and related issues". These issues are done by redefining the "BBB" as the basic investment level in accordance with international practice in capital markets, and cases study shows that how the so-called "CAFÉ Credit" is able to resolve the major issues in the capital market in China for credit rating.

JEL Classification: C53 C58 G21 G24 G32
Keywords: Credit rating; CAFÉ System; Default Matrix; Credit Transition Matrix; ROC and AUC Testing; Non-Structured Feature.

1. The Background and Related Issues

The credit rating is one of the most important parts for today's financial market economy (Altman, 1968; Altman, 1983; Altman & Sabato, 2007; Hull, 2017; Hull, 2018). After nearly 30 years of rapid development in the domestic financial industry, the current domestic credit rating market in China is now facing at least three main problems (BOC, 2021), they are:

1. **The rating is falsely high; and**
2. **The differentiation of credit rating grades is insufficient; and**
3. **The poor performance of predicting early warning.**

Since the establishment of the world's first Credit Rating agency by U.S. Moody's in the early 20th century, the Credit Rating industry with a century of development, has played an important intermediate role in promoting market development, revealing and preventing credit risks, reducing transaction costs, and assisting the government in financial supervision, and of course, has faced many adjustments (Dun & Bradstreet,2014; FICO, 2018; Anderson, 2007; Chi et al., 2019; Thomas et al., 2017; Witzling, 2016; Yuan & Wang, 2019). The Securities and Exchange Commission of the U.S. also believes that Credit Rating results have become more and more important to investors and other market participants in recent years, affecting issuers' access to the capital market, funding costs, financial transaction structure, trustee's investment capabilities, and so on. At the same time, the development and growth of the Credit Rating industry and the formation of a system depend to a large extent on the development of the financial market, especially the bond and securities market (Jing et al., 2003; Du, 2017; Zhang, 2018; Ministry, 2008). For the general introduction and discussion of financial market and corresponding credit rating system comprehensively, please also see Standard & Poor's, Moody's Investors Service, and Fitch Ratings (Baidu, 2020; and related materials wherein).

Thanks to nearly 30 years of development, China's rating business includes almost all aspects of China's financial market. However, there are still many problems existing, where we briefly summarize them as follows: first, the understanding of the Credit Rating industry needs to be deepened; second, the legislative work on Credit Rating is obviously lagging behind; third, the failure to form an effective supervision system; fourth, the imbalance between supply and demand in the Credit Rating market is prominent, which is, on the one hand, there are many rating agencies, on the other hand, the credit rating industry has few quality professional products available for rating and poor business stability, which makes many credit rating companies rely on non-credit main business support; fifth, the independence and impartiality of credit rating agencies needs to be enhanced, and they are subject to more administrative interventions in the process of conducting credit ratings and lack objective independence; sixth, there is a lack of objective and credible rating behavior, and this is the most deadly problem at present, especially in the objective assessment and handling of the actual situation of China's financial market itself, generally existing immature professional rating technologies, and the problem that the quality of credit ratings urgently needs to be improved professionally (Jing et al., 2003; Du, 2017; Feng, 2019).

[1] The Corresponding author is: George Yuan with email address: george_yuan99@yahoo.com

Incorporating with the modern rating theory and approach in the practice from international Credit Rating agencies, China's Credit Rating companies are gradually exploring Credit Rating methods and technologies that are suitable for China's national conditions, and have initially formed rating methods that can cover basic rating theories, Credit Rating models, and Credit Rating system based on classification of industries, products, and subjects. Nevertheless, the gap between China's Credit Rating companies and its international counterparts is also very prominent: for instance, Credit Rating is a necessary disclosure factor to promote the issuance of credit bonds in our country, and an important reference basis for bond issuance pricing, but during the rapid development of China's bond market, a large number of potential risks have been accumulated, and default events have occurred frequently, and the risks have shown normalization of default events, diversification of the nature of the subject, diversification of bond varieties, diffusion of industry distribution, and diversification of default area distribution and so on often happened. A typical incredible case is that the ratings of most China's companies by domestic rating agencies are still high with the rating grades from AA to AAA levels mostly, for example, according to the data reported by the platform “Wind” at the end of year 2020, almost more than 90% of issuers and their bonds/debts’ rating are in the range of AA and above, compared with the Credit Ratings of international rating agencies, the overall rating in Chinese capital markets is so higher and the distribution is more concentrated, which is not conducive to domestic and foreign investors to identify different risk of bond/debts in guiding investment decisions. If the quality of current credit rating market for domestic financial market in China is not improved and adjusted in time, this will affect the healthy development of China's financial market, especially the capital market for ever.

