sustainability

MDPI

Article

Trust Evaluation Method of E-Commerce Enterprises with High-Involvement Experience Products

Kun Liang *, Jun He and Peng Wu

School of Business, Anhui University, Hefei 230601, China
* Correspondence: 17014@ahu.edu.cn

Abstract: Purpose: High-involvement experience products (HIEP) are generally characterized by a high value and difficult purchasing decision for customers, and a wrong decision will bring large losses to consumers, severely affecting their trust in enterprises. The purpose of this paper is to solve the problem of trust evaluation of HIEP e-commerce enterprises. Tasks and research methods: First, given the heterogeneity of trust information in the big data context, this paper collects the reputation data of HIEP enterprises and the trust big data of enterprises in industry, commerce and justice by a crawler program. Next, we use the dictionary and pattern matching methods to extract the trust features in text big data and construct the trust evaluation feature set integrating judicial information. Then, based on machine learning methods, we propose a LAS-RS model, which aims to solve the problem of trust evaluation in an imbalanced and high-dimensional data context. Finally, by introducing signal theory, this paper reveals the differential influence mechanism of big data feature variables on the trust of HIEP e-commerce enterprises. Originality: This study further enriches the relevant theories and methods of e-commerce trust evaluation research and is conducive to a better understanding and control of potential trust risks.

Keywords: e-commerce; trust; high-involvement experience products; signal theory

Citation: Liang, K.; He, J.; Wu, P. Trust Evaluation Method of E-Commerce Enterprises with High-Involvement Experience Products. *Sustainability* **2022**, *14*, 15562. https://doi.org/10.3390/su142315562

Academic Editors: Young-Woong Ko, Se-hak Chun, Moon Young Kang and Jinho Choi

Received: 28 October 2022
Accepted: 20 November 2022
Published: 23 November 2022

1. Introduction

Consumer trust can substantially improve customer satisfaction and loyalty and is the key to help e-commerce enterprises win market competition. The literature mainly studies the trust problem in e-commerce from the characteristics, types, influencing factors and calculation methods of trust. Liu et al. (2015) studied the domain and dynamic characteristics of trust and proposed a domain-aware dynamic trust network construction method by using a time-varying hypergraph model [1]. Isaeva et al. (2020) studied the types of trust from different dimensions. Specifically, trust is divided into cognitive trust, emotional trust and behavioral trust from the perspective of the source of trust; however, from the angle of trust level, trust is divided into institutional trust and interpersonal trust [2]. Verma et al. (2020) reviewed the factors influencing trust in e-commerce from the perspectives of content, emotion and the sender and receiver of online word-of-mouth and explored the degree, direction and moderating effect of these factors on trust [3]. Zhang and Yang (2021) proposed a trust calculation method for cross-border e-commerce using Bayesian networks and the machine learning method [4].

The product type and purchasing decision process can substantially influence consumer trust. For experiential products, such as clothing and skin care products, obtaining information and comparing product attributes are more difficult because consumers usually need to see, touch, smell or use them [5–7], and different consumers may form different opinions based on a single attribute of a product [8]. Therefore, for e-commerce companies with experience products, consumers have difficulty evaluating the credibility of products and companies through online word-of-mouth (OWOM) before purchasing. In addition, because of differences in product value, importance and consumer interest and knowledge, user purchasing decision costs are different. High-involvement experience products (HIEP) generally have the characteristics of a high value and difficult purchasing decision for customers, and a wrong decision will bring large losses to consumers, severely affecting the trust in enterprises. Therefore, trust evaluation of HIEP e-commerce enterprises is an important issue worthy of study, but few scholars have paid attention to it.

Developing big data and information technology provides new opportunities for the trust evaluation of HIEP e-commerce enterprises. In the big data environment, trust information not only exists in e-commerce platforms, but also may be distributed in different regions of cyberspace, such as industrial and commercial departments, judicial departments, supply chain enterprises and third-party social media platforms. The fusion of cross-platform trust information can effectively improve the authenticity and comprehensiveness of trust evaluation, but few scholars have paid attention to it. However, the trust evaluation of HIEP e-commerce enterprises based on big data also faces challenges, mainly in the following three aspects. First, in the big data environment, there are tons of unstructured trust data, such as user comments on e-commerce platforms, judgment documents on judicial websites and news reports on third-party social media. Extracting the appropriate feature variables from this unstructured information to evaluate the credibility of HIEP e-commerce companies is challenging. Second, the influence mechanism of big data feature variables on the trustworthiness of HIEP e-commerce enterprises is complex. Relevant theories must be introduced to reveal the differentiated effects of different types of big data features on trust to improve the interpretability and guidance of the trust evaluation model. Third, in the big data environment, the trust data of HIEP e-commerce enterprises are characterized by high dimensionality and imbalance. The feature set dimension of trust evaluation is high, including both structured and unstructured data. Furthermore, the quantity of trusted enterprises and untrusted enterprises presents an unbalanced distribution, which severely affects the evaluation model performance. Therefore, how to build the trust evaluation model of HIEP e-commerce enterprises in high-dimensional data and imbalanced data environments still needs further research.

Given the above research shortcomings and challenges, this paper collected the reputation data of HIEP e-commerce enterprises and the trust big data of enterprises in industry and commerce, judiciary and other aspects. On this basis, unstructured data analysis methods, such as dictionary matching and pattern matching, are used to extract trust features from text big data. By introducing signal theory, this paper explains the differential influence mechanism of big data feature variables on HIEP e-commerce enterprise credibility. The results show that the trust features of different signal costs have differential effects on the credibility of HIEP e-commerce enterprises. The higher the signal cost of the trust feature is, the stronger its ability to explain whether the HIEP e-commerce enterprise is trustworthy. Finally, we proposed a LAS-RS trust evaluation model for high-dimensional trust data and imbalanced trust data. The LAS-RS model is able to establish a broad, well-defined decision region for instances of the minority class. The effectiveness and robustness of the proposed model are verified by designing an experimental research scheme.

This study has important theoretical and practical values. Theoretically, we integrate big data analysis and signal theory into the study of trust evaluation of HIEP e-commerce enterprises, promoting the integration of theories, methods and achievements in related research fields, such as risk management [9], and providing theoretical and methodological guidance for scholars to study the credibility of HIEP e-commerce enterprises. In practice, the research results of this paper are helpful for consumers to have a more comprehensive understanding of the risks of e-commerce transactions; they are beneficial for HIEP e-commerce enterprises to handle the crisis of trust in a timely manner and improve market competitiveness; they are also conducive to industrial and commercial regulatory authorities to accurately grasp the business risks of HIEP e-commerce enterprises and scientifically formulate intervention measures.

2. Related Studies

2.1. Signal Theory

Signal theory was proposed by Spence in 1978 [10]. This theory is to study how to reach a transaction under asymmetric information through activities such as signal transmission and signal discrimination.

Enterprise trust evaluation mainly solves the problem of information asymmetry [4]. HIEP e-commerce companies clearly know their product situation and trust risks better than customers do. Therefore, companies transmit relevant information as signals to customers to help them make purchasing decisions under the condition of information asymmetry. For instance, the reputation and margin of an e-commerce enterprise can be used as signals for customers to judge whether the enterprise is trustworthy [2,11]. Here, the HIEP e-commerce enterprise is the party that sends the signal, and the consumer is the party that receives the signal. E-commerce enterprises send multiple signals to allow customers to understand the basic situation of products and companies. At the same time, customers can identify trustworthy merchants by identifying various obtained signals [12]. Müller et al. (2020) revealed how trust-enhancing signals influence buying intentions in B2C e-commerce [13].

