Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches

Jie Sun *, Hui Li, Qing-Hua Huang, Kai-Yu He

School of Economics and Management, Zhejiang Normal University, Jinhua, Zhejiang 321004, PR China

A B S T R A C T

As a hot topic, financial distress prediction (FDP), or called as corporate failure prediction, bankruptcy prediction, acts as an important role in decision-making of various areas, including: accounting, finance, business, and engineering. Since academic research on FDP has gone on for nearly eighty years, there are abundant literatures on this topic, which may appear chaotic to the researchers of the field and make them feel confused. This paper contributes to the current review researches by making a full summary, analysis and evaluation on the current literatures of FDP. The current literatures of FDP are reviewed from the following four unique aspects: definition of financial distress in the new century, FDP modeling, sampling approaches for FDP, and featuring approaches for FDP. By considering the new state-of-the-art techniques in this area, FDP modeling are classified and reviewed by the following groups: namely, modeling with pure single classifier, modeling with hybrid single classifier, modeling by ensemble techniques, dynamic FDP modeling, and modeling with group decision-making techniques. Sampling methods for FDP are classified and reviewed by the following paired groups, namely: training sampling and testing sampling, single industry sampling and cross-industry sampling, balanced sampling and imbalanced sampling. Featuring methods for FDP are categorized and reviewed by qualitative selection and combination of qualitative and quantitative selection. We comment on the current researches from the view of each category and propose further research topics. The review paper is valuable to guide research and application of the area.

1. Introduction

The word of “Early Warning” is originated from military area. Nowadays, the word is widely used in some respects such as: macroeconomics, business administration, environmental monitoring, among others. Early warning of financial distress, corporate failure, or bankruptcy is an important research field for corporate finance and its core is financial distress prediction (FDP), which is an extensive ongoing research topic. Generally, FDP is to predict whether or not a company will fall into financial distress based on the current financial data, through mathematical, statistical, or intelligent models. It is also called as financial failure discrimination, bankruptcy prediction, business failure prediction, corporate failure prediction, among others. It plays an important role on managerial decision-making for firms, investment decision-making for investors, credit decision-making for creditors, customer credit rating for banks, and so on.

During the last fifty years, numerous researches focus on this topic and attempt to provide a useful solution for the problem. Literatures on this topic may appear chaotic to the researchers of the field and make them feel confused. A thoughtful review on related articles is useful to help people understand the research and development in FDP. Thus, this paper attempts to review the state-of-the-art researches from a new perspective, which is different from the several review articles on FDP in recent years. Balcaen and Ooghe [11] extensively elaborated on the application of classical statistical techniques in FDP from the following categories, namely: single variable analysis, risk index models, multivariate discriminant analysis (MDA), and conditional probability models. This review made a contribution to the current literature by clarifying the characteristics of the classical statistical techniques in FDP and their related problems. However, intelligent techniques are more frequently used in FDP than statistical techniques in recent years. Thus, Kumar and Ravi [58] considered both statistical...
and intelligent techniques in FDP, and classified articles of FDP from the following groups, namely: statistical techniques, neural networks (NN), case-based reasoning (CBR), decision trees (DT), operational research, evolutionary algorithms (EA), rough set (RS) based techniques, other techniques, and soft computing techniques from the hybridization of all the above-mentioned techniques. They also reviewed each article on the source of data, financial variables, country, time line of the data, and performance of techniques. This review presented that the hybrid modeling is a hot topic in FDP. By considering researches in this category, Bahrammirzaee [10] conducted a comparative review of three famous intelligent techniques for financial applications, i.e., artificial neural networks, expert systems and hybrid intelligent systems. Credit evaluation and financial prediction are two chief problems for financial market. They concluded that these intelligent methods were superior to traditional statistical ones in dealing with financial problems, although such superiority is not absolute. Verikas et al. [125] presented a comprehensive review of hybrid and ensemble-based soft computing techniques for FDP. Their focus was on how different techniques were combined. A technique was named as hybrid if several soft computing approaches were applied and only one predictor was used to make the final prediction. A technique was called as ensemble if outputs of several predictors were combined to obtain a prediction. Lin et al. [74] made a review of FDP literatures between 1995 and 2010. They carried out statistical analysis on the FDP literatures in terms of single and soft classifiers, baseline classifiers, and datasets, and considered soft classification techniques as the direction for future research of FDP. Marques et al. [78] summarized the application of evolution computing to credit scoring, which included classification, variable selection, and parameter optimization. They also reviewed performance evaluation criteria, statistical significance tests, credit databases, and some data preprocessing issues.

Though the six reviews have been published, this review is valuable and significant for the following reasons:

1. From the above six review articles we can find that hybrid intelligent techniques are potentially useful models for FDP with high predictive performance. Thus, how to construct hybrid models for FDP needs to be analyzed by grouping the current work.
2. Meanwhile, with the development of computing and modeling techniques, some new techniques may also be useful for solving the problem of FDP, e.g., dynamic prediction of financial distress, class imbalance classification techniques, feature selection approaches, group decision-making techniques, among others. How these new emerged techniques can be used in FDP needs to be analyzed by summarizing existing researches.
3. Third, none of the existing review articles provides a clear indication on FDP with different sampling techniques and feature selection approaches. This topic needs to be investigated.
4. Fourth, some review articles in the last century discussed the definition of financial distress, which was seldom focused on in recent reviews. However, the definition of financial distress should be re-considered since some new concept or phenomena come into the area of FDP in the new century. The current reviews need some supplement to help researchers understand the area of FDP more thoughtfully.

This review is to fetch up the gap of the current reviews. We reviewed the current FDP literatures from the following four unique aspects: definition of financial distress in the new century, FDP modeling, sampling methods for FDP, and featuring approaches for FDP. In detail, FDP modeling articles are classified and reviewed by the following groups: namely, modeling with pure single classifier, modeling with hybrid single classifier, modeling by ensemble techniques, dynamic FDP modeling, and modeling with decision implementations. Sampling methods for FDP are classified and reviewed by the following paired groups, namely: training sampling and testing sampling, cross-industry sampling and single industry sampling, balanced sampling and imbalanced sampling. Feature selection methods for FDP are categorized and reviewed by qualitative selection and quantitative selection.

This paper is organized as follows. Section 2 overviews and comments on the definition of financial distress in the new century. Section 3 reviews and comments on FDP modeling methods from the category of pure single classifier methods, hybrid single classifier methods, ensemble methods, dynamic modeling for FDP, and FDP-based decision implementations. Section 4 reviews and comments on sampling methods for FDP. Section 5 presents the overview and comments from the perspective of featuring methods for FDP. Section 6 concludes and analyzes some further research topics in FDP.