Taking into account the fact that the available number of default (also, called “bad”) entitles observed for defaulted entities (companies or enterprises) in the market is very small, we must consider to find a new path to establish reasonable credit rating method suitable to Chinese markets with international standards. On the other hand, in the current era of digital economy (ecology), especially in today's rapid development of big data with the financial technology (Fintech), under the premise of fully considering the information provided by both traditional structure and unstructured data, using new approach in dealing with non-structure data which is called “Hologram” approach (Yuan & Wang, 2019) as a fundamental tool, we are able to extract (non-structured) risk feature factors based on un-structured data (instead of only traditional structure data) as breakthroughs to establish the so-called "CAFÉ Risk Assessment System" (in short, CAFÉ system) to conduct rating for almost 10,000 companies in China by including all listed companies and bonds/debts’ issuers (Yuan, 2019a; 2019b). At the same time, combining the international standards that must be considered in the financial credit market, the basic investment level recognized in the financial industry is with “BBB” grade as the starting level, we are able to resolve the issue for the problem without enough default (also, called “(default) bad”) samples by creating enough required “bad samples” under the category of non-structure data types, this would help us to establish a so-called “CAFÉ Credit Rating system”(in short, “CAFÉ”, or “CAFÉ System”) for China's corporate entities and bonds (debts) that are in line with international standards (e.g., see Yuan, 2019a; Yuan, 2019b; Yuan, 2021).

In this paper, we first point out the shortcomings of current China's ratings, and then discuss the idea how the framework of “CAFÉ Risk Assessment System” is established by applying the Hologram approach. The application of the “CAFÉ Risk Assessment System” in the credit rating is reflected as the “Intelligence Stone Rating System” (in short, ‘IS’) (Yuan, 2019a; 2019b). The foundation of our CAFÉ system is a multi-dimensional risk assessment under the framework of big data analysis by using the so-called Hologram approach (Yuan & Wang, 2019) applying to “heterogeneous” data with combining the concept so-called “dynamic ontology” to achieve the extraction of entities (corporate companies)’s risk genes by using AI algorithms (mainly, the Gibbs Sampling method) to resolve the issue “not sufficient (defaulted） bad samples” (Yuan et al., 2020a; 2020b; 2020c; Yuan, 2021). In this way, we are able to achieve the comprehensive dynamic assessments for companies’ credit risk from the four dimensions which consists of “Corporate structure hologram” (denoted by “C”), “Accounting behavior hologram” (denoted by “A”), “Financial behavior hologram” (denoted by “F”) and “Ecosystem Hologram” (denoted by “E”), thus in short, “ CAFÉ” system to form the “CAFÉ Risk Assessment System” for financial markets in China.

In this report, we focus on the application of the CAFÉ System (mainly, the “Intelligence Stone Rating System”) to conduct specific analysis for two listed companies, especially by combining the actual market performance for each case in the past from one to three years back in history to against our risk assessment results with one-to-one interpretation of event screening and risk assessments derived by our Credit Risk system “IS”, to show how the framework of “CAFÉ Risk Assessment System” has at least the ability to overcome the current three major problems in Chinese markets, which are：1) The rating is falsely high; 2) The differentiation of credit rating grades is not sufficient; and 3) The poor performance of predicting early warning”.

This paper consists of four parts: The first section is an introduction to the background and issues we face in current Chinese markets; The second section discuss the basic framework and the key ideas of our "CAFÉ Risk Assessment System" established under the framework of big data by applying the Hologram approach as a tool; The third part is for the case study for two entitles from the markets, in which we conduct the discussion one by one to against each company’s actual performance with the risk assessment derived by our CAFÉ system; and finally the fourth section is with conclusion and comment.

