Finance Research Letters 58 (2023) 104305

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Credit default prediction of Chinese real estate listed companies based on explainable machine learning

Yuanyuan Ma [a,b,*], Pingping Zhang [a,c], Shaodong Duan [a,b], Tianjie Zhang [a,b]

[a] *School of Economics, Northeastern University at Qinhuangdao, Qinhuangdao, 066004, China*
[b] *School of Business Administration, Northeastern University, Shenyang, 110819, China*
[c] *School of Management, Northeastern University at Qinhuangdao, Qinhuangdao, 066004, China*

ARTICLE INFO

Keywords:
Credit default
MD&A(Management discussion and analysis)
Investor sentiment
Explainable boosting machine (EBM) model
Feature importance
AdaBoost model
Distance to default (DD)

ABSTRACT

It is essential to accurately forecast the credit default of real estate businesses and provide interpretable analysis. The intrinsic interpretable glass-box model and the post-hoc black-box model are used to predict and explain the credit default status of China's real estate listed businesses. Chinese annual reports, stock bar investor remarks, financial indicators and Distance to Default (DD) are taken into consideration when forecasting credit default. The AdaBoost model and the intrinsic Explainable Boosting Machine (EBM) model are determined to have the best prediction results, respectively. We present the explainable prediction results to clearly understand the ranking of feature importance and the impact on the prediction results.

1. Introduction

Systemic risks in the financial system can be avoided by doing well in risk perception and default prediction analysis of financial markets with a "barometer" function. From 2021, China Evergrande has experienced a significant debt default, which started a chain reaction. Numerous well-known real estate firms have fallen behind on payments, which is extremely detrimental to the concerned parties. Thus, it is crucial to develop a more precise credit default warning system for companies that are listed on real estate exchanges and to offer concise justifications for the results. The latest research is to use financial and non-financial factors to make predictions (Donovan et al., 2021; Huang et al., 2022; Islam et al., 2022). According to the recently references, texts in social media, annual reports, online stock exchange forums, and other formats provide incremental information on business credit (Zhao et al., 2022). Including emotional tone of MD&A text can considerably increase financial firms' capacity to recognize crises. Additionally, it can greatly enhance the model's capacity to determine whether a company is in bankruptcy or not (Hajek et al., 2014; Huang et al., 2022; Zhang et al., 2022c). Investor sentiment obtained from stock bar reviews can exacerbate stock market volatility and is positively correlated with the annual productivity of enterprises (Qian et al., 2022; Zhang et al., 2022b).

Due to imperfect text information supervision, managers of listed companies frequently engage in "tone manipulation" for the text information of the annual report, which lowers the overall annual report preset (Donovan et al., 2021). The prevalence of illogical investment behaviors will be caused by an abundance of noise investors, and a significant number of these events may even result in a financial crisis (Gaies et al., 2022). According to research, market-based default distances can give early warning of financial distress before a business goes bankrupt (Dinh et al., 2021). The approach presented in this paper attempts to address this issue. Since

* Corresponding author.
E-mail address: ma_yuanyuan@163.com (Y. Ma).

https://doi.org/10.1016/j.frl.2023.104305
Received 31 March 2023; Received in revised form 16 July 2023; Accepted 6 August 2023
Available online 17 August 2023
1544-6123/

Management Discussion and Analysis(MD&A) and stock bar comment text reflect the information of listed companies from inside and outside the enterprise respectively, political risk at the enterprise level is negatively correlated with distance to default (DD), when making default prediction, besides considering the financial indicator, two kinds of text information and the distance to default indicator reflecting the market information can reflect the actual situation of the enterprise more comprehensively (Islam et al., 2022).

Machine learning models have a huge benefit over the glass box model in that they can learn models by processing more complicated data, ultimately leading to the development of more accurate model prediction skills (Huang et al., 2022). Decision makers have concerns about the validity of the prediction results because the machine learning model is an unexplainable black-box, so additional explanations such as Shapely additive interpretation, partial dependence graph, counterfactual interpretation, and individual conditional expectation graph(ICE) are required (Charlton et al., 2023; Goldstein et al., 2013; Qian et al., 2022; Zhang et al., 2022b). According to the summary of the aforementioned literature, there are two contradictory issues: first, using textual data from various sources, such as MD&A, can significantly increase the ability of listed companies to predict credit defaults; second, the machine learning models currently being used in the most recent research on credit default risk prediction are often not interpretable, which reduces the validity and applicability of the findings. This will be fatal for complex and high-risk financial default prediction models that could influence decision-maker behavior. Also, researchers discovered that the Explainable Boosting Machine (EBM), a glass-box model with strong interpretability, has a comparable predictive value to the current advanced models (Hegselmann et al., 2022; Sarica et al., 2022; Thimoteo et al., 2022; Zhou et al., 2020). In a word, compared to the glass box model, the black box model can achieve