2. Development of definition of financial distress in the new century

2.1. Development of definition of financial distress

Financial distress is the situation that an enterprise has certain kind of financial difficulties. In some classical literatures, such financial difficulties include inability to pay debts or preferred dividend and the corresponding consequences such as overdraft of bank deposits, liquidation for interests of creditors, and even entering the statutory bankruptcy proceeding. Such definition of financial distress is based on the theoretical framework of “cash flow” or “liquid assets” model. As Beaver [13] mentioned, an enterprise is like a reservoir formed by the cash flow, composed of cash inflows and outflows. An enterprise in financial distress is just like a reservoir whose water is drained.

Carminie [18] believed that financial difficulty is a situation that an enterprise encounters frustration in fulfilling its obligations. These frustrations include: insufficiency of liquidity, insufficiency of equity, default of debt, and insufficiency of liquid capital. Foster [35] defined financial distress as a serious liquidity problem which is unable to be resolved without large-scale restructuring of the operation or structure of economic entities. In Doumpos and Zopounidis [32], financial distress not only contains inability to re-pay important obligatory payments and the corresponding consequences mentioned above, but also includes the situation of negative net asset value, which means an enterprise's total liabilities exceed its total assets from the view of accounting.

Ross et al. [100] summarized previous studies and concluded that financial difficulties consist of the following four conditions: (1) business failure, that is, a company cannot pay the outstanding debt after liquidation; (2) legal bankruptcy, namely, a company or its creditors applies to the court for a declaration of bankruptcy; (3) technical bankruptcy, namely, a company cannot fulfill the contract on schedule to repay principal and interest; and (4) accounting bankruptcy, namely: a company's book net assets are negative.

In the 21st century, most FDP literatures that collect data from developed countries or areas concentrate on the prediction of bankruptcy, which is the ultimate and most serious form of financial distress [76, 106, 89, 22]. Bose [14] defined financial distress as the condition that a company's stock price is less than 10 cents, which is followed by Ravisankar et al. [96]. When studying FDP for Taiwan companies, Lin [72] defined financial distress as the inability of a firm to pay its financial obligations as they mature.
Operationally, a firm is said to have been failed when any of the following events have occurred: bankruptcy, bond default, an overdraft of bank account, events signifying an inability to pay debts as they come due, entrance into a bankruptcy proceeding, an explicit agreement with creditors to reduce debts, or being classified as “full delivery stock” by Taiwan Stock Exchange or Gre Tai Securities Market.

While in developing countries like mainland of China and Iran, financial distress is usually defined as certain degree of financial deterioration ruled by the national security management institution. For example, Chinese listed company’s financial distress is defined as special treatment (ST) by Chinese Stock Exchange for the reason that their profits continue to be negative for two consecutive years or their per-share net assets are lower than per-share stock face value [109,31,110]. The Iranian companies whose retained losses are more than 50% of their capital are labeled as financial distress according to commercial law of 141 act of Tehran Stock Exchange [93]. In addition, Sun et al. [108] proposed the concept of relative financial distress, which is defined as relative deterioration of financial situation for a certain enterprise with the process of its life cycle.

2.2. Comments on definition of financial distress

Above all, for definition of financial distress, there exist many different points of view. Different scholars may give different explanations according to their own study purpose. Broadly speaking, there are two main ideas:

From the perspective of theoretical analysis, financial distress has different degrees. Mild financial distress may just be temporary cash flow difficulty, while serious financial distress is business failure or bankruptcy. An enterprise in financial distress may experience a dynamic changing process of various states between the above two extreme forms of financial distress. Actually, financial distress is a dynamic ongoing process, and is the result of continuous abnormality of business operation for a period of time (from months to years or even longer).

From the perspective of empirical research, in order to make clear the criteria of research sampling, or for the restrictions of data availability, financial distress is often defined as some situations which clearly show an enterprise's financial difficulty, such as statutory bankruptcy, and ST for Chinese listed companies. Current studies considered single criteria of financial distress rather than the intensity of it, and future studies need to explore a metric that can classify the distressed companies into different degrees such as mild, intermediate, and bankrupt.

3. Review on FDP modeling methods

FDP is the core process of financial distress early warning. Its research progress can be briefly summarized as follows: from single variable analysis to multivariate prediction; from traditional statistical methods to machine learning methods based on artificial intelligence; from pure single classifier methods to hybrid single classifier methods and classifier ensemble methods; from stationary modeling to dynamic modeling considering time process; from quantitative prediction methods to decision implementations.

3.1. Pure single classifier methods for FDP

3.1.1. Statistical single classifier methods

Statistical single classifier methods, including the single variable analysis, the MDA, and the Logit model, are based on statistical theory. Such statistical methods of FDP are usually simple, easy-to-use and time-saving.

3.1.2. Artificial intelligence single classifier methods

Artificial intelligence single classifier methods for FDP include NN, EA, RS, CBR, SVM, etc. They are applied to FDP on the basis of the fruitful research results of computer and artificial intelligence technology. They are not subject to the stringent assumptions required for statistical methods.
3.1.2.1. Neural networks. In early 1990s, NN began to be introduced into the research on FDP, e.g., Tam [118], Tam and Kiang [119]. A NN model consists of an interconnected group of artificial neurons, and processes information using a connectionist approach for computation. A wealth of related researches studied on NN methods for FDP, and the most widely used is the three-layer feed-forward back-propagation NN (BPNN), in which the hidden layer determines the mapping relationships between input and output layers and the relationships between neurons are stored as weights of the connecting links [84,23,72]. The probabilistic NN was also widely used for FDP and it employs Bayesian decision-making theory based on an estimate of the probability density in data space [132,129]. The FDP performances of NN models were often compared with those of the MDA and Logit models, and most researches provided evidence for the conclusion that NN performed better than the statistical methods [34,17,60,135,89,71,129,93]. The advantages of NN over the statistical methods are often attributed to its strong mapping ability based on the network structure. In addition, the statistical relations among variables need not to be considered in the process of NN model construction. Such an advantage of NN is supported by the conclusion of Lin [72] that indicates NN can achieve higher prediction accuracy if the data does not satisfy the assumptions of the statistical approaches. In addition, Chen and Du [23] found that BPNN also performs better than data mining clustering technology for the purpose of FDP.