Signal cost is an important measure of signal effectiveness [14]. Signals that are expensive to produce are more effective for decision making [15]. Rational signal senders do not generate, release or consume expensive signals for petty gain, especially when the cost of generating the signal outweighs the benefit of the signal. For example, a long-term accumulated reputation means expensive signaling costs. Studies show that e-commerce enterprises generally do not ruin their reputation to gain short-term benefits from fraudulent behavior [12]. This paper further studies the prediction ability of the big data features of different signal costs on HIEP e-commerce enterprises' credibility and uses signal theory to explain the results. Our study provides empirical evidence that big data features with high signal costs have better explanatory power for whether HIEP e-commerce enterprises are trustworthy, reveals the differential influence mechanism of cross-platform big data features on HIEP e-commerce enterprises' trust and improves the interpretability and guidance of the trust evaluation model.

2.2. Influencing Factors of Trust

Trust is mainly affected by the trust tendency of consumers and the trustworthiness of enterprises [2]. In the network environment, consumers and e-commerce enterprises have difficulty with communicating face-to-face in the transaction process. In most cases, consumers judge the credibility of e-commerce enterprises through online word-of-mouth. Verma and Dewani (2020) found that the content, sender, receiver and context of OWOM can affect the trust perception of e-commerce enterprises [3]. In terms of content, visual online word-of-mouth (such as pictures and videos) is generally considered to have a stronger effect on the trust perception of e-commerce enterprises than comments without any visual information [16]. When information senders and receivers have similar values, educational backgrounds and social statuses, online word-of-mouth is more likely to affect consumers' trust in enterprises [17]. Product type and platform type are the main contextual factors affecting consumer trust [18,19]. Consumer trust is often influenced by OWOM from a variety of platforms, including e-commerce platforms, online review sites and social networks. Word-of-mouth on social networks has a stronger impact on customer trust than comments from firm-owned channels [20].

In summary, the literature contains a certain amount of studies on the trust evaluation of e-commerce companies. However, word-of-mouth on e-commerce platforms and social media also has flaws, such as fake reviews, which lead to a need to further verify the reliability of trust evaluation based on word-of-mouth. In addition, few studies have focused on the trust evaluation of HIEP e-commerce enterprises. Therefore, this paper combines word-of-mouth information with credible judicial adjudication information in the big data environment, and the significance of building a HIEP e-commerce enterprise trust evaluation model based on cross-platform information is considerable.

2.3. Trust Evaluation Model

The trust evaluation of e-commerce enterprises can be considered a classification problem; that is, a classification model is established by using statistics and other methods to divide e-commerce enterprises into trustworthy and untrustworthy categories.

Common statistical methods include the discriminant analysis model, logistic regression model and decision tree model [21–23]. Statistical analysis methods generally require that the original data obey relevant statistical assumptions (such as normal distribution), and are usually effective only when the sample size is large [24]. With the deepening of research in the field of machine learning, intelligent algorithms have been applied to trust evaluation and decision making [25,26]. Morton's (2011) study shows that neural network (NN) models have significant advantages when there are complex nonlinear relationships between trust evaluation data [27]. The intelligent model loosens the statistical assumption of trust evaluation data, while the intelligent algorithm has the feature of a black box, and the interpretability is poor [28]. The combination method becomes the new trend of evaluation and prediction research. It enables multiple single models to form complementary advantages [29–31]. Studies have shown that the learning performance of combination methods is significantly higher than that of single statistical models and intelligent models [32].

A typical problem in the trust evaluation of HIEP e-commerce enterprises is the high-dimensional imbalance of data. Therefore, trust evaluation model construction must be studied in an imbalanced data environment. Imbalanced classification technologies contain a sampling method [33], cost-sensitive learning method [34] and ensemble learning method [35]. The sampling method balances the imbalanced dataset through oversampling and undersampling, which helps to improve the trust risk identification ability of the model [36]. Cost-sensitive learning (CSL) trains cost-sensitive classifiers by using minority class misclassification costs higher than majority class misclassification costs to solve the imbalanced classification problem from the algorithm level [34]. Although this method is helpful for the trust evaluation of e-commerce enterprises, the accurate misclassification cost in the application is difficult to determine [24]. It is worth mentioning that the ensemble learning method outperforms many single classifiers. However, how to design a rational integrated trust evaluation model for the high-dimensional imbalance of HIEP e-commerce enterprise trust data needs further research.

3. Research Framework

To cope with the research gap and challenges, the following research framework is proposed to construct a trust evaluation method suitable for HIEP e-commerce enterprises. The research framework consists of three parts, as shown in Figure 1. The first part aims to collect big data trust information about HIEP e-commerce enterprises from e-commerce platforms and third-party network platforms, mainly including the basic information of enterprises, the reputation data of e-commerce platforms and the data of industry and commerce and judicial departments. The second part includes trust feature extraction and time window determination, which aim to extract appropriate trust features from structured and unstructured data collected in the first part and determine the best time window. This approach is used because the influence of trust features on the credibility of HIEP e-commerce enterprises differs significantly between time windows. The third part proposes an LAS-RS model for the high-dimensional imbalance of trust data and compares it with several benchmark models. At the same time, according to the signal theory, the influence strength and mechanism of various features on the credibility of HIEP e-commerce enterprises are discussed in depth. Finally, robustness analysis is carried out on the trust prediction ability of the proposed model and various features.

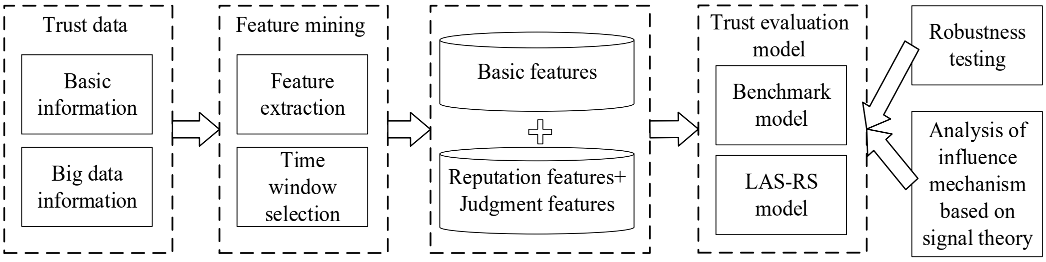

Figure 1. Research framework. Source: developed by the authors.

3.1. Trust Information Collection

This section mainly introduces the trust information of HIEP e-commerce enterprises contained in the judgment document, which is rarely considered in the literature. Then, in Section 3.2, we briefly introduce the enterprise basic information commonly used in the trust evaluation literature and the trust information contained in the e-commerce reputation system. We extract the administrative penalty information of industrial and commercial departments as the target variable to judge whether the HIEP e-commerce enterprises are trustworthy to consumers in their business activities.

3.1.1. Judgment Information

Judgment documents can effectively reflect disputes arising in the business process of HIEP e-commerce enterprises, which is of great help in judging their trustworthiness to consumers. Judgments can also affect a company's ability to gain user trust. For example, a court ordered an e-commerce enterprise to compensate the other party with CNY 10 million. This ruling will directly affect the funds the company spends on improving the quality of its products and services, which in turn affects consumer trust. The following four aspects are mainly considered in extracting trust features from judgment documents (See Appendix A Figures A1–A4).

1. Lawsuit status

Lawsuit status can be divided into plaintiffs and defendants. In general, defendant status may have a greater effect on the trust risk of HIEP e-commerce enterprises. We extract the lawsuit status information of enterprises according to trigger words. Commonly used negative trigger words include defendant, respondent, appellee, etc. Other trigger words indicate that the business is in plaintiff status.

2. Cause of action

Not all causes of action significantly affect the trust risk of HIEP enterprises. We select specific types of disputes as keywords to find causes of action related to consumer trust, such as false publicity and poor quality. We extract these specific types of causes of action based on legal dictionaries and pattern matching methods.