2. The Framework of CAFÉ Risk Assessment System by using Bigdata Method

As the world's second largest economy, China needs to establish a risk assessment system suitable for its national conditions and in line with the international credit rating system. In order to achieve this goal, the first work is to build a credit evaluation model by scientific means, and enough bad samples are needed to realize modeling, model proofreading and testing. Then, according to the international credit rating standard, return to the most basic through the professional definition of "AAA" to "C", and take "BBB credit rating as the most basic investment grade" as the standard to establish the credit rating system for the subject or bond: that is, through the following five steps to establish a scientific standard system for credit rating with the default rates as the core base for the entity (subject or company, or, saying, the issuer of debts) and bonds (debts) (yuan, 2019a;2019b):

Step 1: Defining credit rating standards with default rate as the core

Step 2: Constructing the "default matrix" required by credit risk

Step 3: Construct the "credit transfer matrix" of the change of subject (company) and debt credit qualification;

Step 4: Supporting the "ROC" (and "AUC") testing for credit model performance in Steps 2 and 3 above;

Step 5: Build the most basic "(default) bad samples to support Steps 2, 3 and 4 above (so allow us to complete the credit rating system mentioned in Step 1) to reflect the real situation of China's capital market.

Here, we especially point out that the most difficult point in establishing China's credit rating (evaluation) system is that for the domestic financial market, the performance of the "bad default sample" corresponding to the capital market is different from that of the "default" sample in the financial market of western countries in genera! In addition, the database of default bad samples is not enough (for example, from 2007 to 2020, the number of samples of all default subjects is only about 200, and the (risk) information that should be disclosed by the corresponding bad samples is often insufficient), which requires us to develop new methods to solve the construction of bad samples suitable for China's situation for the domestic market. That is, China's credit evaluation system must build a description of China's bad samples, and in fact, the idea from the Fintech by using bigdata method can help us to realize this job as discussed by Yuan 2019a; 2019b）.

Today, the capital markets in China maintains of more than 50,000 bonds(dents), plus around 4,800 listed companies, but around 92% of the bonds are rated at AA credit rating or above. From the perspective of international standard in the practice, this is obviously inconsistent with the real market situation, which means that most of companies rated as "AA" or "AAA" are actually not true! On the other hand, since 2018, a number of companies (enterprises) rated as "AA" or "AAA" with their bonds/debts with “AAA” credit ratings (such as the state-owned platforms) directly went to bankrupt or announcing no plan to pay back the face amount and or accrued interest of bonds/debts, these events absolute are against the current domestic credit rating system in China. Therefore, it is so urgently to have an internationally accepted Credit Assessment

System suitable for China's financial market. In order to achieve this goal, as mentioned above, our starting point is first to find a scientific way to construct enough bad samples that required by the credit rating modelling to identify good or bad companies with the clear definition for the standard being "BBB" credit rating as the fundamental investment level associated with criteria consisting of the "Default Probability" and the "Credit Transition Matrix" for all classes from "C" to "AAA" credit rating grades. In this way, we first construct around more 1,200 bad samples for years since 2017 which was around 20% of listed companies in China's exchange stock markets, which support us to give a clear definition for the standards being "BBB" credit rating grade as the fundamental investment level; and finally we are able to establish a genera framework called "CAFÉ Risk Assessment", with its application in credit rating, called "Intelligence System" (IS) (Yuan., 2019a; 2019b).

In summary, our "Intelligence System" (IS) no longer uses the "AA" credit rating as the starting point for the basic investment-grade standard that has been popular in China in the past almost 30 years (BOC, 2021), instead, our IS system solves the problem as mentioned above, and this is also verified and we will see the discussion given supported by the case study below.

2.1. The Framework for the Construction of the CAFÉ Hologram Risk Assessment System

In the process of implementation, our starting point is to define the credit rating of entitles (companies, and also their bonds/debts) from the following four dimensions: First, the company's financial performance; Second, whether the company has fraud and how good or bad of the corresponding financial management quality; Third, how health of the company in terms of financial assets and liabilities, immediately indicating the degree of risk of company failure or default; and the fourth: The quality of the company's ecological environment and business risk (i.e., how good of the companies' ecosystems). This is done by using our Hologram approach (Yuan & Wang, 2019), which leads us to conduct the rating distinction more clearly, and the credit rating assessments can better reflect the company's actual credit status in the market.