However, compared to statistical methods, far more sample data are needed to train a relatively stable NN model, and too much duplication of training easily leads to over-fitting, which decreases the stability of cross-sample prediction. In addition, NN is often criticized by practitioners for its difficulty of understanding because the complex network structure seems to be a black-box for the decision makers. To overcome this drawback, Baesens et al. [9] captured the learned knowledge embedded in the networks and expressed it as explanatory rules and a decision table that facilitates easy consultation. Setiono et al. [104] proposed a novel approach to train minimal NNs that are easy to generate concise and comprehensible classification rules for the user. Such a research direction for NN-based FDP is very important because it makes NN-based models easier to be applied for the real-world FDP problem.

3.1.2.2. Support vector machines. Support vector machine (SVM) is a relatively new artificial intelligence method that is based on the structural risk minimization principle rather than the empirical risk minimization principle. SVM is a powerful and promising data classification and function estimation tool [130], which is not easy to run into over-fitting even for relatively small sample. Shin et al. [107] and Min and Lee [80] both used SVM to predict bankruptcy for South Korean companies, and drew the conclusion that this method outperformed MDA, Logit and NN. Hui and Sun [49] and Ding et al. [31] respectively adopted SVM model to do empirical study on FDP for Chinese listed companies, and reached a similar conclusion. Li and Sun [68] used a straightforward wrapper approach to help SVM model produce more accurate prediction of business failure, which effectively improve efficiency of SVM in FDP. Gesel et al. [39] demonstrated that least square SVM was the most preferred predictive model for FDP compared with the Logit, the Fisher MDA and the quadratic MDA. Their results also indicated that bankruptcy problems were weakly non-linear. However, Bose and Pal [15] compared the performance of SVM in FDP with those of MDA and NN, and concluded that SVM performed worse than NN. More studies used SVM with some other techniques in a hybrid or ensemble means, instead of independent use of SVM. These literatures are reviewed in Sections 3.2 and 3.3.

3.1.2.3. Evolution algorithms. EAs are generic population-based meta-heuristic optimization algorithms that use mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, and selection. Varetto [124] applied Genetic Algorithm (GA) to extract linear functions without statistical restrictions and their corresponding discriminant rules. However, its FDP performance was not as good as MDA. Shin and Lee [106] adopted GA to search split point (threshold) of financial ratios, and then extracted the quantitative discriminant rules of bankruptcy prediction. Kim and Han [55] used GA to mine experts’ qualitative bankruptcy prediction rules that had the advantages of easy understanding. However, such rules-based methods could generate FDP results for companies only when at least one rule is activated, and the coverage rate of rules were relatively low. In addition, Rafiei et al. [93] also found that GA had lower FDP accuracy than NN. Etemadi et al. [33] investigated the application of Genetic Programming (GP) for bankruptcy prediction for Iranian listed firms, and McNemar test showed that GP outperformed MDA. Some other EAs, e.g., Particle Swarm Optimization, Ant Colony Optimization, are potential alternatives for FDP. For example, Martens et al. [75] developed AntMiner + based on Ant Colony Optimization to produce reasoning rules for FDP, but this tool was inferior to SVM in term of predictive accuracy. Hence, the FDP methods based on EAs can generate rules that are more interpretable than NN or SVM models, but the performance based on EAs alone is limited and they are more suitable to be combined with other classification algorithms for FDP.

3.1.2.4. Case-based reasoning. FDP based on CBR generally uses K-nearest neighbor algorithm, which determines or predicts the target cases’ labels according to similar cases’ labels. The evident advantage of CBR is that it is easy to understand and its prediction accuracy is also relatively high with the development of CBR technique. By comparing CBR, NN and MDA, Jo and Han [52], Jo and Han [53] found that there was no real difference between CBR and MDA, and CBR performed well when data was not sufficient. Park and Han [87] applied AHP weighted K-nearest neighbor algorithm to bankruptcy prediction, and significantly improved the prediction accuracy by integration of qualitative and quantitative indicators. Sun and Hui [109] proposed a similarity weighted voting CBR method based on gray correlation distance, which performed effective short-term FDP one or two years in advance. In recent years, Li and Sun [62], Li and Sun [63], Li and Sun [64], Li et al. [61] respectively injected principles of ELECTRE, PROMETHEE, ORESTE, and TOPSIS methods into CBR to generate various CBR methods with reasonable FDP performances. Borrajo et al. [16] constructed a multi-agent system for business control process and failure prediction, and the core agent incorporates a CBR system.

3.1.2.5. Rough set. Dimitras et al. [30] and McKee [76] respectively applied RS theory to establish bankruptcy prediction model on Greece data and US data. They found that this method had many advantages, e.g. easily understandable decision rules in natural language, case support for decision rules, combination of qualitative and quantitative variables, no statistical evaluation of the probability and fuzzy membership, etc. However, different samples and decision-makers’ knowledge would generate different set of decision rules. Bose [14] also applied RS to decide financial health of dot-coms with an acceptable validation accuracy of 72.08%, but RS often resulted in many rules associated with each class and most of them (90%) were redundant. Therefore, RS has the shortcomings of unfixed structure and poor universality.
3.1.2.6. Decision tree including RSP, CART and See 5.0. Frydman et al. [36] and McKee and Greenstein [77] adopted recursive partitioning (RSP) method which created DTs for FDP. Sun and Li [111,112] proposed a data mining method based on attribute induction, information gain, and DT, and applied it for FDP of Chinese listed companies. Gepp et al. [38], Li et al. [65] also proved that DT or classification and regression tree (CART) or See 5.0 had better FDP performance than MDA. Chen [24] empirically compared DT with Logit for Taiwan firms, and found that DT got higher accuracy than Logit in short run (less than one year), while Logit performed better in long run (above one and half year). Olson et al. [86] found C5 and CART DT obtained better FDP performances than NN and SVM for the particular US bankruptcy data set, but DT models were easy to generate too many rules. They controlled this by adjustment of minimum support parameter, which helped make a trade-off between average accuracy and DT size.

3.1.2.7. Other techniques. Besides the above major FDP methods, there are some other ones. For example, Sarkar and Sriram [103], Gesel et al. [40], Sun and Shenoy [116], and Wu [127] applied Bayesian (kernel) classifier for early warning of business failures. Pendharkar [88], Ciejen et al. [27], and Premachandra et al. [92] used data envelopment analysis as a tool for corporate failure prediction. Kwak et al. [59] proposed a multiple criteria linear programming (MCLP) method for FDP with Korean data, and indicated that the MCLP performed as well as MDA and Logit and was comparable to DT and SVM. Ryu and Yue [102] developed a new mathematical programming model named Isotonic Separation (IS) for FDP, which outperformed MDA, NN, DT, and RS in the bankruptcy prediction experiment with North American data. IS is different from the well-known statistical survival analysis methods in that it is able to predict a sample's survival time frame [101]. Chandrasekaran et al. [20] further revised IS to make it applicable on regression problems.