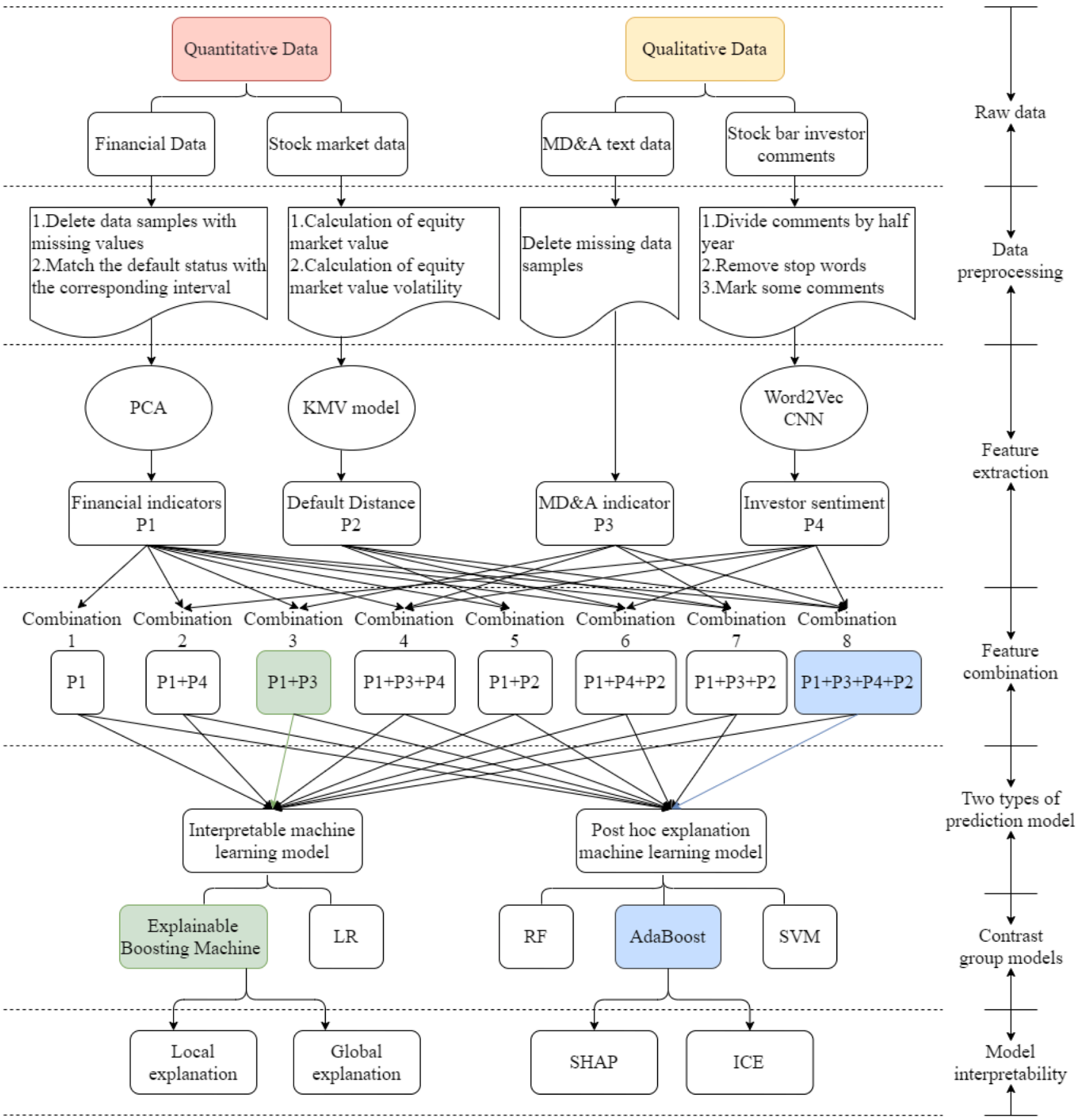


Fig. 1. Research framework of credit default prediction.

higher prediction accuracy. While the glass box model has inherent interpretability for default prediction and, in comparison to the black box model, can achieve equivalent accuracy to the advanced machine learning model. Consequently, research uses the glass box model and the black box model.

This study chooses the real estate firms in the Eastmoney.com as the research object and collects financial information, MD&A, stock bar investor remarks and distance to default characteristics for each company from 2017 to 2021. Word2vec is used to vectorize the stock bar text comments and a convolutional neural network (CNN) is utilized to categorize the vectorized text for computing investor sentiment indicators. The four characteristics are combined and used as the input of the glass-box model (Logistic Regression (LR) and Explainable Boosting Machine (EBM)) and the black-box model (Random Forest (RF), Support Vector Machine (SVM) and AdaBoost) respectively for credit default prediction. The best glass-box model and the best black-box model are identified, and the best input for each model to produce the best prediction effect is identified and then giving the local and global interpretation. Provide decision-makers access to both glass-box and black-box models so that stakeholders may make decisions that best suit their needs. In Fig. 1, the research procedure is displayed.

For the first time, this research gives decision-makers two options: a glass box model and a black box model. The study contrasts two different Explainable Artificial Intelligence (XAI) techniques for predicting credit default risk, which involve the following two steps: (1) Increasing the black box and glass box models' predictive accuracy; and (2) offering interpretable frameworks for the predictions of the two best models, respectively. The contributions of this research are reflected in three aspects. First off, the glass box and black box models suggested in the study offer a variety of solutions to satisfy the demands of stakeholders and decision-makers. Second, the research methodology used in this work can increase model prediction accuracy while providing results that are easy to understand by combining different input features. Finally, in order to show the reliability of the research methodology, empirical analysis on Chinese real estate listed businesses was done. This analysis offered a novel point of view for XAI to apply in default prediction.

The rest of this article is organized as follows. Section 2 reviews the relevant literature. Section 3 introduces the data and methods. Section 4 shows empirical analysis, and Section 5 depicts conclusion.

2. Research methodology

2.1. Multi-feature source data preprocessing

2.1.1. Distance to default data preprocessing

The KMV model is based on the traditional Merton model and the Black-Scholes option pricing theory. The calculation of KMV model can be divided into two parts: the calculation of asset value and its volatility and the calculation of distance to default.

KMV model calculates the company's asset value V_A and its volatility σ_A as follows Eqs. (1)-(6):

$$V_E = V_A N(d_1) - De^{-rT} N(d_2) \tag{1}$$

$$\sigma_E = \frac{N(d_1) V_A \sigma_A}{V_E} \tag{2}$$

$$d_1 = \frac{\ln\left(\frac{V_A}{D}\right) + \left(r + \frac{\sigma_A^2}{2}\right)T}{\sigma_A \sqrt{T}} \tag{3}$$

$$d_2 = d_1 - \sigma_A \sqrt{T} \tag{4}$$

$$DD = \frac{V_A - DP}{V_A \sigma_A} \tag{5}$$

$$DP = SD + \frac{1}{2} LD \tag{6}$$

Where V_E is the market value of enterprise equity; σ_E is the volatility of enterprise equity market value; r is the risk-free interest rate; T is the debt repayment period; $N(d_1)$,$N(d_2)$ is the cumulative probability function of normal distribution; DP represents the default point; DP is composed of the linear combination of short-term debt SD and long-term debt LD.

With reference to the setting of Chen (Chen et al., 2010), the equity value V_E is calculated as follows. According to the actual situation of China's securities market, the equity market value of listed enterprises can be divided into the market value of tradable shares (V_1) and the market value of non-tradable shares (V_2). The market value of tradable shares can be determined by multiplying the closing price by the number of tradable shares. This paper uses the semi-annual closing price S and the number of semi-annual tradable shares N_1 to show. The market value of tradable shares can be expressed as follows Eq. (7):

$$V_1 = S \times N_1 \tag{7}$$

The market value of non-tradable shares V_2 can be obtained by net profit per share, net assets per share, and basic earnings per share. BEP is for basic earnings per share for the half-year, BVP is for net assets per share, and NI stands for net profit per share for the half-year. Non-tradable share market value can be expressed as Eq. (8) and (9):

$$V_2 = \left(\frac{NI}{BEP} - N_1\right) \times BVP \tag{8}$$

$$V_E = V_1 + V_2 \tag{9}$$

Since the forecast in this study is made on a semi-annual basis, the forecast term T is chosen as 0. 5. The central bank's interest rate on one-year fixed deposits is used as the risk-free interest rate.

The volatility of the equity market value is calculated as follows Eq. (10). $\sigma_y$represents the simple moving average of the 60 day equity returns. The following equation calculates semi-annual volatility, where n is the number of trading days in a half-year.

$$\sigma_E = \sigma_y \times \sqrt{n} \tag{10}$$

2.1.2. Management discussion and analysis indicator

The Chinese Stock Market & Accounting Research Database (CSMAR) supplied the MD&A emotional indicators data. Every six months, the indicator is released. It contains information on the text's similarity to the previous year (TextualSimilarity), the number of positive and negative words (PositiveVocabularyNum, NegativeVocabularyNum), the number of sentences (SentencesNum), and the emotional tones 1 and 2 (EmotionTone1, EmotionTone2). EmotionTone1 is calculated by Eq. (11) and EmotionTone2 is calculated by Eq. (12). For the convenience of the following description, TextualSimilarity, PositiveVocabularyNum, NegativeVocabularyNum, SentencesNum, EmotionTone1and EmotionTone2 are respectively simplified to TS, PVN, NVN, SN, ET1, ET2

$$EmotionTone1 = \frac{PositiveVocabulary - NegativeVocabularyNum}{TotalWordsNum} \tag{11}$$

$$EmotionTone2 = \frac{PositiveVocabularyNum - NegativeVocabularyNum}{PositiveVocabularyNum + NegativeVocabularyNum} \tag{12}$$

2.1.3. Investor sentiment indicator data preprocessing

The research first uses Python crawler technology to crawl all the comments released by the real estate of the Eastmoney.com Stock Bar from January 1, 2017 to December 31, 2021.