3. Number of lawsuits

If HIEP e-commerce enterprises face many lawsuits from multiple stakeholders in a short period, the enterprise may have serious trust risks. If lawsuits are only sporadic, then negligence in some aspects of management may be the cause, which does not necessarily affect consumer trust. To this end, we extract the number of lawsuits and study its relationship with the credibility of HIEP e-commerce enterprises.

4. Judgment result

Judgment results include winning and losing, and different results have different impacts on the trust risk of HIEP e-commerce enterprises. In addition, we extracted the keywords of receipt and payment information from the judgment result section. This approach is used because receipt and payment information directly affects the financial status of HIEP e-commerce enterprises, which in turn affects the quality of products and services and consumer trust. As with litigation status, we divide the verdict into winning or losing, depending on whether the verdict is in favor of the applicant.

3.1.2. Time Window Selection

In the big data environment, judgment information changes rapidly. HIEP e-commerce enterprises may be involved in multiple judgment documents within a period of time. However, judicial decisions in different time windows have different effects on consumer trust. In general, the influence of judgment documents on the credibility of e-commerce enterprises will gradually decay over time. Therefore, an appropriate time window must be chosen to evaluate the credibility of HIEP e-commerce enterprises and to obtain all the judgment documents during this period.

To select an appropriate time window, we calculate the time interval between the judgment and target variable (whether HIEP e-commerce enterprises have industrial and commercial administrative penalties). In our data, the maximum time interval is 5 years. Therefore, for the judgment information Ji (i = 1–5), a five-dimensional vector (J1, J2, . . . , J5) is defined, which represents the number of judgments of the HIEP e-commerce enterprises in the previous i years of industrial and commercial administrative punishment.

For example, for judgment information, there is a five-dimensional vector (2, 3, 5, 6, 8), where J1 = 2 and J2 = 3 indicate that two and three judgment documents have accumulated in one and two years before the enterprise is punished by industrial and commercial administration, respectively. The Chi-squared test was used to select the time window corresponding to the dimension most associated with consumer trust.

3.2. Trust Feature Set Construction

We use the data collected from the judgment documents, combined with the enterprise basic situation data commonly used in the literature and the e-commerce reputation system data, to establish a feature set to evaluate the credibility of HIEP e-commerce companies.

The big data environment contains unstructured text data, such as the cause of action description text in a judgment document. There are challenges in extracting appropriate trust features from these unstructured data. Because of the standard expression of judgment documents and the use of many professional terms to describe the process of litigation and judgment results, we combine the legal dictionaries and use related trigger words and rule templates to extract information such as lawsuit status, cause of action and judgment results and then design the relevant trust features.

Table 1 shows the established feature set, which mainly includes three types of trust features, namely basic features, e-commerce reputation features and judgment document features.

Table 1. Trust feature set of HIEP e-commerce enterprise.

Type	Features	References
FA: Basic features	Fa1 Registered capital Fa2 Tax credit rating Fa3 Amount of insurance Fa4 Age (years) Fa5 Number of insured	[37,38]
FB: E-commerce reputation features	Fb1 Total number of reviews Fb2 Number of positive reviews Fb3 Number of negative reviews Fb4 Authenticity of product information Fb5 Speed of logistics Fb6 After-sales service	[3,39]

Table 1. *Cont.*

Type	Features	References
	Fb7 Number of positive reviews with visual information Fb8 Number of negative reviews with visual information Fb9 Number of additional positive reviews Fb10 Number of additional negative reviews	
FC: Judgment document features	Fc1 Total number of lawsuits Fc2 Number of wins as plaintiff Fc3 Number of wins as a defendant Fc4 Number of lost cases as plaintiff Fc5 Number of lost cases as a defendant Fc6 Compensation Fc7 Compensation as a percentage of profit Fc8–Fc16 Number of disputes of various types	[37]

Source: systematized by the authors.

In the judgment document features (FC), we extract sixteen trust features based on the four aspects analyzed in Section 3.1.1. Among them, the total number of lawsuits (Fc1) reflects the overall situation of disputes arising in the operation of HIEP e-commerce enterprises, which potentially can indicate whether the enterprises are trustworthy. We further divide litigation into four types according to lawsuit status and judgment results and study the relationship between different types of litigation and enterprise trust (Fc2–Fc5). We believe that the compensation awarded in the judicial judgment can substantially affect the financial status of the enterprise, thus affecting the ability of HIEP e-commerce enterprises to gain consumer trust (Fc6 and Fc7). In addition, this paper intends to analyze the relationship between causes of action and consumer trust (Fc8–Fc16). The types we focus on include the dispute about construction projects (Fc8), commercial contract (Fc9), loan contract (Fc10), shareholders (Fc11), real right (Fc12), guarantee (Fc13), personality rights (Fc14), labor (Fc15), production and business operation (Fc16). We believe that not all causes of action will substantially affect the credibility of the enterprise. This paper selects special types of disputes as keywords to find causes of action related to consumer trust, such as false publicity and low quality. We extract these specific types of causes of action based on legal dictionaries and pattern matching methods.

3.3. Trust Evaluation Model Construction

The trusted HIEP e-commerce enterprises greatly outnumber the untrusted HIEP e-commerce enterprises [24]. Therefore, constructing a trust evaluation model involves classifying imbalanced data. Generally, imbalanced data classification methods include two categories. One type is the sample-level method, which resamples by means of oversampling and undersampling to balance the data distribution. Another type is the algorithm-level approach, which mainly addresses class imbalances by improving existing models or developing new algorithms. The ensemble approach is an effective algorithm-level approach when imbalanced datasets contain high-dimensional data [40,41].

Feature selection can transform data into a low-dimensional space and is an effective way to avoid misclassification of minority class instances. If the irrelevant features can be filtered out and the features that have a significant impact on the classification effect can be extracted, the performance of the ensemble model will be effectively improved. From this perspective, the Lasso method involves a shrinkage estimation method that promises to improve the accuracy of the ensemble method. Unlike most feature selection methods (each feature is evaluated separately), Lasso can reduce the global feature set during the construction of a trust evaluation model.

Our proposed LAS-RS method consists of three steps, as shown in Figure 2. Firstly, the original feature set is divided into several subsets through two important parameters: the subspace ratio r and the penalty parameter λ estimated by the Lasso. r determines the ratio of each feature subset to the global feature set, while the penalty parameter λ affects the shrinkage of the feature set. The feature weight is determined by Lasso estimation, which works as follows. To obtain a model, Lasso minimizes the sum of the squared residuals. The constraint condition is that the sum of the absolute values of the regression coefficients is less than a constant. Given a set of instances, the constraint condition is expressed as an instance matrix $D = \{(x_1, y_1), \cdots, (x_i, y_i), \cdots (x_n, y_n)\}^T$, where $x_i = \{x_{1,i}, x_{2,i}, \cdots, x_{c,i}, \cdots, x_{m,i}\}$ is the prediction variable, y_i is the state variable, n is the number of instances and m is the number of prediction variables. On the condition that the regression coefficient of the predictor does not exceed a threshold parameter on the L1 norm, a linear regression model with prediction variables is established. The observations are independent or y_i is conditionally independent on a specific $x_{c,i}$. Additionally, $x_{c,i}$ is standardized and satisfies $\frac{\sum_{i=1}^{n} x_{c,i}}{n} = 0$ and $\frac{\sum_{i=1}^{n} x_{c,i}^2}{n} = 1$. Therefore, Lasso estimation can be defined as a quadratic optimization problem according to the following conditions:

$$\underset{\gamma}{\operatorname{argmin}}\{\sum_{i=1}^{n} (y_i - \sum_{c=1}^{m} (\gamma_c x_{c,i}))^2\} + \lambda \sum_{c=1}^{m} |\gamma_c| \tag{1}$$

where γ_c is the set of regression coefficients of the prediction variables x_c, and λ is the penalty parameter that controls the shrinkage degree. We determine the degree of association between feature x_c and class label y_i by Lasso estimation. Then, we assign a set of importance scores $y = \{\gamma_1, \gamma_2, \cdots, \gamma_c, \cdots, \gamma_m\}$ for each feature. Finally, the feature weight w can be determined as $w_c = \frac{|\gamma_c|}{\sum_{c=1}^{m} |\gamma_c|}$. By controlling the parameters w and r, the features are randomly extracted, and the probability that feature x_c is extracted is equal to its weight (w_c). Therefore, several random feature subspaces formed by the original feature set can be expressed as $L_{sub}^j = \left\{(x_1^j, y_1), \cdots, (x_i^j, y_i), \cdots, (x_n^j, y_n)\right\}, \quad 1 \le j \le s.$