3.1.3. Comments on pure single classifier methods for FDP

Pure single classifier methods are the base for FDP modeling and each one has its own superiority and inferiority. Statistical methods are restricted by statistical assumptions, but they can obtain a fixed model structure in different times of training on a certain data set. On the contrary, artificial intelligence methods are not subjected to statistical assumptions and can construct FDP models for more complex data context, but their training processes are much more complex too. Table 1 shows the comparison of training characteristics among six typical artificial intelligence classification algorithms. Firstly, CBR does not need training process before implementation of FDP, which greatly differs from the other five algorithms that need training models/rules before FDP. Secondly, DT has an advantage that it needs not setting any parameter before training, but the other five artificial intelligence methods need parameter setting before training, which significantly affects the performance of FDP models trained. Third, for NN and EA, different times of training on a certain data set may result in different FDP models for certain degree of randomness in the training process.

3.2. Hybrid single classifier methods for FDP

Many researchers studied on hybrid FDP methods based on two or three algorithms, and the most popular forms are integrating NN, CBR or SVM with other techniques. They are categorized into three types as shown in Fig. 1.

3.2.1. One algorithm is applied to choose features for another classification algorithm

It is a very common hybrid paradigm to use the first technique for feature selection and the second for classification, which is a two-stage modeling process. As shown by Fig. 1(1), if the first stage applies certain evolution algorithm to optimize the feature portfolio and uses a classification performance measure as the object function, it will embed the classification modeling into the feature selection process. Other than this situation, feature selection and classification modeling are usually two independent sequential stages.

3.2.1.1. Hybrid NN for FDP. Back et al. [8] and Anandarajan et al. [4] combined NN with GA. They used GA to select input variables for NN and established GA-NN model to further improve FDP performance. Ravisankar et al. [96] and Ravisankar and Ravi [97] also constructed several hybrids in such a paradigm. In the former, each hybrid comprised two of the multilayer feed forward NN, probabilistic NN, RS and genetic programming (GP), and GP-GP hybrid showed the best performance at the significance level of 10%. In the latter, features were firstly selected by t-statistic, f-statistic, or Group Method of Data Handling (GMDH, a relatively unexplored NN), and classifiers were trained by GMDH, counter propagation NN and fuzzy adaptive resonance theory map. As a result, GMDH-GMDH and t-statistic-GMDH outperformed others.

3.2.1.2. Hybrid SVM for FDP. Yeh et al. [133] built a two-stage hybrid model that integrates RS theory and SVM, called as RST-SVM. Redundant attributes were reduced by RS and then business failure was predicted by SVM, and it was proved to outperform RST-BPN. Lin et al. [73] trained a SVM FDP model after feature dimensionality reduction though isometric feature mapping (ISO-MAP) algorithm, one of the most developed dimensionality reduction techniques. It produced better performance when compared with the PCA-SVM hybrid.

3.2.1.3. Hybrid CBR for FDP. Ahn and Kim [11] put forward a CA-CBR hybrid bankruptcy prediction method, in which GA was used to optimize feature weighting and instance selection simultaneously. Such a hybrid obtained higher prediction accuracy than conventional CBR, but it needed much more modeling time and computer resource. Li et al. [66] solely used GA to select feature for CBR and applied the hybrid algorithm in FDP. Cho et al. [26] used DT to

<table>
<thead>
<tr>
<th>Training models/rules before FDP</th>
<th>Needing parameter setting before training</th>
<th>Fixed model structure for different times of training based on the same data and parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>SVM Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>EA Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>CBR No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>RS Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DT Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1 The comparison of training characteristics among six typical artificial intelligence classification algorithms.
3.2.2. One algorithm is applied to optimize parameters for the other classification algorithm

This hybrid paradigm is usually designed for the classification algorithms that need parameter setting before model training, especially for the classification algorithms such as NN and SVM whose parameter values are critical to the final FDP models. As shown by Fig. 1(2), although parameter optimization and classification modeling are two sequential stages, classification algorithms should also be used in the parameter optimization stage, in which the evolution methods use classification performance measures as the object function for optimizing parameters.

3.2.2.1. Hybrid SVM for FDP. Wu et al. [128] integrated real-valued GA (RGA) with SVM, so that the optimal parameters of SVM can be automatically optimized by RGA considering the predictive accuracy and generalization ability simultaneously. RGA-SVM showed better experimental results than the pure SVM as well as DA, Logit, Probit, and BPN. Min et al. [81] also proposed a GA-SVM hybrid model for bankruptcy prediction, but they used GA to optimize both feature selection and SVM parameters, to significantly improve the predictive ability of SVM.

3.2.2.2. Hybrid NN for FDP. Chaudhan et al. [21] used differential evolution (DE) for optimizing weights of wavelet neural network (WNN), with FDP results indicating that DEWNN outperformed initial WNN and threshold accepting WNN (TAWNN). Hu [44] developed an ELECTRE-based single-layer perceptron (SLP) approach and GA was then designed to determine its connection weights. Its application to FDP demonstrated that it performed better than the traditional SLP, the multi-layer perceptron, and some other single classifier methods. Pendharkar [80] focused on FDP by integrating cost-sensitive learning with GA-NN through using posterior probability for minimizing total misclassification cost as the fitness function of GA, and it effectively helped the model improve the FDP performance with imbalanced samples.

3.2.3. A new classification algorithm is produced by integrating principles of two or more methods

Different from the former two hybrid types that belong to two-stage integration, this type of hybrid paradigm is an inner fusion of two algorithms to produce a new classification algorithm (see Fig. 1(3)), which is expected to be more effective for FDP modeling than either original algorithm.

Cheng et al. [25] proposed a FDP method by embedding Logit into radial basis function NN, in which the logistic function was used as the active function between the hidden nodes and the output node, and it outperformed both the Logit model and the BPNN. Hua et al. [48] developed a method called as binary discriminant rule for corporate FDP, which modified the outputs of the SVM classifiers according to the result of logistic regression analysis, and its prediction effect was better than the pure SVM. Chaudhuri and De [22] integrated SVM with fuzzy membership functions to propose a new classification tool called as FSVM, and it was effective in their bankruptcy prediction experiment for large organizations.