The Chinese word segmentation component Jieba, which is frequently used in Python 3.7 software, is used to segment the comment text (Jang et al., 2019); In order to decrease the quantity of text that needs to be processed by the model, stop words, punctuation marks, spaces, meaningless words like "we," "are," and other bytes are eliminated from the comment text after word segmentation. Each word segmentation for the text vectorization processed by the aforementioned stage is converted into a particular word vector, wherein the skid-gram model in the word2vec is chosen. The selected window size of the model parameters is 3 and the dimension of the word vector is 150. It represents the vector of each word segmentation, namely$\omega_i = (\omega_{i1}, \omega_{i2}, \cdots, \omega_{i150})$, where$\omega_i$represents the i-th word segmentation in the document.

CNN is composed of input layer, convolution layer, pooling layer, full connection layer, and output layer. The following are the steps to acquire the classification outcomes of convolutional neural networks: Step 1: For each comment input, suppose the comment consist of m word segmentations. The comment can be represented as a m×150 word vector matrix, meaning that there will be N m×150 word vector matrices as input for a document containing N comments; Step 2 : The convolution layer's convolution kernel is combined with the input word vector matrix. Here, we pick four convolution kernels: 2×2, 3×3, 4×4 and 5×5. Here is the i-th convolution process as Eq. (13):

$$c_t = \mathrm{Re}LU(C \cdot W_{i:i+l-1} + b_0) \tag{13}$$

Where $W_{i:\, i+l-1}$ denotes the matrix of continuous row vectors from the i th word vector to the i + l-1 word vector in the m×150 word vector matrix; $c_t$indicates the degree of semantic proximity to the central word from the ith word to the i+l-1 word; C represents the convolution kernel, ReLU is a Rectified Linear Unit (ReLU), and b_0 is the deviation parameter.

Step 3: The maximum pooling method is used to the features. Step 4: The softmax function was utilized as the activation function in this study to categorize and the calculation equation is as follows Eq. (14):

$$softmax(D_j) = \frac{e^{D_j}}{\sum_j e^{D_j}} \tag{14}$$

Where $D_j$stands for the j-th node's output value and the result of the softmax function denotes the probability that the comment belongs to a particular class. The final classification outcome is determined by the model, which chooses the class with the highest probability.

The investor sentiment indicator is calculated as follows Eqs. (15)-(17):

$$posg = \frac{x_p}{x_p + x_n} \tag{15}$$

$$neg = \frac{x_n}{x_n + x_p} \tag{16}$$

$$ET = posg - neg \tag{17}$$

Where x_p represents the number of positive comments in half a year and x_n represents the number of negative comments. ET represents the mood tone of shareholders during the half year.

2.2. Interpretable machine learning methods

2.2.1. Intrinsic interpretable method -EBM model

Explainable boosting machine (EBM) model have been improved with the Generalized additive models (GAMs) (Caruana et al., 2015). The model's prediction accuracy is increased by including interaction terms (Lou et al., 2013), based on the original attributes, while also taking interpretability and good predictive ability into mind. The function goes like Eq. (18):

$$g(y) = \text{log}it^{-1}\left(\beta_0 + \sum_i f_i(x_i) + \sum_{i\neq j} f_{i,j}(x_i, x_j)\right) \tag{18}$$

Where $x_i$is the feature, $g(\cdot)$ is the connection function for classification, $\beta_0$is the intercept term, $f_i(x_i)$is the function used to quantify the contribution of each feature to the final prediction, and $\sum_{i\neq j} f_{i,j}(x_i,x_j)$is the function of the contribution of interaction features to the final prediction.

2.2.2. Post hoc explanation method-AdaBoost model

AdaBoost classification algorithm is an ensemble learning method. It repeatedly iteratively trains various classifiers on the same training set, then combines the inferior classifiers to modify the associated weight parameters to produce a stronger final classifier. Here is the formula for Eqs. (19)-(21):

$$f(x) = sign\left[\sum_{t=1}^{T} \log\left(\frac{1-e_t}{e_t}\right) f_m(x)\right] \tag{19}$$

$$e_t = E_w\left[1_{y\neq f_m(x)}\right] \tag{20}$$

$$\omega_i = \frac{\omega_{i-1} \times e^{-y_i \times f_m(x) \times \log\left(\frac{1-e_t}{e_t}\right)}}{\sum_{i=1}^{m} \omega_i} \tag{21}$$

Where m is the number of training samples $(x_1,y_1),(x_2,y_2),\cdots,(x_m,y_m)$,$x_i$ is the default influencing factor, $y_i$is the category label, T is the number of iterations, $f_m(x)$is the prediction result of the weak classifier and $\omega_i$is the adjusted classifier weight.e_t indicates the error rate.

In this work, the glass-box model's comparison analysis used the logistic regression. In order to establish the black-box model's best prediction outcomes, support vector machine (SVM) and random forest (RF) were chosen for comparison analysis with the AdaBoost model.

2.3. Explainable methods

It is challenging to clearly describe the forecast results for the black-box model. The post hoc explanation methods are mostly used in recent studies to explain the outcomes of predictions. This study examines the significance of features and the marginal relationship between features and output results using the SHAP algorithm and the individual conditional expectation diagram (ICE).

Shapley Addition Interpretation (SHAP), an algorithm based on Local Interpretable Model-agnostic Explanations (LIME) and Shapley values, is used to explain the contribution of each feature to the final output. It's an after-the-fact interpretation of the model. The following formula Eq. (22) can be used to predict the SHAP value of the model.

$$\hat{y}_j = shap_{base} + shap(X_{1j}) + shap(X_{2j}) + \cdots + shap(X_{pj}) \tag{22}$$

Where$\hat{y}_j$represents the prediction result of the j-th feature.$shap_{base}$represents the average value of all prediction samples.$shap(X_{ij})$refers to the SHAP value of the ith feature of the jth sample which reflecting the marginal contribution of the feature to the prediction result. The calculation formula of shap $(\cdot)$ is as follows Eq. (23):

$$shapley(X_j) = \sum_{K\subseteq M\{j\}} \frac{K!(p-K-1)!}{p!} (f(K\cup\{j\}) - f(K)) \tag{23}$$

$p$refers to the total number of features.$M\{j\}$representing all feature subsets.Krefers to a subset of $M\{j\}$.$f(K)$refers to the model prediction of features in K. $f(K\cup\{j\})$refers to the model prediction of featuresX_j and features in subset$K$. The final output shows the marginal contribution of the feature to the prediction.

Individual conditional expectation graph (ICE) is developed on the basis of partial dependence graph (PDP). PDP cannot illustrate the heterogeneity of each sample; it can only reflect the marginal effect of features on all samples. The marginal effect and sample heterogeneity are both easily visible with ICE.