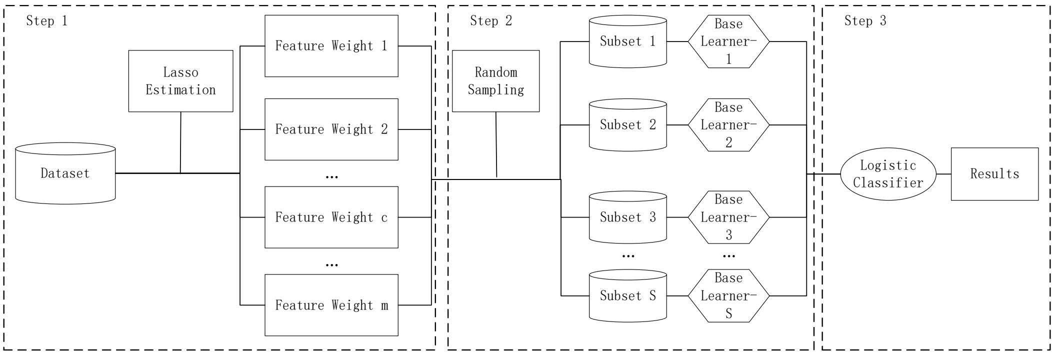

Figure 2. LAS-RS algorithm process. Source: developed by the authors.

The second step is to train the selected base classifier. Decision trees are widely used as one of the best classifiers for prediction because of their superior classification performance and good interpretability [42]. Therefore, we choose the C4.5 decision tree as the base classifier of LAS-RS.

The third step is to integrate the initial results of each base classifier. Commonly used integration strategies are voting methods and learning methods. Because of the lack of relevant expert knowledge to determine the voting weight of each base classifier, we adopt the learning method. Inspired by the stacking ensemble strategy, logistic regression was adopted as the meta-classifier, taking the output of each base classifier as the input of the meta-classifier. The results of the base classifier are learned to reduce the classification error of the ensemble model and improve the generalization ability.

The above contents are the three basic steps of the LAS-RS algorithm. Figure 2 shows the process of the LAS-RS algorithm.

4. Experimental Design

4.1. Experimental Dataset

Our data come from China's famous e-commerce platform JD.com, as well as the China Judgment Documents Network and "www.qcc.com". Among them, basic information about the registered capital, deposit amount and age of HIEP e-commerce enterprises is obtained from Jingdong Mall, and information, such as user reviews and ratings, is obtained from the e-commerce reputation system. In addition, we obtained the judgment document information about the HIEP e-commerce enterprise from the China Judgment Documents Network and linked the cross-platform big data with the Jingdong Mall data through the enterprise name. This paper selects jewelry e-commerce companies and high-end customized apparel e-commerce companies as representatives of HIEP e-commerce enterprises. We collected data from the above websites using a crawler program. Data from June 2015 to June 2020 for a total of 2586 e-commerce enterprises are included. After the stage of data preprocessing, 2532 e-commerce enterprises remained, including 1947 trustworthy e-commerce enterprises and 585 e-commerce enterprises with integrity problems. This paper uses the industrial and commercial administrative penalty information obtained from "www.qcc.com" as the target variable to judge whether HIEP e-commerce enterprises are trustworthy and divides the enterprises into two categories: trustworthy and untrustworthy. Among them, credible refers to the fact that the enterprise has not experienced industrial and commercial administrative penalties in its business activities in the past year. This paper uses the LAS-RS model and multisource trust information to predict whether HIEP e-commerce enterprises will experience a trust crisis in the operation process in the next year.

4.2. Evaluation Indicators

To measure and compare the performance of trust evaluation models, we use two commonly used metrics, namely average accuracy (AA) and area under the ROC curve (AUC). The untrusted enterprise is a positive class, and the trusted enterprise is a negative class. The average accuracy metric is defined below.

$$\text{AverageAccuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \tag{2}$$

By showing the binary classification results under different thresholds, the relationship between TP and FP can be seen, thus forming the ROC curve. The value of AUC is between 0 and 1, and the larger the AUC value is, the better the performance of the model is.

4.3. Experimental Procedure

We selected several benchmark models for comparison with LAS-RS, including sample-level methods (oversampling, OS; undersampling, US; SMOTE) and algorithm-level methods (cost-sensitive methods, CSL; ensemble methods). To implement the above benchmark approach, WEKA modules, such as SMO, Bagging, ADBoostM1, Stacking and CostSensitiveClassifier, were selected. We verified the performance of the proposed LAS-RS model and each benchmark model under different types of feature sets and feature set combinations to reflect the explanatory ability of various features for the credibility of HIEP e-commerce enterprises. To reduce the negative effects of the changes in the training set, we conducted ten tenfold cross-validation experiments.

5. Results and Discussion

5.1. Time Window Selection

To select the best time window, we study the relationship between judgment information in different periods and the credibility of HIEP e-commerce enterprises at observed time points. The Chi-squared test method was used to calculate the correlation between judgment information and HIEP e-commerce enterprise trust status in different time windows. The results are shown in Table 2. The time window "Having judgment within 2 years" obtains the maximum Chi-squared value, which indicates that the judgment information within 2 years has the strongest impact on the credibility of HIEP e-commerce enterprises. This result shows that in the short term, the influence of judicial judgment on enterprise credibility has a lag, and it takes some time (one year later) to appear; in the long run, the influence of judicial decisions on enterprise credibility will gradually weaken. We did not conduct time window analysis on reputation information (FB) and enterprise basic information (FA). Given that the online reputation data change in real time and their impact on trust is instantaneous, all the data before the observation time point need to be collected. However, enterprise basic information is generally time-independent, so time window analysis was not performed.

Table 2. Results of the Chi-squared test for different observation periods.

Observation Period	Number of Judgment	Chi-Value	*p*-Value
1 year	20711	0.000	0.997
2 years	38928	0.963	0.328
3 years	47154	0.205	0.642
4 years	52212	0.132	0.716
5 years	57905	0.098	0.763
6 years	61532	0.086	0.782

Source: calculated by the authors.

5.2. High-Dimensional Feature Selection

We use the Lasso method to select subfeature sets that are closely related to the credibility of HIEP e-commerce enterprises from the high-dimensional feature set. Feature selection can effectively reduce the complexity of machine learning models and improve the generalization ability of models.

In terms of enterprise basic information, the five basic features (Fa1–Fa5) all significantly affect the consumer trust of HIEP e-commerce enterprises. Therefore, we use them as the basic control variables.