3.2.4. Comments on hybrid single classifier methods for FDP

Studies on hybrid single classifier methods for FDP are almost present in the 21st century, and began to boom especially in the recent five years. This indicates that hybrid method for FDP is a new subfield of FDP research that appeals to the current researchers. Among the three hybridization forms, the first two paradigms that are two-stage integration of two algorithms are more common than the third paradigm, which needs inner embedding of two algorithms with much more innovation. Therefore, it is worth exploring how to produce some new hybrid classification algorithms that are effective for FDP, and more than one form of hybridization can be further integrated to obtain more effective FDP tools.

3.3. Ensemble methods for FDP

In recent years, ensemble methods became another fruitful topic in FDP field. An ensemble system can exploit each base classifier's unique information for classification, and it is expected to produce smaller error variance than each base classifier in standalone mode. Though the pioneering work on ensemble was done early by Bates and Granger [12], its application to FDP firstly emerged in Jo and Han [52] as far as we know and began to rise.
only recently after 2005. There are mainly two forms of classifier ensemble: parallel and serial.

3.3.1. Parallel ensemble methods for FDP

In a parallel ensemble system, the classification results of multiple classifiers are combined according to the schemes of majority voting, weighted majority voting, product rule, minimum or maximum, Borda count, simple average, etc. [43]. There exist mainly three approaches to generate base-level classifiers [120]: (1) generating classifiers by applying different learning algorithms (with heterogeneous model representations) to a single data set, (2) applying a single learning algorithm to different versions of a given data set with different samples and features, and (3) applying a single learning algorithm with different parameter settings to a single data set.

3.3.1.1. Ensemble with different algorithms. This ensemble approach was early used in FDP. Its overall framework is shown in Fig. 2. It uses several classification algorithms to generate several FDP models with different representations, which are believed to contain complementary classification information. By combining the outputs of these FDP models, the ensemble system obtains the final FDP result that is more accurate and stable than member models.

For example, Jo and Han [52] utilized MDA, CBR and NN to generate base classifiers and integrated their predicted values by weighted sum scheme to reduce the bankruptcy prediction error. Ravikumar and Ravi [98] developed a set of ensemble classifiers by simple majority voting scheme based on original seven algorithms. Each time, two, three, four, five or six classifiers with highest classification rates and least Type-I error were selected to construct the best ensemble classifier, which outperformed constituent stand-alone models. Sun and Li [111,112] investigated weighted majority voting combination of multiple diversified classifiers that were generated by MDA, Logit, NN, DT, SVM and CBR, and obtained higher average accuracy and lower variation coefficient than any base classifier. Ravi et al. [94] combined a multi-layer feed forward BPNN, a probabilistic NN, a radial basis function NN, SVM, DT, and a fuzzy rule classifier to be an ensemble system for FDP, and demonstrated that the ensemble yield lower Type I and II errors.

3.3.1.2. Ensemble with one algorithm under different samples or features. The framework of ensemble FDP system with one algorithm under different samples or features is shown in Fig. 3. This ensemble approach tries to produce multiple diverse training datasets from the original one by a sampling or featuring mechanism, so that a classification algorithm can train on them to construct multiple FDP models that have certain degree of diversity among each other. Certain combination mechanism can also be used to combine the outputs of different FDP models for the final FDP result. This kind of ensemble method tries to make full use of different fragments of information contained in the original data set and thus improves the FDP performance for a given classification algorithm.

Bagging and Boosting are the two most popular ensemble methods that apply a single learning algorithm to different versions of a given data set. Bagging and Boosting (including AdaBoost) ensembles based on NN were widely applied in FDP researches such West et al. [126], Alfaro et al. [2], and Kim and Kang [56]. They had the generalization ability superior to the single NN model. Sun et al. [110] constructed AdaBoost ensemble respectively with single attribute test (SAT) and DT for FDP, and found that AdaBoost-SAT outperformed AdaBoost-DT, single DT and single SVM, indicating that AdaBoost is more suitable for weak learners. Kim and Upneja [57] compared the predictive and discriminatory performances of AdaBoosted DT models with single DT models for publicly traded US restaurants for the period from 1988 to 2010, and AdaBoosted DT model demonstrated the best prediction performance with the smallest error in overall and type I error rates. Besides, Random Subspace is another ensemble method that uses different versions of a given data set to produce base classifiers. Namely, each individual classifier uses only a subset of
all features for training and testing. Nanni and Lumini [82] found that Random Subspace outperformed Bagging, Class Switching, and Rotation Forest. Li et al. [67] integrated Random Subspace method with Logit for FDP, and improved the FDP performance of Logit models significantly. Li and Sun [69] proposed a principal component CBR ensemble method that produced diverse base models by different features and different algorithms simultaneously to effectively improve the performance of CBR in FDP. The different feature sets were the initial ones and those respectively selected by stepwise MDA, stepwise Logit and t-test. Two CBR algorithms respectively with Euclidean Manhattan metrics were used to construct CBR classifiers.

3.3.1.3. Ensemble with one algorithm under different parameters. This ensemble method is applicable to classification algorithms that can train different FDP models with different parameter values, e.g. NN and SVM. Because improper parameter settings may lead to terribly bad model performance and the number of FDP models trained with optimized parameter values is usually limited, an ensemble system constructed under the thought of one algorithm under different parameters alone usually does not have enough diversity. As a result, it is usually combined with ensemble under different samples or features and the framework is shown in Fig. 4.

Tsai and Wu [122] constructed an NN ensemble by mainly using different parameters and also using different versions of the same training set. However, it did not outperform the single best NN classifier in many cases, and it was attributed to too small training dataset for diversified single classifiers and the binary classification problem of FDP. Sun and Li [137] constructed a SVM ensemble whose candidate single classifiers were trained by SVM algorithms with different kernel functions and parameters on different feature subsets of one initial dataset, and indicated that it significantly outperformed individual SVM classifier when the number of base classifiers in SVM ensemble was properly set.