3. Credit default prediction based on EBM and AdaBoost

3.1. Data preprocessing

3.1.1. Sample and data

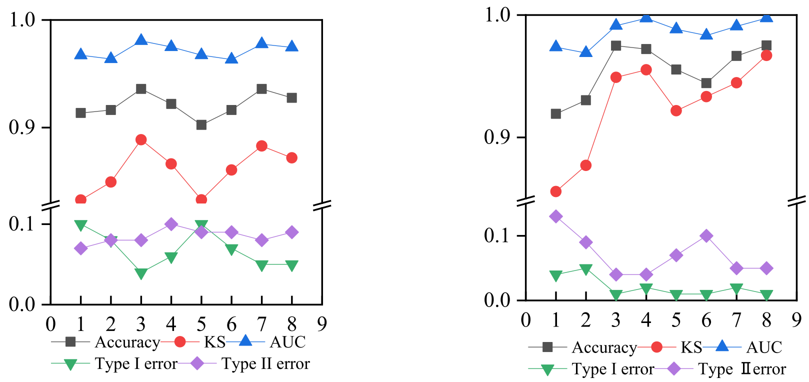
In this paper, Special treatment (ST) and entering the list of enforcement for trust-breaking are taken as the signs of credit default of listed companies. The annual report in the t range and the default in the t-1 range occur at the same time. As a result, a model is created using the data from the t-2 interval to forecast the likelihood of a credit default which will occur in the t interval. For the sample interval every half year from 2017 to 2020, 116 listed real estate businesses from Eastmoney.com were chosen. The final sample consists of 94 firms with 940 sample intervals, of which 44 are designated with ST and 896 are non-ST after the companies with missing data have been eliminated. For the problem of data imbalance, SMOTE algorithm is used in this paper.

The financial data of the enterprise in this article comes from the Resset Financial Database, the implementation of dishonesty and the management's discussion and analysis indicators come from China Stock Market Accounting Research(CSMAR) and the shareholder comments come from the stock bar of Eastmoney.com.

3.1.2. Feature extraction and combination

Due to the multicollinearity issue among financial indicators, principal component analysis was conducted on 24 financial indicators first (Chen et al., 2020; Zhang et al., 2022a). In the principal component analysis, Bartlett's spherical test is significantly 0 and the KMO value is 0.769, indicating that the correlation between variables is strong, which is suitable for the principal component analysis(PCA). Eight principal component factors are extracted in this study. The cumulative rate of variance, which is 67.418%, can capture the majority of the financial indicators. The first principal component (F1), according to the component matrix, primarily indicates the enterprise's profitability and earnings per share; The second principal component (F2) mainly reflects that the current ratio represents the solvency of the enterprise; The third main component (F3) mainly reflects the sales net interest rate represents the main business capacity of the enterprise; The fourth principal component (F4) mainly reflects the growth rate of operating profit, indicating the development ability of enterprises; The fifth principal component (F5) mainly reflects the turnover of shareholders' equity and represents the operating ability of the enterprise. The sixth principal component (F6) mainly reflects the accounts payable turnover rate and represents the enterprise's operating ability, the seventh principal component (F7) mainly reflects the operating revenue growth rate represents the enterprise's capital turnover level, and the eighth principal component (F8) reflects the total assets growth rate represents the enterprise's development ability.

This paper chooses listed real estate companies with and without ST from 2017 to 2020 to predict whether default will happen from 2018 to 2021. Make the ST minority sample and non-ST majority sample balance 896 after the SMOTE oversampling method. Use five indicators of AUC, KS, accuracy, type I error, and type II error to measure the predictive power of each prediction model. Investor sentiment characteristics, MD&A characteristics, and DD characteristics are added on the basis of principal components to observe the changes in the values of five test indicators in order to test the concept proposed above. Financial indicators serve as the foundation for creating P1 combinations when features are combined, namely: (1) Financial indicators (FI); To investigate the impact of adding a single text indicator on the predicted results, two combinations, P1+P4 and P1+P3, were generated, namely (2) FI and investor



(a) Comparison results of EBM combination indicators (b) Comparison results of AdaBoost combination indicators

Fig. 2. Comparison results of each combination indicator of the model.

sensitivity (3) FI and MD&A; To study the comprehensive effects of two emotional indicators, a combination of P1+P3+P4 was generated, namely (4) FI, MD&A, and investor sensitivity; Finally, the default distance indicator was added to all four previously generated combinations in order to analyze the influence of the default distance indicator on the prediction results. This made it simpler to compare the results with and without the default distance indication in each scenario. Including (5) FI and DD. (6) FI, investor sentiment, and DD (7) FI, MD&A, and DD (8) FI, MD&A, investor sentiment, and DD. Various combinations of these characteristics include the following: (1) Financial indicators(FI) (2) FI and investor sentiment (3) FI and MD&A (4) FI, MD&A and investor sentiment (5) FI and DD (6) FI, investor sentiment and DD (7) FI, MD&A and DD (8) FI, MD&A , investor sentiment indicators and DD.

3.2. Experimental result

3.2.1. Analysis results based on EBM model

Input the eight combinations constructed above into the EBM model to compare the accuracy, KS, AUC, type I error, and type II error of the prediction results. Fig. 2(a) shows the performance of eight combinations of EBM model on different indicators, and AppendixIIshows the corresponding values. AppendixIIdemonstrates that the model's highest prediction accuracy, KS value and AUC value are 93.59%, 88.88%, and 98.06%, respectively, and the lowest error rate of type I is 4% when the input feature is combination 3. The findings demonstrate that the addition of text characteristics enhances the model's capacity to classify data and make predictions, and that the MD&A is more useful than the investor comment text. If the decision-maker opts for the glass-box model to predict credit defaults, he can select the EBM model and take into account the company's financial features and MD&A text characteristics to get the best prediction outcomes.

3.2.2. Analysis results based on AdaBoost

Input 8 combinations into the AdaBoost model, comparing and analyzing the accuracy and classification ability indicators of the model prediction. Table 1 and Fig. 2 (b) describe the results of credit default risk prediction for several portfolios based on the AdaBoost model. Table 1 demonstrates that the model's type I error, classification ability, and prediction accuracy are all superior when portfolio 8 is used as the input feature. It demonstrates that incorporating financial features, MD&A, stock bar investor comments, and distance to default into the AdaBoost model's credit default test can maximize the model's ability to forecast outcomes. Fig. 2(b) shows that MD&A provides more information than investor sentiment on credit risk. Adding distance to default to financial indicators and investor sentiment can significantly improve the forecasting effect.

3.2.3. Contrast analysis result

For glass box model EBM and black box model AdaBoost, Logistic regression, RF and SVM were selected for comparative analysis. Table 2 lists the prediction results of the eight combinations made by the Logistic model, EBM model, RF, SVM, and AdaBoost. Table 2 shows that the prediction impact of the input EBM model is clearly superior to that of the logistic model. Black-box models are more accurate at predicting real estate enterprise credit default than EBM models.

When the input feature is combination 8, RF, SVM, and AdaBoost models have the highest accuracy and KS values. It demonstrates that the model's current prediction and classification capabilities are the best. AdaBoost model has superior KS value, AUC value, and Type I error than RF and SVM models. If the decision-maker opts for a black-box model to forecast real estate credit default, he can select the AdaBoost model and take distance to default, MD&A, investor comments, and financial characteristics into account while selecting input characteristics.