Our "CAFÉ" is indeed the system that integrates all kind of information including the management structure of the corporation itself, the operational and business disclosure information, and related parts in terms of networks. All of these are done by classifying as four dimensions:

1) Corporate structure hologram (denoted by "C");
2) Accounting behavior hologram (denoted by "A");
3) Financial behavior hologram (denoted by "F"); and
4) Ecosystem Hologram (denoted by "E").

In short, denoted by " CAFÉ" to represent the "CAFÉ Risk Assessment System" for financial markets in China.

The two biggest features of this system are that it can convert static analysis into dynamic analysis, and then combine dynamic analysis with corporate ecology to form a more objective characterization of corporate ratings. But to form such a characterization, we need to integrate heterogeneous and heterogeneous big data with the hologram approach developed since in year of 2015 as a tool to implement the "data fusion" to conduct the risk feature factors (in terms of risk genes embedded in the complex network of entities.

We like to specifically share with the readers is that in order to have a good performance of risk assessment system for entitles in Chinese market, it is essentially to first establish the criteria for the modelling of company's financial fraud risk system, which should be the integration of financial statement analysis, governance structure analysis, audit and internal control analysis (for example, based on the general framework of the so-called "fraud triangle theory" as the starting point), and then to extract reliable risk characteristics (features) which include the following three parts (in terms of traditional structured and unstructured data in our CAFÉ system): They are, 1) the pressure/motivation dimension includes financial stability, external pressure, personal needs and financial goals, and so on; 2) the opportunity/vulnerability dimension Including industry attributes, invalid supervision, organizational structure; and 3) the excuse/attitude dimension includes the auditing dimension.

Based on the data from 2016 to 2020 from China's listed companies and non-listed companies, we first classify the fraud types into eleven cases: 1) occupation of company assets; 2) false disclosures (other); 3) illegal guarantees; 4) fraudulent listings; 5) unauthorized changes to the use of funds; 6) general accounting mishandling; 7) false records (or misleading statements); 8) postpone disclosure; 9) fictitious profits; 10) major material omissions; and 11) false listings of assets. Indeed these eleven types of fraud information support us to construct the “bad” sample collection, and by combining with audit information which allow us to form a dynamic and ecological enterprise risk assessment (see the discussion by the literature of [18]-[25]) under the help of Gibbs sampling algorithms for the extract of risk features.

2.2. The Extraction of Risk Features Based on AI Algorithm under the Digdata Framework

The quality for the construction of China's credit system is actually essentially with one key thing: how to deal with the situation where well-defined default (bad) samples are normally not available! That is, how we to “construct” a reasonable number of bad samples based on the unstructured samples to support about 9,000 companies (including around 4,800 listed companies, or around 5,500 issuers (for issuing bonds/debts))!

As discussed above, based on around 4,800 listed companies, we like to have around 1000 to 2000 bad samples. However, the actual situation is that until the end of 2020 (from 2007), the total number of default entitles (samples) that can be used to describe company failures in China is not more than 200, so the only choose we have is to consider around 2,700 events (which are mainly non-structure data) issued or released by the regulatory bodies of “China Banking and Insurance Regulatory Commission (CBIRC)), and china Securities Regulatory Commission (CSRC)’s punishment those listed companies and bond-issuing companies as the raw (unstructured) data! These raw samples basically consists of 11 categories as listed above and indeed, they basically appeared either embedded in the financial report in the form of document statements, or in other forms, thus qualified as big data samples. These around 2,700 bad bigdata samples are the original sources for us to construct at least 1,000 or about 2,000 bad samples required for modeling by using our big data approach in the practice. Once we have these 11 types of big data samples, we need to extract the risk characteristics embedded in these 11 types of bad samples to construct at least 1,000 or around 2,000 bad samples to supporting modeling for CAFÉ system.