3.3.2. Serial ensemble methods for FDP

Different from parallel combination, a serial ensemble system arranges several base classifiers in sequence and selects the result of one base classifier as the final output according to certain principles. However, the methods to produce candidate FDP models in parallel ensemble systems are still suitable for serial ensemble systems. Fig. 5 shows the framework of serial ensemble FDP system, in which a base model except the last one has an effective output only when it is triggered by the base model in front of it and its output class has enough credibility simultaneously. The last base model \( n \) in the serial FDP system can be either a single classifier or a parallel ensemble classifier, which is usually considered to have overall good FDP performance. For example, Hung and Chen [50] proposed a selective ensemble for bankruptcy prediction, which selects some suitable classifiers on the base of the expected probability of each individual classifier and uses a parallel voting ensemble as the last base model. Sun and Li [113] applied the serial ensemble method for FDP by designing two selection operators for each class’s best classifier and the wholly best single classifier. However, serial ensemble methods did not show much superiority to the best base classifier for the FDP problem with two categories.

3.3.3. Comments on ensemble methods for FDP

Classifier ensemble, which is also named as combination of multiple classifiers, has shown much superiority for FDP. However, the current researches almost focused on the effectiveness and feasibility of different forms of ensemble classifiers for FDP by comprehensive comparisons, and some problems still need to be further explored. For example, how to selectively choose partial base classifiers from candidate ones to construct more effective classifier ensembles for FDP? What kind of measures is more suitable for evaluating candidate classifiers considering both predictive ability and diversity? Whether other combination mechanisms instead of voting work better for FDP?

3.4. Dynamic FDP modeling for disposal of financial distress concept drift

It is a real-world problem how to update the FDP model dynamically when the new sample data batches gradually emerge with time going on. The sample data that flow in at intervals may be col-
lected by two ways: one from different companies and the other from different time points of certain company. Dynamic FDP modeling in the former case is considered as lateral and that in the latter case is considered as longitudinal. Either lateral or longitudinal dynamic FDP modeling tries to construct a model updating mechanism that can make the FDP model keep effective with time going on. The core thought of dynamic FDP modeling is shown in Fig. 6, and the key problem is how to design the FDP updating mechanism.

3.4.1. Lateral dynamic modeling for FDP
Sun and Li [115] firstly explored the financial distress concept drift (FDCD) that may behave in two ways: real and virtual, and attempted to treat the virtual one by dynamic FDP modeling using five instance selection methods, namely full memory windowing, no memory windowing, windowing with fixed size, windowing with adaptable size, and batch selection. They indicated that virtual FDCD existed and the dynamic FDP models (particularly windowing with fixed size and batch selection) outperformed static ones. Sun et al. [138] further proposed a new FDP method for FDCD that is called as adaptive and dynamic ensemble (ADE) of SVM (ADE-SVM). It was divided into three steps: incremental construction of candidate SVMs, adaptive selection of base SVMs for the current financial distress concept, and dynamic combination of multiple SVMs. It was proved that ADE-SVM had better FDP performance than the above instance selection methods.

3.4.2. Longitudinal dynamic modeling for FDP
Sun et al. [108] proposed a dynamic FDP method named as SFFS-PC-NN optimized by GA with an individual enterprise’s longitudinal data streams, so that a company’s ex-post financial situation evaluation was combined with its ex-ante FDP. In the process of longitudinal dynamic FDP modeling, the label for a company’s financial condition at certain time point may varies with the drift of relative FDCD, which makes it more effective and suitable for a specific company’s financial risk management.

3.4.3. Comments on dynamic FDP modeling methods
Both lateral and longitudinal dynamic FDP modeling methods use incremental sample data to update FDP models with time going on, trying to solve the FDCD problem. Since the external and internal environment for an enterprise always changes constantly, such study on dynamic FDP modeling is particularly important for keeping the adaptability of FDP models. However, such a research topic in FDP just began to obtain focus of the area, and how to dispose the problem of FDCD by proposing more creative and effective dynamic FDP modeling approaches that consider the factor of time leaves to be further studied.

3.5. FDP based on decision implementations and comments
3.5.1. FDP based on decision implementations
Existing quantitative FDP methods almost only took financial ratios into consideration and paid little attention to the important role of experts’ experiential knowledge and non-financial information. Sun and Li [114] brought forward a group decision making method for financial distress early warning by designing qualitative attribute system and the mechanism for integrating multiple experts’ opinion for financial distress diagnosis.

Some researches applied social network analysis to study the links and effects of one company’s failure on others related. When the customers of a sound firm are financially distressed, then this firm gets into financial difficulties with a probability ranging from 3.4% to 11.3%, depending on the level of collection costs and the degree of diversification [6]. Such bankruptcy propagation among suppliers/customers in supply chains connected by trade credit can be upward, downward, or both upward and downward, which is determined by whether payments are delayed and whether firms can instantly adapt their resources to the input received [7]. In a credit network model that includes firms and banks, the bankruptcy of a borrower may lead to the lender’s bankruptcy or an increase of interest rate provided by the surviving lender, which further causes more borrowers already on the verge of bankruptcy to reach a tipping point and default [41]. Thus an avalanche of bankruptcies may ensue. Xu et al. [131] verified that supply chain coordination like information sharing and vendor-managed inventory is effective in reducing the risk of bankruptcy. Furthermore, Hua et al. [47] studied on bankruptcy propagation in a supply chain, and found that operational interactions between firms and operational decisions made by firms in a supply chain were important causes of such bankruptcy propagation. Meanwhile, impacts of these operational issues and decisions depended on financial decisions.

3.5.2. Comments
The problem of FDP can be regarded as a group decision-making problem, and the process of FDP should integrate knowledge of various experts related to the firm, e.g., managers, stockholders, governmental officials. Therefore, group decision-making methods for FDP should gain more focus and become a necessary supplement of quantitative FDP techniques. Meanwhile, investigation on FDP from the view of credit chain and supply chain management is valuable, because one firm’s failure is possible to result in the failure of an entire chain. In other words, identification of key actors whose distress can affect the others will further help analysis of predictors of FDP, and if some factors related to credit chain and supply chain operations are considered as the input variables for FDP, the effectiveness of FDP may be further improved.

![Fig. 6. The core thought of dynamic FDP modeling.](image)
4. Sampling and featuring methods for FDP

4.1. Sampling methods for FDP

4.1.1. Training and testing sampling

Experimental samples of each class are often divided into training and testing samples randomly. Training data are used to establish FDP model based on statistical or artificial intelligence techniques. Testing data are used to test the model’s generalization ability and robustness. Many literatures [3,7,77,42,103,54,51], etc.) on FDP used accuracies on testing data to measure the efficiency of FDP models, because hold-out testing ensures the independence of training data and testing data. However, the randomness of sample division in hold-out testing may result in stochastic FDP performance, making the conclusion less persuasive. Therefore, some studies used thirty or more times of hold-out testing to obtain thirty or more testing accuracies for a given FDP approach, so that statistical mean comparison and deviation analysis can be implemented for more persuasive conclusions [112,110,70].