3.3. Model interpretability results

3.3.1. The SHAP and ICE interpretation results of AdaBoost

When applying the AdaBoost model to forecast four feature inputs Fig. 3(a) displays the importance ranking of the various factors. Fig. 3(b) illustrates how each significant feature affects the model output. Each line is a feature, ranked by feature relevance from top to bottom. Each point represents a sample, and the bluer the color, the lower the feature value. These two illustrations offer a global interpretation of the test set. The likelihood of a default is most strongly influenced by ET2 in MD&A, and the more positive it is, the less likely it is to happen. Furthermore, NVN's feature is crucial for model prediction. The bigger the feature, the simpler it is to default. Fig. 3(c) and (d) provide local explanations for the prediction findings of a single sample by illustrating the influence of various

Table 1
AdaBoost prediction results.

AdaBoost	Accuracy	KS	AUC	Type I error	Type II error
1	0.91922	0.855771	0.9737	0.04	0.13
2	0.93036	0.877093	0.969	0.05	0.09
3	0.97493	0.949130	0.9913	**0.01**	**0.04**
4	0.97214	0.955248	0.9974	0.02	**0.04**
5	0.95543	0.921774	0.9884	**0.01**	0.07
6	0.94429	0.933340	0.9833	**0.01**	0.1
7	0.96657	0.944621	0.9908	0.02	0.05
8	**0.97514**	**0.966982**	**0.9975**	**0.01**	0.05

Table 2
Prediction effect of each model in different combinations.

		1	2	3	4	5	6	7	8
Acc	LR	0.6825	0.7214	0.7409	0.7437	0.7242	0.7604	0.7855	**0.7939**
	EBM	0.9136	0.9164	**0.9359**	0.922	0.9025	0.9164	**0.9359**	0.9276
	RF	0.9164	0.9297	0.9304	0.9499	0.9276	0.9443	0.9471	**0.9526**
	SVM	0.9192	0.9276	0.9694	0.9721	0.9359	0.9415	0.9721	**0.9766**
	AB	0.9192	0.9304	0.9749	0.9721	0.9554	0.9443	0.9666	**0.9751**
KS	LR	0.3899	0.4530	0.5416	0.5656	0.4523	0.5284	0.5941	**0.6387**
	EBM	0.8328	0.8495	**0.8888**	0.8663	0.833	0.8607	0.883	0.8719
	RF	0.8550	0.8665	0.8885	0.9136	0.8648	0.8898	0.9058	**0.9286**
	SVM	0.8719	0.8806	0.9445	0.9493	0.9033	0.8995	0.9616	**0.9620**
	AB	0.8558	0.8771	0.9491	0.9552	0.9218	0.9333	0.9446	**0.9670**
AUC	LR	0.6894	0.719	0.8181	0.822	0.701	0.7472	0.8506	**0.8601**
	EBM	0.9672	0.9637	0.9806	0.9749	0.9673	0.9633	**0.9775**	0.9745
	RF	0.9753	0.9765	0.9782	0.9904	0.9741	0.9842	0.987	**0.9891**
	SVM	0.9801	0.9762	0.9948	**0.9963**	0.9771	0.9888	0.9947	0.9958
	AB	0.9737	0.969	0.9913	0.9974	0.9884	0.9833	0.9908	**0.9975**
I	LR	0.44	0.38	0.25	0.25	0.38	0.33	0.21	**0.19**
	EBM	0.1	0.08	**0.04**	0.06	0.1	0.07	0.05	0.05
	RF	0.06	0.09	**0.04**	0.05	**0.04**	0.05	0.05	**0.04**
	SVM	0.04	0.07	**0.02**	0.04	**0.02**	0.03	**0.02**	0.04
	AB	0.04	0.05	0.01	0.02	0.01	0.01	0.02	**0.01**
II	LR	**0.18**	0.19	0.27	0.26	0.19	0.15	0.22	0.22
	EBM	**0.07**	0.08	0.08	0.1	0.09	0.09	0.08	0.09
	RF	0.1	0.11	0.1	**0.05**	0.1	0.07	0.06	0.06
	SVM	0.13	0.07	0.05	**0.02**	0.1	0.08	0.03	**0.02**
	AB	0.13	0.09	**0.04**	**0.04**	0.07	0.1	0.05	0.05

Note : AB is AdaBoost model. I is Type I error, II is Type II error, Acc is Accuracy.

features in the two samples on the prediction results. The sample base value in Fig. 3(c) is 0.499, while the sample anticipated value is 0.907. F2 is the feature that most influence the prediction outcome for this sample. The sample base value in Figure (d) is 0.499. The anticipated value is reduced by the model from the initial value to 0.06. The PVN characteristic has the biggest impact on the sample, and it contributes 0.14 percent to the prediction of business default. In addition to the influence of a single feature on the prediction results of the model, the (e) graph studies the influence of the interaction between two features. The main influence is on the diagonal, and the interaction between features is outside the diagonal. It can be found that the feature interaction is not obvious.

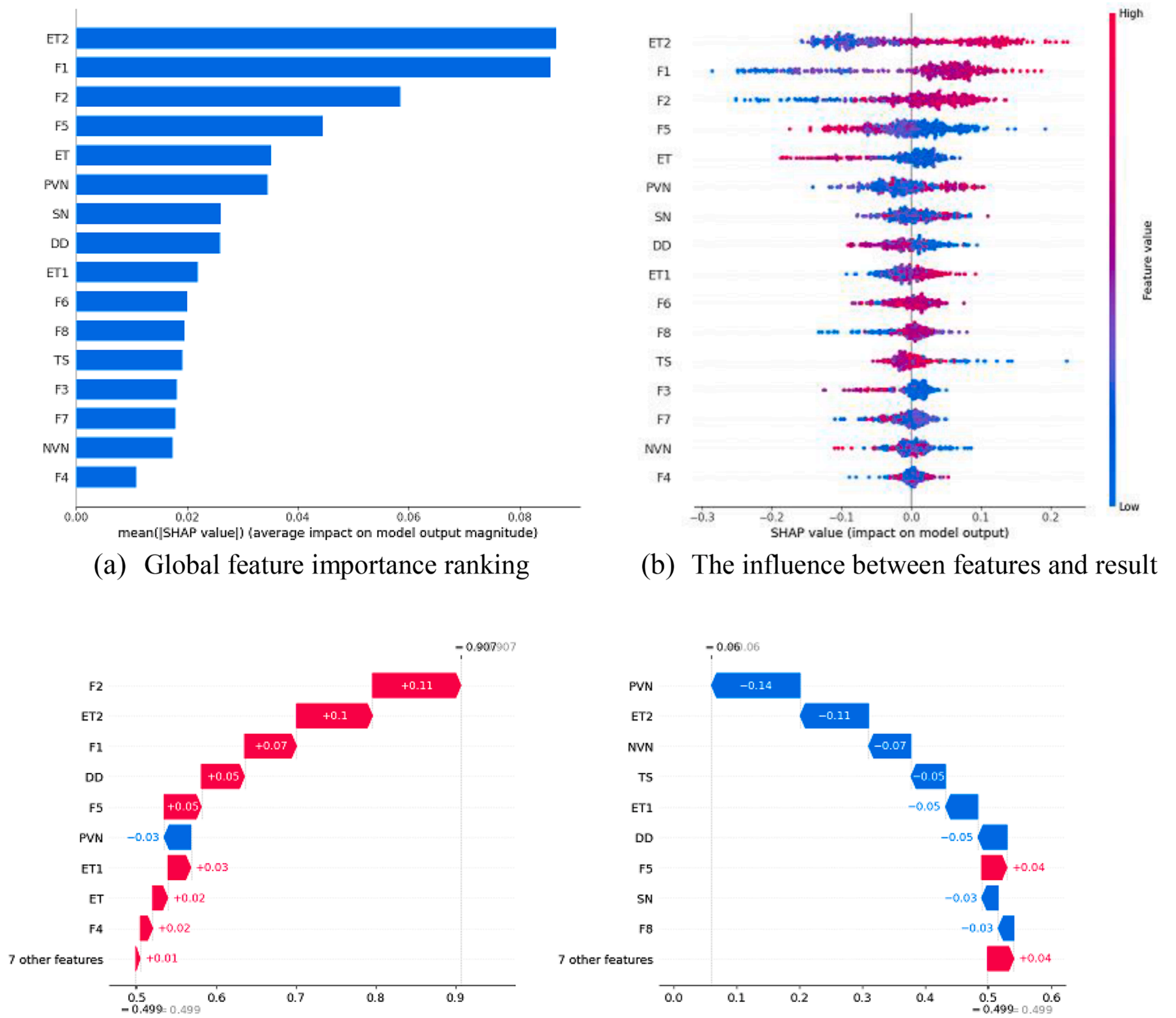
Fig. 3(f) shows the individual conditional expectation graph (ICE) of the test sample. It illustrates how the output result varies depending on the characteristic value for each instance sample. The EmotionTone2 feature is used as an example to demonstrate the circumstance of 300 test samples, where each blue line represents an instance sample and the orange dotted line the average result. The enterprise default probability of more than 0.14 decreases for the majority of samples using ET2. However, when ET2 did not reach 0.1, a few samples started to exhibit a trend of the default probability reducing initially and subsequently increasing. When ET2 exceeded 0.3, the default forecast changed slightly.