In order to achieve this highly related risk feature extraction that characterizes fraudulent embedded from the un-structured data, we need to use an artificial intelligence-based algorithm tool: the so-called Gibbs sampling algorithm (see the discussion given by Yuan et al. (2020a; 2020b; 2020c) and Yuan (2021). The Gibbs sampling method (German & German, 1984) indeed is a Monte Carlo algorithm based on the Markov framework (MCMC) in statistics, which is very important and useful in deal with simulation in the practice (which is also listed as one of the top ten human algorithms in the 20th century). Since the launch of its basic algorithm prototype in the 1950s, in the 1970s and 1980s, the integration of the AIC (Akaike Information Criteria) (Akaike, 1974) and BIC (Bayesian Information Metrics) (Schwarz, 1978) test standards for the amount of information was used and developed, it is currently a very effective method to support feature extraction and statistical inference. Through the use of the Gibbs sampling algorithm, we can extract the highly correlated non-institutional risk characteristic factors that characterize financial fraud, which helps us to build a Risk Assessment System that distinguishes bad samples from good samples, and achieves the corresponding “the receiver operating characteristic curve”(referred to as “ROC” curve) of the model test, in which, “ROC curve value” is in the range from 0.72 to 0.75! To put it simply, if the ROC test result is 0.7 or above, it indicates that the features used have strong discrimination and interpretation capabilities) (Yuan et al., 2020a; 2020b; 2020c; Yuan, 2021) for the supporting the claim here), which thus supports our CAFÉ system has the ability to evaluate companies and debts more effectively.

Then, considering the classification and analysis of the company's major shareholders, management, board of directors, and board of supervisors according to the proportion of shareholding, identity, and internal and external ratios, using machine learning algorithm with the weight of evidence (in short, WOE) and information value (in short, IV) to explain the impact of the amount of information on the risk of fraudulent behavior for the entities, we have the following general conclusion with the basic assessment criteria (Yuan et al., 2020a;2020b; 2020c): The company’s shareholding structure affects the company’s financial fraud risk Important factors, and the following four characteristics can be used to warn the typical performance characteristics that may lead to fraud from the perspective of the corporate governance framework (Yuan, 2021):

1) The shareholding ratio of major shareholders and corporate legal persons is between 5% and 50%;
2) The cumulative shareholding ratio of major shareholders does not exceed 60%;
3) The shareholder of the management holds less than 1% of the shares; and
4) The proportion of major shareholders in the board of directors should not exceed 12%.

In this way, we incorporated the establishment of a feature system based on internal and external audit data.

Actually, we may ask this question: why there is always fraudulent behavior? It should have many reasons to answer this one, but according to the results of weight of evidence (WOE) binning and corresponding analysis of audit indicators (Yuan, 2021), we found that the number of audit committees and inconsistencies in opinions also have a corresponding impact on the company's fraud. Based on the report of our study by Yuan (2021), because the ROC curve of the audit model factors is linear, we have therefore used a rational method to demonstrate the following facts for the first time: External audits can only find out whether there is fraud, but cannot form inferences (if external audits can form inferences, the corresponding ROC curve should be a non-linear convex function form); secondly, based on the study of the relationship between the number of companies on the board of directors and the company's qualifications, we also found that regardless of the company's rating, the number of board committees of 7 to 9 people accounted for 80% of the total number of companies (except for large financial institutions and large group companies). Therefore, under the circumstances required by the company registration law, the number of board members within a reasonable range does not have an impact on the quality of the company. When combined with the factors of the board of directors, internal audit and external audit, we can also find that there are always auditing companies that give wrong opinions (that is, unqualified opinions are given when the company is fraudulent), and in turn, our hologram portrait technology can be used to verify whether the external audit is doing well or not.

Based on the CAFÉ system, the following is our analysis for the financial fraud of a listed company in China called "Guangzhou Langqi Industrial Co., Ltd." (In short, "Guangzhou Langqi" or "Langqi") in 2020 (see also the Case study in next section with more in details). By comparing the risk characteristics of financial fraud with the industry medians we established, we can clearly see the following two fact.

The Fact 1: In three years of 2019, 2018, and 2017, its financial indicator data is far from the median of the industry (see Table 1 below).