If the whole sample is randomly divided into K folds and each time one fold is used for testing and the remaining (K – 1) folds are used for training, the K-fold cross-validation accuracy can be computed after K times of iterations. K-fold cross-validation procedure is preferred to hold-out procedure especially for small sample size because it is a more objective performance measure [36]. When K equals to the total sample number, it is unbiased leave-one-out validation, which needs much more computational cost and is only suitable for less complicated approaches with a relatively small data set [105]. The K-fold cross-validation can also be repeated for more than thirty times for statistical analysis.

4.1.2. Cross-industry sampling and single industry sampling

To ensure sample size, most FDP researches used cross-industry samples, in which each distressed company was matched with a healthy one from the same industry but different pairs may belong to different industries. Some studies also considered the impact of industry on FDP modeling. Platt and Platt [91] showed that industry adjustment for financial ratios improved the effectiveness of FDP modeling, and the model trained from one industry was often ineffective for FDP of other industries. Grice and Ingram [42] studied on the sensitivity of Altman’s model, and concluded that the model trained with only manufacturing samples had significantly higher overall accuracy than the model trained with both manufacturing and non-manufacturing samples.

With the development of research database, some researches began to focus on FDP modeling for a specific industry such as manufacturing [135], retail [45,46], hotel [70], restaurant [57], and dotcom [19,96]. Since different industries face different levels of competition, the likelihood of financial distress can differ for firms in different industries [57]. FDP modeling based on single industry sampling can help construct more efficient FDP models for a specific industry.

4.1.3. Balanced sampling and imbalanced sampling

Most FDP researches used balanced samples [3,60,52,53,87,107,112,115]. However, Zmijewski [136] showed that if the ratio of distressed to non-distressed samples evidently deviated from the real-world population, it would distort the model’s prediction capability. Specifically, the proportion of distressed samples is negatively related with the type I error rate (to recognize the distressed as non-distressed), while it is positively related with the type II error rate (to recognize the non-distressed as distressed). That is, if the proportion of distressed samples is higher than that in the real-world population, it will lead to the underestimation of the type I error and the overestimation of the type II error, and vice versa.

Later on, some FDP studies applied imbalanced samples, also called proportional samples, in which the proportion of distressed sample should be closer to that in the real-world population. Namely, the FDP experiments should use fewer distressed and more non-distressed companies. Table 2 shows the sample information of some representative FDP literatures based on imbalanced datasets [77,42,103,54,51]. For example, McKee and Greenstein [77] used five groups of imbalanced training and testing datasets, and the proportion of distress to non-distress in testing samples reflects the real-world proportion. While the proportions in some literatures are much lower than 1:10 [51,31,45,46]. Besides Gepp et al. [38] used an imbalanced dataset consisting of 58 failed and 142 non-failed samples and carried out fivefold cross-validation for the DT FDP method.

The above literatures did not take treatment measures for imbalanced training data sets, but tuned the models’ cut-offs to ensure the prediction performance for skewed testing data sets. Such an approach has the limitation that the cut-off should be determined before prediction, which is a difficult task in real world, and the prediction performance will be low if the cut-off is not set properly. Li and Sum [70] used an over-sampling method to generate new minority samples to make the training dataset balanced, and found that the FDP model constructed on the corrected balanced training data set significantly outperformed the FDP model trained on the original imbalanced data set. In fact, there are several sampling methods that can be used to treat the problem of imbalanced datasets. Crone and Finlay [28] applied both over-sampling and under-sampling methods to balance the

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Training sample</th>
<th>Testing sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distressed</td>
<td>Non-distressed</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McKee and Greenstein</td>
<td>54</td>
<td>288</td>
<td>12</td>
</tr>
<tr>
<td>[77]</td>
<td>66</td>
<td>336</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>240</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>72</td>
<td>292</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>76</td>
<td>309</td>
<td>18</td>
</tr>
<tr>
<td>Grice and Ingram [42]</td>
<td>148</td>
<td>824</td>
<td>148</td>
</tr>
<tr>
<td>Sarkar and Sriram [103]</td>
<td>118</td>
<td>793</td>
<td>25</td>
</tr>
<tr>
<td>Hwang et al. [51]</td>
<td>154 (78 solvency difficulty and 116 failure samples)</td>
<td>2838</td>
<td>229 (119 solvency difficulty and 110 failure samples)</td>
</tr>
<tr>
<td>Ding et al. [31]</td>
<td>50% of 79</td>
<td>50% of 156</td>
<td>50% of 79</td>
</tr>
</tbody>
</table>
original imbalanced credit scoring datasets. They provided evidence that over-sampling significantly increases the accuracy relative to under-sampling across all algorithms, and Logit is less sensitive to the balancing preprocessing than MDA and CART DT.

4.2. Comments on sampling methods for FDP

Training and testing sampling relates to how to train and assess FDP models with available samples. It provides a way to simulate real-world FDP process, which produces statistical results for performance estimation of FDP models.

A cross-industry data set is usually applied for PDF modeling when the size of single industry data set is too small, especially when the number of distressed samples is too small. The knowledge of FDP differs among different industries, which have different financial characteristics. Therefore, a single-industry dataset is preferred when the volume of interested samples is not too small.

The real-world problem of FDP related to imbalanced datasets. Solutions for imbalanced data sets at data level include over-sampling, under-sampling and a hybrid of them. Traditional FDP researches used pairing strategy to balance datasets, which essentially belongs to over-sampling. Besides, imbalanced problems can also be solved at model level, which adjusts the predictive modeling by introducing a threshold parameter or cost-sensitive learning. For FDP with imbalanced datasets, assessing models only by predictive accuracies may provide misleading information. For example, a model can produce an accuracy of 99% when predicting all the samples as non-distressed, if the dataset consists of 99% non-distressed samples and 1% distressed samples. Hence, FDP researches based on imbalanced datasets should consider the Type I and II error rates and some other measures such as sensitivity, specificity, and F-measure.

5. Featuring methods for FDP

Financial experts have found that there are some abnormal changes in financial or non-financial indicators before financial distress occurs. Therefore, financial distress can be predicted by observing the characteristics of some feature indicators. Theorists and practitioners are struggling to establish an effective feature set for FDP.