3.3.2. EBM local and global interpretation

The EBM model provides the best classification and prediction capabilities when taking into account financial indicators and MD&A features. Fig. 4 (a) illustrates the feature importance of each feature and its interaction item. It can be noticed that the first three significant characteristics are the interaction item of EmotionTone2 (ET2), F1×F2 and F5×ET2. MD&A is very significant for credit default prediction. The interaction diagram of F1 and F2 is shown in Fig. 4(b), with low F1 and low F2 or high F1 and high F2 indicating a higher likelihood of credit default. The relevance ranking of local features in a single sample and how it affects prediction outcomes are shown in Fig. 4(c). Among them, EmotionTone1 (ET1) feature contributed the most to the final prediction result, and this feature increased the risk of sample default. The global and local interpretability of EBM model can explain the importance of global features and the final predicted score of a single sample.

4. Conclusion

The relationship between social media sentiment and the stock market, the relationship between MD&A and credit risk, and the relationship between default distance and credit risk have all been taken into account in recent studies on credit default risk warnings (Islam et al., 2022; Qian et al., 2022; Zhao et al., 2022). However, no one has examined the relationship between investor sentiment, MD&A, default distance, and credit default risk. This article examines this relationship and looks at how each of the three features influences the final prediction results separately. For policymakers, estimates of credit default are extremely important in terms of accuracy and interpretability. This study predicts the credit default situation of listed real estate enterprises in China from two aspects of intrinsic interpretable model and blackbox model, and explains the results locally and globally. The findings of the empirical investigation are as follows. EBM model's predictive power is considerably superior to that of the Logistic regression model in the glass box model. When financial indicators and MD&A were used as input data, the EBM model exhibited the best prediction accuracy. The



(a) Global feature importance ranking

(b) The influence between features and result

(c) Local interpretation of sample 1

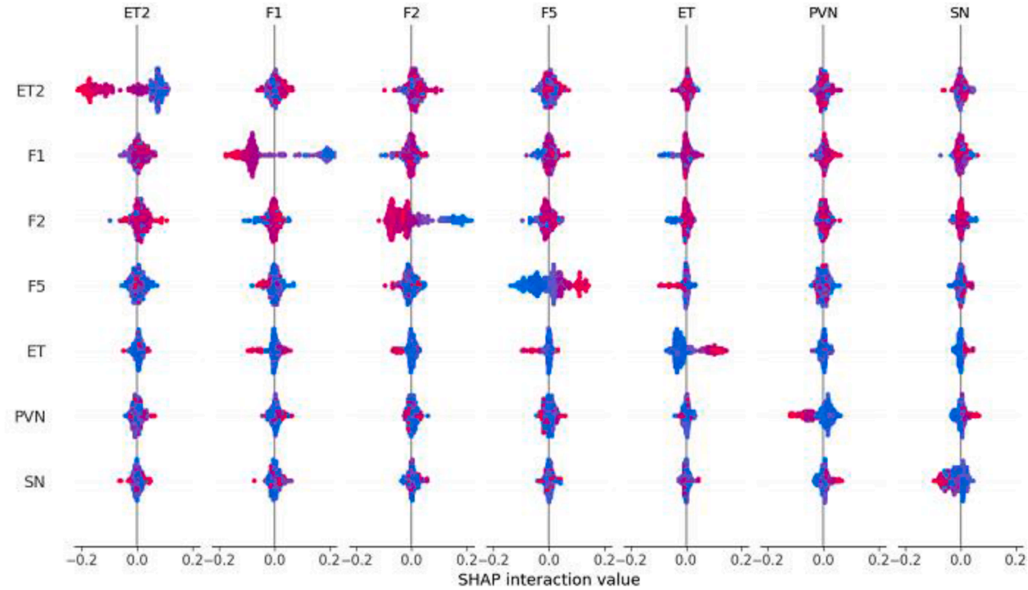
(d) Local interpretation of sample 2

Fig. 3. SHAP local and global interpretation diagram and ICE diagram of EmotionTone2.

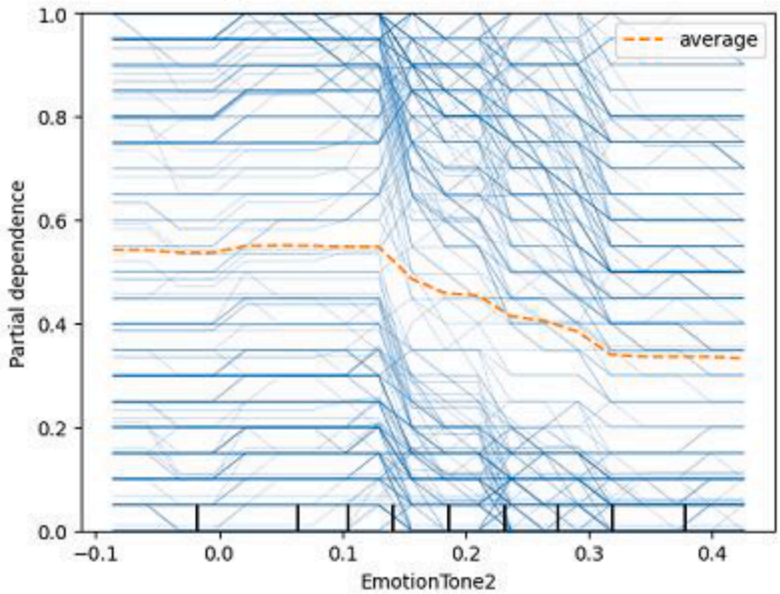
AdaBoost model offers the best predictive power among black box models. The model produces the best forecast when given four characteristics as input.