Table 1: During 2017-2019 Langqi's Financial Data with Industry Sector's Median Comparison

No#	Items	Median of Chem. ector	Yr 2019	Yr 2018	Yr 2017
1	Deduction of return on non net assets	6. 64%	0. 49%	0. 57%	1. 04%
2	Operating cash flow liability ratio	21. 44%	−8. 37%	−8. 14%	−6. 44%
3	Total asset turnover days	502. 27	244. 53	172. 61	138. 8
4	Prepayment turnover days	6. 38	28. 00	19. 25	13. 88
5	Growth rate of other receivables	16. 33%	−11. 23%	111. 56%	−47. 40%
6	Total growth rate of owner's equity (or shareholder's equity)	7. 78%	−14. 95%	21. 69%	3. 08%
7	Operating cost / total operating income	75. 65%	95. 15%	95. 88%	97. 18%
8	Interest expense (financial expense) / total operating income	0. 95%	1. 02%	0. 61%	0. 39%
9	Non operating net income / total operating income	1. 04%	0. 22%	0. 00%	0. 03%
10	Monetary capital / total assets	12. 88%	11. 61%	11. 50%	9. 33%
11	Other receivables (including interest and dividends) / total assets	0. 42%	0. 50%	0. 65%	0. 49%

12	Payroll payable / total assets	0.79%	0.34%	0.33%	0.38%
13	Paid in capital (or share capital) / total assets	13.35%	7.06%	8.13%	10.87%
14	Undistributed profit / total assets	19.19%	3.45%	3.40%	4.11%
15	Subtotal of cash paid to and for employees / cash inflow from operating activities	9.27%	1.24%	1.37%	0.85%
16	Subtotal of cash received from other financing activities / cash inflows from financing activities	0.00%	5.51%	11.10%	0.00%

The Fact 2: Based on the characteristic indicators that we have established to identify the company's possible fraudulent behaviors extracted from the corporate governance structure, we found that Guangzhou Langqi has a high risk of financial fraud in the following four characteristic indicators:

1) The shareholding ratio of major shareholders and corporate legal persons is between 5%-50%, and Langqi Company is 45%;

2) The cumulative shareholding ratio of major shareholders does not exceed 60%, and Langqi Company's 49%;

3) Management The proportion of major shareholders holding less than 1% of the shares of Langqi Company is 0;

4) The proportion of major shareholders on the board of directors does not exceed 12%, and the proportion of Langqi Company is 0.

Therefore, although the audit opinion did not reflect the risk of financial fraud, through the analysis of the company's statements and governance structure data, Langqi Company has a higher risk of fraud.

2.3. The Construction of Credit Transition Matrix by CAFÉ System for China's Financial Market

Credit rating adjustment is one of the most important rating actions of credit rating agencies. Credit rating adjustment behaviors include upscaling, downscaling, and maintenance. In a certain period (inspection period), the result of credit rating agency's adjustment of the debt issuer's credit rating can form the debt issuer's credit migration path, which reflects the change in the debt issuer's credit quality, this is done through the so-called Credit Transition Matrix (Hull, 2018; Jarrow et al., 2004), in order to do so, we first discuss the mapping rules for the construction of the CAFÉ credit transition matrix with two parts of A) and B) below.

A) The mapping rules for the construction of the CAFÉ credit transition matrix

In order to maintain the necessary stability of the transition matrix of the CAFÉ Credit Assessment System, we process the construction of the rating transition matrix in the following way (see the specific mapping in Table 2):

1. Divide the default model into one-year and two-year periods, and conduct ROC verification to determine the final model;

2. Give different rating results based on the 1-year and 2-year default models;

	AAA-A	BBB	BB	B	CCC-C
AAA-A	71.55%	24.76%	3.24%	0.39%	0.04%
BBB	15.11%	69.19%	13.77%	1.76%	0.15%
BB	3.68%	22.62%	55.66%	16.52%	1.51%
B	2.05%	8.94%	36.30%	42.93%	9.74%
CCC-C	1.52%	5.05%	16.41%	39.39%	37.37%

3. Integrate the 1-year and 2-year rating results to give the initial rating. The rules are as follows:

Table 2: CAFÉ's initial rating mapping table based on 1 year and 2 year period data

Items	Credit Rating in 1 Year						
Credit Rating in 2 Years	AAA	AA	A	BBB	BB	B	CCC-C
AAA	A	A	A	BBB	BB	BB	CCC-C
AA	A	A	A	BBB	BB	BB	CCC-C
A	A	A	A	BBB	BB	B	CCC-C
BBB	A	BBB	BBB	BBB	BB	B	CCC-C
BB	BBB	BBB	BBB	BBB	BB	B	CCC-C
B	BB	BB	BB	BB	BB	B	CCC-C
CCC-C	BB	B	B	B	CCC-C	CCC-C	CCC-C