In terms of reputation information, the total number of reviews (Fb1) and the number of positive reviews (Fb2) have no significant impact on consumer trust, which coincides with the research of Verma and Dewani (2020) [3]. This result is possibly due to some positive reviews being inspired by the interests of merchants rather than the hearts of consumers. In addition, there are many default positive reviews on e-commerce platforms. However, the number of negative reviews (Fb3) and the authenticity of product information (Fb4) significantly affect the credibility of HIEP e-commerce companies. This result may be due to the greater likelihood that negative reviews and the authenticity of product information reflect the real consumer experience. This conclusion is also consistent with the research of Liang et al. (2017) [39]. However, Liang et al. (2017) focused on ordinary e-commerce enterprises, while we focused on HIEP e-commerce enterprises. Notably, logistics speed (Fb5) does not have a significant effect on consumer trust, which may be because consumers of HIEP e-commerce enterprises pay more attention to the quality of products and services, but have low requirements on logistics speed. Therefore, in our study, after-sales service (Fb6) has a more significant effect on consumer trust. At the same time, we found that for HIEP e-commerce companies, customer reviews with visual information (Fb7 and Fb8) have a significant effect on consumer trust. This result is consistent with the analysis conclusion of Verma and Dewani (2020) [3] on general e-commerce enterprises, but Verma and Dewani (2020) [3] only provide a theoretical analysis and lack an empirical test. Finally, we reveal that additional reviews (Fb9 and Fb10) significantly affect the credibility of HIEP e-commerce enterprises. This result may be due to high-involvement experience products having the characteristics of complex product knowledge, strong professionalism and a long problem exposure cycle, and the follow-up review can reflect the consumer's trust in the product after using it for a period of time.

In terms of judgment information, not all judicial judgment information have been found to significantly affect the financial trust of enterprises [37]. Inspired by this finding, we divide the total number of lawsuits (Fc1) into four categories from the perspectives of HIEP e-commerce enterprises' position in judgment (including plaintiff and defendant) and judgment results (including winning and losing), that is, the number of wins as plaintiffs (Fc2), the number of wins as defendants (Fc3), the number of lost cases as plaintiffs (Fc4) and the number of lost cases as defendants (Fc5). The results show that only the number of lost cases as a defendant (Fc5) has a significant impact on enterprise credibility, possibly because these cases often affect a company's reputation and perceived trust. In addition, companies usually need to pay a certain amount of legal compensation (Fc6) when they lose a lawsuit, which may affect the enterprise's operating funds that could be used for the quality of products and services improvement and ultimately affect consumer trust. However, because of the differences in the profitability of different enterprises, the same compensation will have different impacts on different enterprises [37]. For example, CNY 100,000 may not be worth mentioning for a company with strong profitability and will not affect its operating capital and consumer trust, but may lead to bankruptcy for a company with weak profitability, seriously affecting its consumer trust. Therefore, the ratio of litigation compensation to the total annual profit of a company may have a stronger explanatory power for the credibility of the company. However, many e-commerce companies do not disclose their total annual profits. Using tax credit rating as a proxy variable of total annual profits, we find that the compensation as a percentage of profit (Fc7) has a significant relationship with the consumer trust of HIEP e-commerce enterprises. We further explore the relationship between the causes of action (Fc8–Fc16) and the credibility of HIEP e-commerce enterprises. Nine common causes of action are extracted from the judgment document by using regular expressions and legal dictionaries. Through the Lasso method, we explore the correlation between the number of disputes and consumer trust for the above nine types of lawsuits. The result shows that the causes of action related to commercial contract disputes (Fc9) and production and business operation disputes (Fc16) significantly correlate with the credibility of HIEP e-commerce enterprises.

5.3. Trust Evaluation Results and Discussion

5.3.1. Findings

We further investigate the prediction ability of various features and methods on the credibility of two types of HIEP e-commerce enterprises. The conclusions are shown in Tables 3 and 4.

Table 3. Average accuracy of various features and methods (jewelry e-commerce enterprises).

Feature Sets	FA	FA + B	FA + C	FA + B + C
SVM	71.23	73.18	79.16	80.09
OS	78.57	80.39	88.61	90.66
US	72.65	74.58	83.79	86.61
SMOT	78.46	81.22	90.21	90.87
CSL	78.38	81.96	89.53	91.15
Bag	85.21	87.86	90.50	93.14
Boost	86.34	88.31	92.40	93.27
Stacking	87.76	88.72	91.24	93.93
LAS-RS	88.32	89.79	93.45	96.76

Source: calculated by the authors.

Table 4. AUC of various features and methods (jewelry e-commerce enterprises).

Feature Sets	FA	FA + B	FA + C	FA + B + C
SVM	73.92	75.39	82.86	83.52
OS	81.35	82.40	89.75	87.36
US	68.43	70.83	84.58	85.72
SMOT	82.32	84.26	87.33	86.95
CSL	81.79	83.41	86.17	92.38
Bag	84.46	85.93	87.42	93.43
Boost	86.13	86.72	90.78	92.84
Stacking	88.63	89.32	92.29	93.02
LAS-RS	90.39	91.61	93.82	95.87

Source: calculated by the authors.

Tables 3 and 4 show the performance of various features and models in predicting the trustworthiness of jewelry e-commerce enterprises. From the perspective of features, the addition of e-commerce reputation features (FB) and judgment document features (FC) can effectively improve the performance of each trust evaluation method. In addition, compared with the reputation feature (FB), the judgment document feature (FC) has a more obvious improvement effect on the performance of the trust evaluation method. For example, the AA and AUC of the LAS-RS model proposed in this paper reached 93.45 and 93.82, respectively, after adding judgment document features (FA + FC), which was significantly higher than the model performance after adding reputation features (FA + FB) (AA and AUC were 89.79 and 91.61, respectively). From the perspective of the models, the overall performance of the imbalanced classification model is better than that of the benchmark SVM model, and the performance varies between imbalanced classification models. In general, the performance of sampling-based imbalanced classification models (OS, US, SMOT) and cost-sensitive learning (CSL) is inferior to that of ensemble models (Bag, Boost). The LAS-RS model achieves optimal performance based on improvements to RS. When FB and FC features were added, the AA of LAS-RS reached 96.76, while the AUC reached 95.87.

Tables 5 and 6 show the performance of various features and methods in predicting the trustworthiness of high-end customized clothing e-commerce enterprises. On the whole, the results are alike to that shown in Tables 3 and 4. That is, the imbalanced classification method outperforms the benchmark SVM method, and the ensemble method outperforms the sampling-based imbalanced classification method and cost-sensitive learning. LAS-RS also achieves the best trust prediction ability. In terms of the trust recognition ability of features, the signal cost of the FC is higher than that of the FB.

Table 5. Average accuracy of various features and methods (high-end customized apparel e-commerce enterprises).

Feature Sets	FA	FA + B	FA + C	FA + B + C
SVM	69.15	71.76	77.65	78.71
OS	73.31	76.90	86.44	88.50
US	69.68	70.80	82.42	85.14
SMOT	74.42	77.99	88.07	89.10
CSL	75.56	78.81	87.95	89.91
Bag	80.39	86.10	88.64	91.73
Boost	81.46	86.71	90.40	92.47
Stacking	83.28	86.48	90.05	92.96
LAS-RS	84.30	88.33	92.46	94.61

Source: calculated by the authors.

Table 6. AUC of various features and methods (high-end customized apparel e-commerce enterprises).

Feature Sets	FA	FA + B	FA + C	FA + B + C
SVM	70.69	74.23	81.51	82.03
OS	72.07	81.93	88.02	89.58
US	70.94	71.66	81.29	84.31
SMOT	76.26	82.07	86.18	88.64
CSL	78.38	81.91	86.86	90.47
Bag	79.50	84.34	87.95	92.51
Boost	81.34	85.08	89.92	91.76
Stacking	82.75	85.46	91.46	93.75
LAS-RS	85.60	89.69	93.59	95.92

Source: calculated by the authors.