Using the AdaBoost model as an example, the accuracy increased with merely adding the investor sentiment, MD&A, and DD, from 81.34% to 84.4%, 88.86% to 82.73%, respectively. Compared to stock bar investor comments, MD&A offers more additional information about the credit default of listed real estate firms. The prediction accuracy of the model is 89.14% when MD&A and stock bar investor comments are added simultaneously, which is even better than when only one text indicator is added. This finding suggests that the two text information can complement one another in the prediction of credit default, which can lessen information hiding within the company and the manipulation of comments in the stock bar. When the distance to default is added into investor sentiment and MD&A, the prediction accuracy of the model is improved to 90.81%, indicating that the prediction accuracy of the model is the best when the three kinds of information are included.

This study examines more influencing elements than other studies on financial risk warnings. Compared to Islam's analysis of the connection between default distance and corporate political risk. In this study, the association between text sentiment and default risk is taken into account in addition to the relationship between default distance and default risk. The study demonstrates that using text sentiment can greatly increase prediction precision. This study takes into account the relationship between online stock forum's investor sentiment, MD&A, default distance, and financial indicators as compared to Zhao's research on the association between online stock forums sentiment, MD&A, financial statement sentiment, and financial distress. The AUC value of the ideal model can reach



(e) Interaction between features

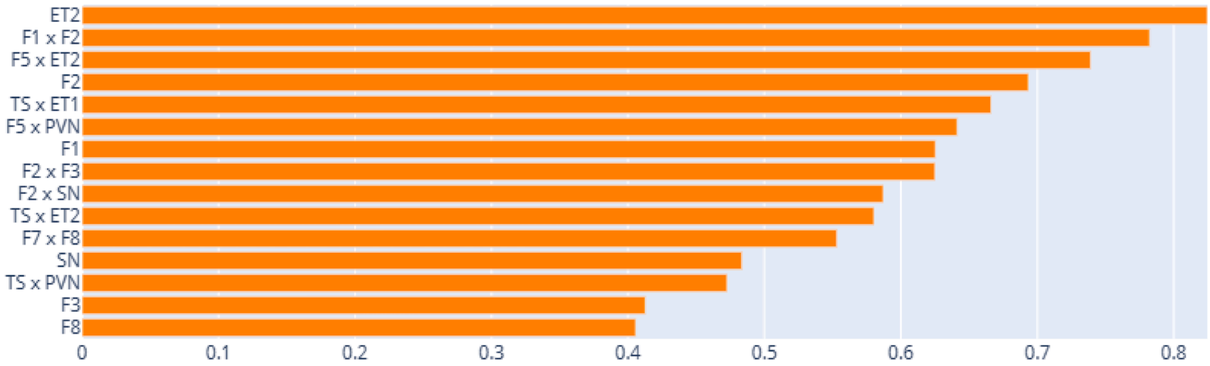


(f) Individual Conditional Expectation Diagram of EmotionTone2

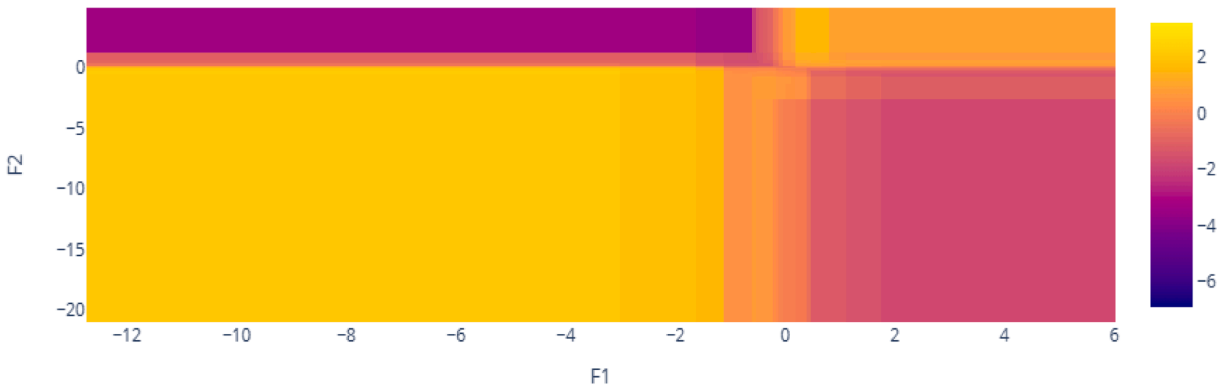
Fig. 3. (*continued*).

99.75%, and the model has better discriminative capability. Show that the many parameters put forth in this study contain additional information on corporate credit and that the SMOTE algorithm for addressing imbalances can boost the model's precision. This study took into account some variables, produced predictions using the black box and glass box models, and performed an interpretability analysis, giving decision-makers a variety of options that may be adapted to their requirements.

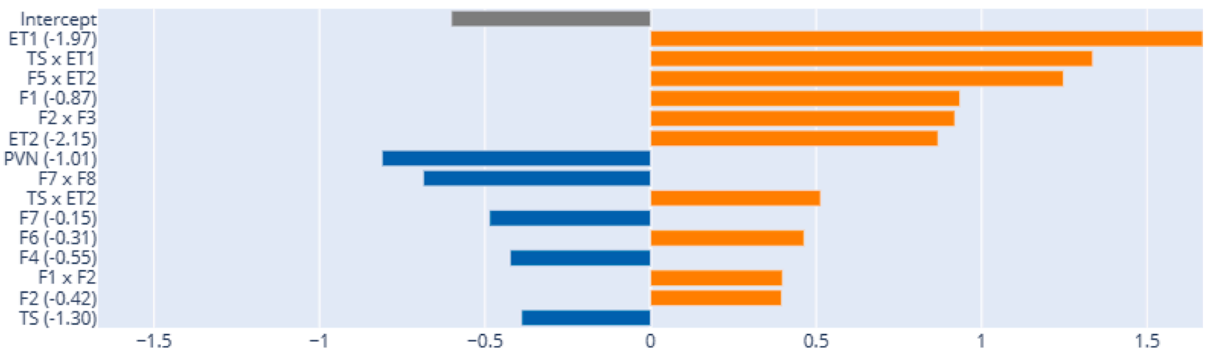
If decision-makers choose intrinsic interpretable model for prediction, they can choose EBM model whose input features are financial indicators and MD&A. After global interpretation of prediction results, it can be found that the importance of features affecting the results is ranked as EmotionTone2 (ET2), F1×F2 interaction term and F5×ET2. If decision-makers choose the black-box model for prediction, the AdaBoost model with input characteristics of financial indicators, MD&A, investor sentiment and DD can be selected. After global interpretation of the prediction results, it can be found that the feature importance that affects the final prediction result is ranked as EmotionTone2, EmotionTone1, F5, NegativeVocabularyNum, PositiveVocabularyNum, In the ranking of the importance of the two types of model features, EmotionTone2 is the most important factor, and F5 and F2 are more important than other factors. EBM model considers the relative importance of feature interaction items in importance ranking.