B) The CAFÉ's Credit Transition Matrix

Taking into account the limitation of data acquisition, the observation limit of CAFÉ on the sample of listed companies is limited to the period from 2014 to 2019. We have obtained 3,000 companies with complete annual financial information from 2014, and 4,500 companies in 2019. Taking the 31st of December, 2020 as the observation day, the number of default samples returned for one year was 44; the number of default samples returned for the second year was 89 (here, we emphasize that we have actually observed "three years of return The number of default samples in this report is 115", but in the construction and analysis of the credit transition matrix in this report, the reason is that other sample information needs to be improved when the default samples of three years backed back are used for feature extraction. Therefore, only the backed One-year and two-year default samples) (Jarrow et al., 2004).

Based on the use of the above mapping rules, after integrating the rating results with the one-year and two-year default models, we will observe the following basic conclusion: the basis of the credit transition matrix from rating AAA-A to CCC-C Basically with BBB as the center, this maintains good stability, thereby solving the problem of instability of the AAA, AA and A-level migration matrix (especially for emerging markets like China). Therefore, from the perspective of the credit transition matrix, if we classify AAA to A into a "AAA-A" category, after adjusting the mapping threshold, we have the following credit transition matrix results (see Table 3):

Table 3: Summary information of BBD 2015-2020 transition matrix

It can be seen from the results of the transition matrix that the AAA-A grades maintain better stability after merging, and the transition matrix also reflects the monotonic decrease in the stability of AAA-C at each level, which supports the relative stability of our CAFÉ system Stability issues, and thus play a role in supporting the industry.

Appendix A: The Definition of Credit Rating Grades for CAFÉ Risk Assessment System

The CAFÉ Assessment System consists of three credit rating grades and nine credit rating for entity’s credit risk. The different levels represent different credit risk levels as shown below:

Table: The definition of Credit Risk for CAFÉ Credit Risk Assessment System

Definition	Ratings	The assessment of the ratings
Investment Grade	AAA	Speculative companies have the best operating conditions, the highest quality of financial reports, and the highest security.
	AA	The company's business conditions are very good, the quality of financial reports is high, and the security is high.
	A	The company's operating conditions are relatively good, the quality of financial reports is high, and the security is high.
	BBB	The company's operating conditions and financial report quality are acceptable, and safety can be basically guaranteed.
Speculative Grade	BB	The company's operating conditions are relatively poor, the quality of financial reports is not high, and the security is not high.
	B	The company's operating conditions are poor, there may be some problems in the financial report, and safety cannot be guaranteed.
	CCC	The company's operating conditions are poor, or the quality of financial reports is low, and the security is poor.
	CC	The company's operating conditions are extremely poor, or the quality of financial reports is extremely low, and the security is extremely poor.
	C	The company's operations are basically difficult to maintain, and there is basically no security at all. level

The References

[1] Altman, EI., 1968.Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, 23(4): 589–609，1968.

[2] Altman, EI., 1983.Corporate Financial Distress. A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy. Wiley Inter-science. John Wiley and Sons.

[3] Altman, EI., G. Sabato, 2007.Modeling credit risk for SMEs: Evidence from the US market. Abacus, 43(3):332–357, 2007

[4] BOC (for Five ministries and commissions including the people's Bank of China), 2021: Notice on promoting the high-quality and healthy development of the credit rating industry in the bond market (Exposure Draft), March 28, 2021 （http://www.pbc.gov.cn/tiaofasi/144941/144979/3941920/4215457/index.html）.

[5] Dun and Bradstreet, 2014. Dun and Bradstreet on credit rating [EB/OL]. http://www.dandb.com/glossary/d-b-rating/, 2014.

[6] FICO, 2018.What is a credit score? https://www.myfico.com/credit-education/credit-scores/.

[7] Hull, J., 2017.Options, Futures and Other Derivatives, 10th edn., Pearson, New York, 2017.

[8] Hull, J., 2018.Risk Management and Financial Institutes, 5th edn., Pearson, New York, 2018.