To make sure the results were not accidental, we examined the significance of the AUC. To be specific, we performed a significance paired *t*-test at the a = 0.05 level on all results to show the statistical significance of each comparison. Table 7 shows the statistical tests of the differences in the trust prediction ability of different types of features. The results show that for jewelry e-commerce enterprises and high-end custom clothing e-commerce enterprises, integrating FCs has a significantly higher trust prediction ability than integrating FBs. Furthermore, incorporating the three types of features FA + B + C makes the model's trust prediction performance significantly better than what can be achieved using only FA + B and FA + C features. Table 8 shows the statistical test of the AUC differences in various methods in the state of fusing the three types of features FA + B + C. The conclusion shows that the imbalanced classification method obtains a significantly higher trust prediction performance than the benchmark SVM method. Among the many imbalanced classification methods, the ensemble methods (Bagging and Boosting) have better trust prediction performance than sampling-based methods (OS, US and SMOT) and cost-sensitive learning methods. The LAS-RS method significantly outperforms the other methods in terms of trust prediction performance.

Table 7. Statistical test of the AUC differences in various feature sets (* $p < 0.05$, ** $p < 0.01$).

Type	Jewelry			Apparel		
Methods	FA + B + C/ FA + C	FA + B + C/ FA + B	FA + C/ FA + B	FA + B + C/ FA + C	FA + B + C/ FA + B	FA + C/ FA + B
SVM	2.43 *	8.60 **	7.15 **	2.66 *	7.93 **	7.43 **
OS	3.62 **	11.63 **	10.70 **	3.31 **	11.26 **	9.09 **
US	3.08 **	14.53 **	13.15 **	5.08 **	13.04 **	10.41 **
SMOT	2.49 *	12.83 **	11.27 **	2.41 *	11.19 **	10.65 **
CSL	2.94 *	12.33 **	10.93 **	2.98 *	9.51 **	8.55 **
Bag	4.59 **	6.79 **	3.05 **	4.35 **	6.68 **	3.35 **
Boost	3.09 **	5.09 **	4.88 **	2.24 *	5.44 **	5.31 **
Stacking	3.50 **	6.93 **	4.81 **	4.84 **	6.21 **	3.38 **
LAS-RS	3.84 **	7.28 **	4.56 **	4.55 **	7.16 **	4.82 **

Source: calculated by the authors.

Table 8. Statistical test of the AUC differences in various methods (FA + B + C, * $p < 0.05$, ** $p < 0.01$).

Type	Jewelry	Apparel	Type	Jewelry	Apparel
OS/SVM	11.29 **	12.27 **	Bag/CSL	2.15 *	2.78 *
US/SVM	6.53 **	6.05 **	Boost/OS	5.06 **	4.85 **
SMOT/SVM	13.36 **	13.61 **	Boost/US	10.18 **	10.61 **
CSL/SVM	12.49 **	13.18 **	Boost/SMOT	3.66 **	2.69 *
Bag/SVM	13.40 **	14.47 **	Boost/CSL	3.58 **	3.75 **
Boost/SVM	15.36 **	16.79 **	Stacking/OS	4.49 **	3.57 **
Stacking/SVM	15.21 **	14.51 **	Stacking/US	9.36 **	9.20 **
Bag/OS	2.79 *	2.58 *	Stacking/SMOT	3.49 **	3.80 **
Bag/US	8.42 **	8.78 **	Stacking/CSL	3.61 **	3.44 **
Bag/SMOT	2.84 *	2.27 *	LAS-RS/Stacking	3.43 **	3.57 **

Source: calculated by the authors.

5.3.2. Discussion

We try to interpret and describe the findings in Section 5.3.1 from the perspective of signal theory. From the signal theory perspective, the results of Tables 3 and 4 were obtained because judgment information usually means high litigation costs, and the enterprise solves trust-related disputes through litigation only as a last resort. However, the cost of HIEP e-commerce enterprises to obtain a short-term (such as within one year) good reputation can be relatively low. Some enterprises even exploit the loopholes in the e-commerce reputation system to create the illusion of good reputation through strategic means such as collusion. Therefore, the signal cost of the FC is significantly more than that of the FB. According to the principle of signal theory, signal FC is more effective in reducing information asymmetry in commercial activities. Therefore, FC has a better and significant improvement effect on the prediction ability of the trust evaluation method. When the above two types of features are added simultaneously, the hologram of trust can be formed from a broader perspective, and the performance of the trust evaluation model reaches the optimal level. Similarly, for the results of Tables 5 and 6, according to the principle of signal theory, feature set FA + C is stronger than feature set FA + B in trust evaluation, and FA + B + C achieves the optimal performance.

6. Conclusions

This paper studies the trust evaluation of HIEP e-commerce enterprises. We obtain the trust data of e-commerce enterprises from multisource and heterogeneous big data platforms, such as the e-commerce reputation system, judicial judgment documents and industrial and commercial administrative departments, and propose an LAS-RS trust evaluation model aimed at the high-dimensional imbalance of trust data in a big data environment. At the same time, to clarify the deep influence mechanism of various features on trust, we introduce signal theory to explain the differential influence of e-commerce reputation features and judgment document features on the credibility of HIEP e-commerce enterprises from the perspective of signal cost.

This study has important theoretical value. First, this paper enriches the related theory and literature research on e-commerce trust evaluation. Most of the existing literature revolve around the trust evaluation of general e-commerce enterprises. From the perspective of product type and purchase decision-making process, this paper focuses on the trust evaluation of HIEP e-commerce enterprises, which is an innovation of the research object.

Second, this paper uses judicial judgment information to evaluate the credibility of HIEP e-commerce companies and further expands the data sources of e-commerce companies' trust evaluation. In addition, we introduce signaling theory to explore the differential impact of various characteristics on trust from the perspective of signaling cost. The results show that judgment features with high signal costs have a stronger impact on trust. Therefore, this paper expands the applicable scenarios of signal theory and promotes the integration of classical theories in related fields to trust evaluation research.

Third, this paper improves the RS-based trust evaluation model. To address the high-dimensional imbalance of e-commerce trust data, the LAS-RS model is designed. Additionally, the enterprise operation management data in some scenarios are also high-dimensional and/or imbalanced. Therefore, the LAS-RS model is also instructive for the literature research in related areas, such as predicting enterprise bankruptcy risk in the context of big data.

At the same time, this study also has important practical value. First, the research results of this paper help business administration departments strengthen the supervision of e-commerce enterprises; detect the signs of trust risks in the business process of enterprises in a timely manner; formulate relevant risk surveillance, forewarning and prevention measures; promote the healthy and sound development of the e-commerce market.

Second, the research conclusions of this paper provide direction for the operation and risk management of e-commerce platforms. The platform can build a better trust evaluation system for HIEP e-commerce enterprises. For example, publicly accessible judicial judgment information can be integrated into the e-commerce reputation systems so that consumers can have a more comprehensive understanding of the potential risks of the enterprise.

Third, the research results of this paper help consumers obtain more trust risk reminders about HIEP e-commerce enterprises, help consumers find trustworthy businesses, reduce consumer losses, improve consumer satisfaction and loyalty and improve the profitability and market competitiveness of enterprises.

This paper also has shortcomings. For example, we chose only two types of HIEP e-commerce enterprises. In the future, we will select more types of HIEP e-commerce enterprises to further verify the research conclusions of this paper. Furthermore, trust is related to the enterprises themselves and consumer attitudes. This paper mainly studies one of the aspects that affects trust, that is, trust data on the enterprise side. In the future, we will consider consumers' risk preferences and build a personalized consumer trust model.

Author Contributions: Conceptualization, K.L.; Methodology, J.H.; Software, P.W.; Validation, P.W.; Formal analysis, K.L.; Data curation, J.H.; Writing—original draft, K.L.; Writing—review & editing, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Science Foundation of the Ministry of Education of China (No.18YJC630082) and the Natural Science Foundation of Anhui Province (No.1908085QG307).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Some or all data and models that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors have declared that no conflict of interest exist.