(a) Overall importance: mean absolute score



(b) F1×F2 interaction diagram



(c) Importance and influence of local features

Fig. 4. Global and local interpretation of EBM model.

In this paper, the study on credit risk prediction methods is expanded using the credit default of Chinese real estate listed businesses as the application environment. Machine learning algorithms were mostly employed for credit risk prediction in earlier studies. We introduced the glass box model for credit default risk analysis in response to stakeholders' distrust and skepticism regarding the high prediction accuracy of complicated machine learning algorithms. The research presented in this article offers experience and evidence for improving the precision of credit default prediction using XAI approaches. It also offers interpretable explanations for the findings, satisfying the demand for interpretable findings from decision-makers. This article can serve as an adequate basis for regulatory bodies to formulate regulatory measures and investors' financial decision-making behavior.

CRediT authorship contribution statement

Yuanyuan Ma: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing. **Pingping Zhang:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft. **Shaodong Duan:** Methodology, Validation, Writing – review & editing. **Tianjie Zhang:** Software, Validation.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Data availability

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (No. 71701036).

Appendix

Appendix I Financial indicators

First grade indicatores	Second indicator
Debt paying ability	Current ratio
	Quick ratio
	Cash ratio
	Debt ratio
Profitability	Return on assets ratio
	Operating profit ratio
	Return on net assets
	Net profit margin of total assets
	Earnings per share
	Net profit ratio
	Return on invested capital
	Gross operating interest rate
	Net operating interest rate
Operating capacity	Accounts receivable turning rate
	Inventory turning rate
	Total asset rate
	Accrued payable rate
	Working capital total rate
	Equity ratio
Development capacity	Total assets grow rate
	Operating income growth ratio
	Operating profit grow rate
	Net profit grow rate
	Equity year to date grow rate

Appendix II EBM prediction results

EBM	Accuracy	KS	AUC	Type I error	Type II error
1	0.9136	0.8328	0.9672	0.10	**0.07**
2	0.9164	0.8495	0.9637	0.08	0.08
3	**0.9359**	**0.8888**	**0.9806**	**0.04**	0.08
4	0.922	0.8663	0.9749	0.06	0.10
5	0.9025	0.8330	0.9673	0.10	0.09
6	0.9164	0.8607	0.9633	0.07	0.09
7	**0.9359**	0.8830	0.9775	0.05	0.08
8	0.9276	0.8719	0.9745	0.05	0.09

References

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). *Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission.* Paper presented at the ACM.

Charlton, C.E., Poon, M.T.C., Fleuriot, J.D., Brennan, P.M., 2023. Development of prediction models for one-year brain tumour survival using machine learning: a comparison of accuracy and interpretability. Comput. Methods and Programs in Biomed. 233, 17. https://doi.org/10.1016/j.cmpb.2023.107482.

Chen, C., Chen, C., Lien, D., 2020. Financial distress prediction model: the effects of corporate governance indicators. J. Forecast.

Chen, X.H., Wang, X.D., Wu, D.D., 2010. Credit risk measurement and early warning of SMEs: an empirical study of listed SMEs in China. Decision Support Syst. 49 (3), 301–310. https://doi.org/10.1016/j.dss.2010.03.005.

Dinh, D.V., Powell, R.J., Vo, D.H., 2021. Forecasting corporate financial distress in the Southeast Asian countries: a market-based approach. J. Asian Econ. 74, 18. https://doi.org/10.1016/j.asieco.2021.101293.

Donovan, J., Jennings, J., Koharki, K., Lee, J., 2021. Measuring credit risk using qualitative disclosure. Rev. Accounting Stud. 26 (2), 815–863. https://doi.org/10.1007/s11142-020-09575-4.

Gaies, B., Nakhli, M.S., Ayadi, R., Sahut, J.M., 2022. Exploring the causal links between investor sentiment and financial instability: a dynamic macro-financial analysis. J. Econ. Behav. Organ. 204, 290–303. https://doi.org/10.1016/j.jebo.2022.10.013.

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2013). Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. In.

Hajek, P., Olej, V., Myskova, R., 2014. Forecasting corporate financial performance using sentiment in annual reports for stakeholders' decision-making. Technol. Econ. Dev. Econ. 20 (4), 721–738.

Hegselmann, S., Ertmer, C., Volkert, T., Gottschalk, A., Dugas, M., Varghese, J., 2022. Development and validation of an interpretable 3 day intensive care unit readmission prediction model using explainable boosting machines. Front. Med. 9, 16. https://doi.org/10.3389/fmed.2022.960296.

Huang, B., Yao, X., Luo, Y.Q., Li, J., 2022. Improving financial distress prediction using textual sentiment of annual reports. Ann. Operations Res. 28. https://doi.org/10.1007/s10479-022-04633-3.

Islam, M.S., Alam, M.S., Bin Hasan, S., Mollah, S., 2022. Firm-level political risk and distance-to-default. J. Financ. Stability 63, 16. https://doi.org/10.1016/j.jfs.2022.101082.

Jang, B., Kim, I., Kim, J.W., 2019. Word2vec convolutional neural networks for classification of news articles and tweets. Plos One 14 (8), 20. https://doi.org/10.1371/journal.pone.0220976.

Lou, Y., Caruana, R., Gehrke, J., Hooker, G., 2013. Accurate intelligible models with pairwise interactions. In: Paper presented at the Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining.

Qian, H., Wang, B., Yuan, M., Gao, S., Song, Y., 2022. Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. Expert Syst. App. 190. Mar.

Sarica, A., Quattrone, A., Quattrone, A., 2022. Explainable machine learning with pairwise interactions for the classification of Parkinson's disease and SWEDD from clinical and imaging features. Brain Imaging and Beha. 1–11.

Thimoteo, L.M., Vellasco, M.M., Jorge, A., Karla, F., Lie, Y.C., Erito, M., 2022. Explainable artificial intelligence for COVID-19 diagnosis through blood test variables. J. Control, Automation and Electric. Syst. (2), 33.

Zhang, K., Wang, X.F., Wang, J.J., Wang, S.A., Hui, F., 2022a. Analysis and prediction of corporate finance and exchange rate correlation based on machine learning algorithms. Comput. Intelligence and Neurosci. 2022, 9. https://doi.org/10.1155/2022/2850604.

Zhang, Z.J., Wu, C., Qu, S.Y., Chen, X.F., 2022b. An explainable artificial intelligence approach for financial distress prediction. Inf. Process. Manag. 59 (4), 24. https://doi.org/10.1016/j.ipm.2022.102988.

Zhang, Z.S., Luo, M.T., Hu, Z.T., Niu, H.Y., 2022c. Textual emotional tone and financial crisis identification in chinese companies: a multi-source data analysis based on machine learning. Appl. Sci.-Basel 12 (13), 24. https://doi.org/10.3390/app12136662.

Zhao, S., Xu, K., Wang, Z., Liang, C., Lu, W., Chen, B., 2022. Financial distress prediction by combining sentiment tone features. Econ. Modell. 106, 105709. -.

Zhou, J., Tse, G., Lee, S., Liu, T., Wu, W., Cao, Z., …, Cheung, B., 2020. Identifying Main and Interaction Effects of Risk Factors to Predict Intensive Care Admission in Patients Hospitalized With COVID-19: a Retrospective Cohort Study in Hong Kong. Elsevier.