[9] Jing, XC., D. Li, JH.Wang, 2003.current situation and development prospect of China's credit rating agencies. China Financial bimonthly, 21:47-48, 2003.

[10] Du, LH., 2017.Plan ahead: the internationalization challenge of Chinese rating agencies. Financial market research, 62 (7): 119-127, 2017.

[11] Zhang, H., 2018.Review and Prospect of the development of China's credit rating market. Financial development research, 29 (10): 29-352018.

[12] Feng, GH., et al., 2019. Principles and pragmatism of credit rating. China finance press, 2019.

[13] Baidu (for Baidu Encyclopedia),2020.The Introduction to the world's three major rating agencies, 2020: Baidu Encyclopedia (https://baike.baidu.com/item/%E4%B8%96%E7%95%8C%E4%B8%89%E5%A4%A7%E8%AF%84%E7%B

A%A7%E6%9C%BA%E6%9E%84/8368411?fr=aladdin）, 2020.
[14] Anderson, R., 2007.The Credit Scoring Toolkit: Theory and Pracice for Retail Credit Risk Management and Decision Automation, 1st edn., Oxford University Press, Oxford, UK, 2007.
[15] Chi, GT, SL. Yu, GX. Yuan, 2019. Facility rating model and empirical for small industrial enterprises based on LM test. Journal of Industrial Engineering and Engineering Management, 33, 170–181, 2019.
[16] Thomas, L, J. Crook, D. Edelman, 2017. Credit Scoring and Its Applications (2nd). SIAM, Philadelphia, PA.
[17] Witzling. D., 2016.Financial complexity: Accounting for fraud. Science, 352（6283）, 301–301, 2016.
[18] Yuan, GX., HQ. Wang,2019.The general dynamic risk assessment for the enterprise by the hologram approach in financial technology. Int. J. Financ. Eng. 6(1):1-48, 2019.
[19] Ministry (for Ministry of finance of the people's Republic of China), 2008. Accounting Standards for Business Enterprises No. 36 - disclosure by related parties, promulgated on June 18, 2008.
[20] Yuan. GX., 2019a. Introduction to CAFÉ Hologram Risk Assessment System, BBD Inc., September 2019.
[21] Yuan. GX., 2019b. The function manual of intelligence stone credit rating system v1.2.0, BBD Inc., October 1, 2019 (https://file.bbdcreditservice.com/M00/00/01/CiDISF_ 22umAEXWBAE42fNtcLOk381.pdf）.
[22] Yuan, GX., 2021. The CAFÉ Risk Assessment by Applying Hologram Approach for Chinese market. TGES, Renmin University of China, Feb.23,2021 (https://mp.weixin.qq.com/s/RwP6UTtk3hMF9gYkImqD7A).
[23] Jarrow. RA., D. Lando, SM. Turnbull, 2004.A Markov Model for the Term Structure of Credit Risk Spreads. Review of Financial Studies, 10(2): 481-523。
[24] Yuan, GX., YP. Zhou, CX. Yan, et al., 2020a.New method for corporate financial fraud early warning and risk feature screening: Based on artificial intelligence algorithm. Proceedings of the 15th (2020) China annual management conference (www.cnki. Com. CN), 709-724.
[25] Yuan, GX., HY. Liu, YP. Zhou, et al., 2020b. The extraction of risk features by applying stochastic search algorithm FOF under the framework of bigdata. Management science, 33 (6): 41-53
[26] Yuan, GX., YP. Zhou, D. Li, et al., 2020c. The framework for risk characteristic factor base on AI algorithms in Fintech[J]. Journal of Anhui University of engineering, 35 (04): 1-13.
[27] Hanley,JA., BJ. McNeil, 1983.A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 148(3): 839–843.
[28] Fawcett, T., 2006: An introduction to ROC analysis. Pattern recognition letters, 27(8): 861-874.
[29] Akaike, H., 1974. A new look at the statistical model identification[J]. IEEE Transactions on Automatic Control, 19(6):716-723.
[30] Schwarz, G., 1978. Estimating the dimension of a model[J]. The Annals of Statistics, 6(2):461-464.
[31] Geman, S., D. Geman, 1984: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Mathematical Intelligence, 6, 721-741.