Appendix A

山东省烟台市中级人民法院

执 行 裁 定 书

Intermediate People's Court of Yantai City, Shandong Province

Execute the judgment

（2020）鲁 06 执 6 号

(2020) Lu 06, No.6

原告：***股份有限公司，住所地：山东省烟台市。

Plaintiff: ***Co., Ltd, domiciled in: Yantai City, Shandong Province.

法定代表人：孙*。

Legal representative: Sun*.

委托代理人：赵*、章*，**律师事务所。

Authorized agent: Zhao*, Zhang*, **Law Firm.

被告：**股份有限公司，住所地：湖北省武汉市。

Defendant: **Co, Ltd., domiciled in: Wuhan City, Hubei Province.

法定代表人：王*

Legal representative: Wang*.

Figure A1. The first part of a judgment. Source: compiled by the authors.

Cause of action

原告**股份有限公司金融借款合同纠纷一案，山东省烟台市中级人民法院于2019年8月21日作出（2019）鲁06民初330号民事调解书，确认……

In the case of a dispute over a financial loan contract with the plaintiff ** Co., Ltd., the Intermediate People's Court of Yantai City, Shandong Province issued a civil mediation letter (2019) Lu 06 Min Chu No. 330 on August 21, 2019, confirming:……

The process of judgement

本院于2020年1月1日立案执行，2020年1月2日向被执行人送达执行通知书和报告财产令。……

The court filed the case for enforcement on January 1, 2020,, and served the enforcement notice and property reporting order on January2,2020…

2020年1月7日、2021年2月7日，本院经两次网络司法查控，被执行人在银行、证券机构、互联网银行、工商总局、不动产登记机关均无相应资产登记的反馈信息。……

On January 7,2020 and February 7,2021,the court conducted two online judicial investigations and found that the subject had no feedback information on asset registration:in banks, securities institutions, Internet banks, the State Administration for Industry and Commerce, and real estate registration authorities. ……

截止2021年2月7日，本案未受偿债权数额为贷款本金5000000000元，以及至借款实际付清之日止的利息（截至2019年4月20日，正常利息、罚息、复利合计13239716.14元，自2019年4月21日起至借款实际付清之日止仍按照合同约定计息），诉讼费1303999.5元，保全费5000元。案件执行情况已告知申请执行人，申请执行人也提供不出可供执行的财产线索。

As of February 7, 2021, the amount of outstanding debt repayment right in this case is RMB 500,000,000 principal of the loan, and the interest till the date of actual payment off the loan (as of April 20, 2019, the total amount of normal interest, penalty interest and compound interest is RMB 13,239,716.14. From April 21, 2019 to the date when the loan is actually paid off, interest will still be calculated according to the contract), legal costs of RMB 130399.5, preservation costs of RMB 5,000. The execution of the case has been informed to the applicant for execution, but the applicant can not provide any property clues for execution.

2021年2月7日，申请执行人提出书面申请，要求终结本次执行程序。不申请发布悬赏公告查找被执行人的财产，亦不同意将本案移送破产审查。

On February 7, 2021, the executor submitted a written application to terminate the execution procedure. He did not apply for a reward notice to find the property of the subject, nor did he agree to move the case to bankruptcy review.

Figure A2. The second part of a judgment. Source: compiled by the authors.

本院认为：申请执行人的申请符合相关法律规定，本案应依法终结本次执行程序。依照《最高人民法院关于适用〈中华人民共和国民事诉讼法〉的解释》第五百一十九条规定，裁定如下：

The court believes that the application for execution conforms to relevant laws and regulations, and the execution procedure should be terminated in accordance with the law in this case. In accordance with Article 519 of the Interpretation of the Supreme People's Court on the Application of the Civil Procedure Law of the People's Republic of China, the ruling is as follows:

终结本次执行程序。

Terminate this execution procedure.

申请执行人发现被执行人有可供执行财产的，可以再次申请执行。

If the person applying for execution finds that the person subjected to execution has property available for execution, he may apply for execution again.

本裁定送达后立即生效。

This ruling shall take effect immediately after it is served.

Figure A3. The third part of a judgment (results of the judgment). Source: compiled by the authors.

审判长	王*	Presiding judge: X, Wang
审判员	徐*	Judicial Officer: X, Xu
审判员	林*	Judicial Officer: X, Lin
二〇二一年二月八日		February 8, 2021
书记员	王*	Clerk: X,Wang

Figure A4. The fourth part of a judgment. Source: compiled by the authors.

Appendix A.1. Lawsuit Status

We select keywords (as shown in Table A1) related to lawsuit status from the legal dictionary to extract the lawsuit status of the applicant, which appear in the basic information of the parties (see Figure A1). We scan the file from the judgment code line to find the first place where the applicant's name appears. If we find a word written before the applicant's name in our word bag, we assume that the keyword represents lawsuit status.

Table A1. Keywords associated with lawsuit status.

Lawsuit Status	Keywords
Negative	Defendant, appellee, respondent, person subjected to enforcement, party against whom execution is filed
Non-negative	Plaintiff, appellant, applicant, execution applicant, retrial applicant, third party, claimant

Source: compiled by the authors.

Appendix A.2. Cause of Action

We use a pattern-based approach to extract the keywords for the cause of action from the second part of the judgment (see Figure A2). The first sentence of a cause of action is usually written as follows: the XXX dispute case with the plaintiff XXX accusing the defendant XXX. We treat the specific dispute as the keyword for cause of action. Therefore, we use regular expressions and legal dictionaries to find specific disputes. There are 1960 terms related to disputes in the predefined dictionary.

Appendix A.3. Judgment Result

This structured information is extracted from the third part of the judgment (see Figure A3). We extract the sentences between the two phrases "the sentence shown below:" and "the total fee ... the lawsuit fee ... the insurance fee ... ". Then, we split these sentences by the sequence number and only retain the items involving the applicant. After analyzing the content of the judgment results in our sample, we construct 24 regular expressions to match each item in the judgment results. The keywords shown in Table A2 are part of these regular expressions, and each keyword belongs to only one regular expression and vice versa. If an item matches one of these regular expressions, we use the keyword of that regular expression to represent the item.

Table A2. The keywords used in the description of the results of a judgment.

Judgment Result	Keywords
Negative	Payment, joint and several liability, frozen assets, detained assets, sealed up assets, returning goods, compulsory execution, issuing false invoice, the notice of discharging obligation
Non-negative	Receiving money, suspension of payment, no compensation shall be paid, preserved assets, release of frozen assets, release of the sealed assets, release of the detained assets, allowed to withdraw a lawsuit, allowed to withdraw an appeal, revoked a case placed on file, execution termination, lawsuit termination, revised the mistake in the judgment, affirmed the original judgment, the notice of assisting in enforcement

Source: compiled by the authors.

References

1. Liu, S.; Jiang, C.; Lin, Z.; Ding, Y.; Duan, R.; Xu, Z. Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach. *Inf. Sci.* **2015**, *306*, 34–52. [CrossRef]
2. Isaeva, N.; Gruenewald, K.; Saunders, M.N. Trust theory and customer services research: Theoretical review and synthesis. *Serv. Ind. J.* **2020**, *40*, 1031–1063. [CrossRef]
3. Verma, D.; Dewani, P.P. Ewom credibility: A comprehensive framework and literature review. *Online Inf. Rev.* **2021**, *45*, 481–500. [CrossRef]
4. Zhang, F.; Yang, Y. Trust model simulation of cross border e-commerce based on machine learning and bayesian network. *J. Ambient Intell. Humaniz. Comput.* **2021**. [CrossRef]
5. Huang, L.; Tan, C.H.; Ke, W.; Wei, K.K. Do we order product review information display? how? *Inf. Manag.* **2014**, *51*, 883–894. [CrossRef]
6. Jiménez, F.R.; Mendoza, N.A. Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. *J. Interact. Mark.* **2013**, *27*, 226–235. [CrossRef]
7. Singh, J.P.; Irani, S.; Rana, N.P.; Dwivedi, Y.K.; Saumya, S.; Roy, P.K. Predicting the "helpfulness" of online consumer reviews. *J. Bus. Res.* **2017**, *70*, 346–355. [CrossRef]
8. Hong, K.Y.; Pavlou, P.A. Product fit uncertainty in online markets: Nature, effects, and antecedents. *Soc. Sci. Electron. Publ.* **2014**, *25*, 328–344. [CrossRef]
9. Lagodiienko, V.; Popelo, O.; Zybareva, O.; Samiilenko, H.; Mykytyuk, Y.; Alsawwafi, F.M.A.S. Peculiarities of the Management of the Foreign Economic Activity of Enterprises in Current Conditions of Sustainability. *Int. J. Sustain. Dev. Plan.* **2022**, *17*, 1215–1223. [CrossRef]
10. Spence, M. Job market signaling—Sciencedirect. *Uncertain. Econ.* **1978**, *87*, 283–306.
11. Zhang, S.; Zhong, H. Mining users trust from e-commerce reviews based on sentiment similarity analysis. *IEEE Access* **2019**, *7*, 13523–13535. [CrossRef]
12. Leonard, L.N.; Jones, K. Trust in C2C Electronic Commerce: Ten Years Later. *J. Comput. Inf. Syst.* **2021**, *61*, 240–246. [CrossRef]
13. Müller, S.; Schwehm, S.; Hewig, J.; Paelecke, M. Influence of Trust-Enhancing Signals on Signal Processing in B2C E-Commerce. In Proceedings of the AMCIS 2020, Virtual, 10–14 August 2020; Volume 6.
14. Bird, R.; Smith, E. Signaling theory, strategic interaction, and symbolic capital. *Curr. Anthropol.* **2005**, *46*, 221–248. [CrossRef]
15. Aggarwal, R.; Gopal, R.; Gupta, A.; Singh, H. Putting money where the mouths are: The relation between venture financing and electronic word-of-mouth. *Inf. Syst. Res.* **2012**, *23*, 976–992. [CrossRef]
16. Lin, T.M.Y.; Lu, Y.; Wu, J. The effects of visual information in eWOM communication. *J. Res. Interact. Mark.* **2012**, *6*, 7–26. [CrossRef]
17. Ismagilova, E.; Slade, E.; Rana, N.P.; Dwivedi, Y.K. The effect of characteristics of source credibility on consumer behaviour: A meta-analysis. *J. Retail. Consum. Serv.* **2020**, *53*, 101736. [CrossRef]
18. Tsao, W.-C.; Hsieh, M.-T. eWOM persuasiveness: Do eWOM platforms and product type matter? *Electron. Commer. Res.* **2015**, *15*, 509–541. [CrossRef]
19. Cheung, C.M.; Thadani, D.R. The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decis. Support Syst.* **2012**, *54*, 461–470. [CrossRef]
20. Cao, C.; Yan, J.; Li, M. The Impact of Different Channel of Online Review on Consumers' Online Trust. In Proceedings of the PACIS 2018, Yokohama, Japan, 26–30 June 2018; p. 213.
21. Lee, T.S.; Chiu, C.C.; Chou, Y.C.; Lu, C.J. Mining the customer credit using classiication and regression tree and multivariate adaptive regression splines. *Comput. Stat. Data Anal.* **2006**, *50*, 1113–1130. [CrossRef]
22. Rosenberg, E.; Gleit, A. Quantitative methods in credit management: A survey. *Oper. Res.* **1994**, *42*, 589–613. [CrossRef]
23. Wiginton, J.C. A note on the comparison of logit and discriminant models of consumer credit behaviour. *J. Financ. Quant. Anal.* **1980**, *15*, 757–770. [CrossRef]
24. Chen, N.; Ribeiro, B.; Chen, A. Financial credit risk assessment: A recent review. *Artifcial Intell. Rev.* **2016**, *45*, 1–23. [CrossRef]
25. Huang, Y.; Chen, M. Improve Reputation Evaluation of Crowdsourcing Participants Using Multidimensional Index and Machine Learning Techniques. *IEEE Access* **2019**, *7*, 118055–118067. [CrossRef]
26. Tulchynska, S.; Popelo, O.; Vovk, O.; Dergaliuk, B.; Kreidych, I.; Tkachenko, T. The Resource Supply of Innovation and Investment Strategies of the Microeconomic Systems Modernization in the Conditions of Digitalization. *Wseas Trans. Environ. Dev.* **2021**, *17*, 819–828. [CrossRef]
27. Morton, H. Municipal credit rating modelling by neural networks. *Decis. Support Syst.* **2011**, *51*, 108–118.
28. Wang, G.; Chen, G.; Chu, Y. A new random subspace method incorporating sentiment and textual information for fnancial distress prediction. *Electron. Commer. Res. Appl.* **2018**, *29*, 30–49. [CrossRef]
29. Lee, T.S.; Chiu, C.C.; Lu, C.J.; Chen, I.F. Credit scoring using the hybrid neural discriminant technique. *Expert Syst. Appl.* **2002**, *23*, 245–254. [CrossRef]
30. Sun, J.; Li, H. Listed companies' fifinancial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Syst. Appl.* **2008**, *35*, 818–827. [CrossRef]
31. Finlay, S. Multiple classifier architectures and their application to credit risk assessment. *Eur. J. Oper. Res.* **2011**, *210*, 368–378. [CrossRef]

32. Liang, K.; He, J. Analyzing credit risk among Chinese P2P-lending businesses by integrating text-related soft information. *Electron. Commer. Res. Appl.* **2020**, *40*, 100947. [CrossRef]
33. Hájek, P.; Olej, V.; Myskova, R. Forecasting corporate financial performance using sentiment in annual reports for stakeholders' decision-making. *Technol. Econ. Dev. Econ.* **2014**, *20*, 721–738. [CrossRef]
34. Chen, N.; Ribeiro, B.; Vieira, A.S.; Duarte, J.; Neves, J.C. A genetic algorithm based approach to cost-sensitive bankruptcy prediction. *Expert Syst. Appl.* **2011**, *38*, 12939–12945. [CrossRef]
35. Tsai, C.F.; Hsu, Y.-F.; Yen, D.C. A comparative study of classiffer ensembles for bankruptcy prediction. *Appl. Soft Comput.* **2014**, *24*, 977–984. [CrossRef]
36. He, H.; Garcia, E.A. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 1263–1284.
37. Yin, C.; Jiang, C.; Jain, H.K.; Wang, Z. Evaluating the credit risk of smes using legal judgments. *Decis. Support Syst.* **2020**, *136*, 113364. [CrossRef]
38. Abramova, A.; Shaposhnykov, K.; Zhavoronok, A.; Liutikov, P.; Skvirskyi, I.; Lukashev, O. Ecosystem of VAT Administration in E-Commerce: Case of the Eastern Europe Countries. *Estud. Econ. Apl.* **2021**, *39*, 1–15. [CrossRef]
39. Liang, K.; Jiang, C.; Lin, Z.; Ning, W.; Jia, Z. The nature of sellers' cyber credit in C2C e-commerce: The perspective of social capital. *Electron. Commer. Res.* **2017**, *17*, 133–147. [CrossRef]
40. Galar, M. A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2012**, *42*, 463–484. [CrossRef]
41. Guo, H.; Li, Y.; Shang, J.; Gu, M.; Huang, Y.; Gong, B. Learning from class-imbalanced data: Review of methods and applications. *Expert Syst. Appl.* **2016**, *73*, 220–239.
42. Panhalkar, A.R.; Doye, D.D. A novel approach to build accurate and diverse decision tree forest. *Evol. Intell.* **2022**, *15*, 439–453. [CrossRef]