5.1. Feature selection in classic statistical-technique-based literatures of FDP

Beaver [13] selected the initial financial indicators by four criteria:

1. Popularity in the literature.
2. Potential relevancy to the study by his own professional judgment.

To establish the final variable profile, the following four procedures are utilized:

1. Observation of the statistical significance of various alternative functions including determination of the relative contributions of each independent variable.
2. Evaluation of inter-correlations between the relevant variables.
3. Observation of the predictive accuracy of the various profiles.

Five variables are finally selected, including working capital/total assets, retained earnings/total assets, earnings before interest and taxes/total assets, market value equity/book value of total debt, and sales/total assets.

According to simplicity principle, Ohlson [85] selected nine financial ratios by including the five financial ratios frequently present in previous literatures, as well as adding asset size, two virtual variables representing whether net income is negative or positive and whether total liabilities exceeds total assets, and a variable measuring net income’s trend. He found that company size was an important predictor for financial distress, followed by the financial structure.

Rose and Giroux [99] compared the means of one hundred and thirty indicators that did not appear previously between forty-six pairs of failed and non-failed companies. They found that thirty-four indicators had significant difference. After combining these indicators with the other twenty-seven indicators commonly used in previous studies, they carried out stepwise MDA to establish a final discriminant equation that included eighteen variables, to improve the prediction accuracy. Hence, innovative use of new indicators for new models may improve the prediction ability.

5.2. Qualitative selection of FDP features

On one hand, qualitative selection means that the features for FDP modeling are selected subjectively by researchers according to their analysis on different indicators. For instance, Kiviluoto [139] subjectively selected the following four financial indicators: operating margin, net income before depreciation and extraordinary items, net income before the depreciation and extraordinary items of the previous year, and equity ratio. Kim and Han [55] extracted expert qualitative decision rules by six qualitative attributes including industry risk, management risk, financial flexibility, credibility, competitiveness, operating risk. Sun and Li [114] put forward seven first-level qualitative attributes for early warning of financial distress based on group decision making.

On the other hand, qualitative selection also means directly using the indicators that were used in previous literatures [132,76,77], or subjectively and selectively adding or deleting some indicators based on them. For example, Cielien et al. [27] directly selected eleven financial indicators from Foster [35] and other previous researches. Pendharkar [89] directly chose five financial indicators from Altman’s Z-Score model and Zeta score mode. Zhang et al. [135] combined current ratio with the five variables in Altman’s Z-score model. More recently, Wu et al. [128] se-
lected nineteen financial ratios that were organized into four groups and Chen [24] selected thirty-seven variables covering six major types both based on prior FDP researches. While, Bose [14] firstly identified fifteen popular financial ratios used in literatures, and then added the other nine to capture the novelty of dot-com companies.

5.3. Quantitative selection of FDP features

We considered a feature selection method quantitative as long as quantitative approaches were employed. Thus, this section also includes the combination use of qualitative and quantitative feature selection.

5.3.1. Quantitative feature selection

Based on the initial features subjectively or qualitatively selected by researchers, quantitative statistical methods, artificial intelligence methods or other methods are utilized to streamline features. For instance, Jo and Han [52], Jo and Han [53], Park and Han [87], Shin and Lee [106], and Ding et al. [31] selected variables by the statistical methods of stepwise regression or t-test. Back et al. (1996) [8] selected variables for MDA and Logit models by stepwise regression, while selected variables for NN by integrating GA. Galvao et al. [37] optimized the feature set for FDP with the criteria of greatest information content and minimum multi-collinearity by GA.

Atiya [5] developed two feature systems: one based on financial ratios alone and another based on financial ratios and stock-price-based indicators. They obtained the best five or six inputs from a pool of about 120 candidate inputs by using an initial prescreening procedure based on individual indicator prediction accuracy and correlation matrix and a subsequent cross-validation procedure. The five indicators included: book value/total assets, cash flow/total assets, rate of change of cash flow per share, gross operating income/total assets, and return on assets. The six indicators included: book value/total assets, cash flow/total assets, price/cash flow ratio, rate of change of stock price, rate of change of cash flow per share, and stock price volatility. Li and Sun [68] integrated a straightforward wrapper with SVM for FDP. A forward feature selection method was used to implement the wrapper. Performance of SVM on each feature was used to ranking indicators, which is similar to the operation of Atiya [5]. Meanwhile, they also investigated the feasibility of using linear SVMs as assessment function in feature selection. The result indicated that the RBF non-linear SVM with features selected by linear SVM performed significantly better than all the other SVMs. Sung et al. [117] employed a data mining approach to solve the problem of FDP, and revealed that the major features in predicting bankruptcy included: cash flow to total assets, and productivity of capital under normal conditions. Meanwhile, the indicators of cash flow to liabilities, productivity of capital, and fixed assets to stockholders equity and long-term liabilities were effective to predict financial distress under crisis conditions. Their findings also indicated that the effective FDP models in crisis economic conditions and normal economic conditions were different, which was also supported by the finding of Li et al. [61].

In addition, Leshno and Spector [60] streamlined variables by multi-step exclusion procedure and selected variables on the basis of seventy initial financial parameters and ratios by the following four steps:

1. Include all variables in Altman’s Z-score model.
2. Retain only one variable from each pair of variables with a correlation coefficient of 0.7 or over.
3. Exclude the variable which has lots of missing values from each pair of highly correlated variables.
4. Exclude the variable that is intuitively identified as less relevant to FDP from each pair of highly correlated variables which has the same number of missing values.

5.3.2. Quantitative feature extraction

Factor analysis (FA) and principal component analysis (PCA) are also applied to feature extraction for FDP [134,24,110,69]. Different from t-test or stepwise regression that selects efficient variables from initials ones, FA or PCA reduces dimensions by extracting factors or principal components, which are new variables computed from initial data. Tsai [121] compared five well-known feature selection methods for FDP, which are t-test, correlation matrix, stepwise regression, PCA and FA, and found that on average t-test is superior to others and stepwise regression is on the second position. This indicates that t-test is not only a simple feature selection method but also an efficient one. Ravi and Pramodh [95] improved performance of TAWNN in FDP by using features extracted by PCA. The result indicated that the PCTANN outperformed the compared techniques.

Besides, some other examples of quantitative feature methods for FDP also include the references [79,1,96,97,133,73] in the Section 3.2.1, where one algorithm is applied to choose features for another classification algorithm to generate a hybrid FDP method.

5.4. Comments on featuring methods for FDP

5.4.1. Advantage and disadvantage of qualitative and quantitative featuring methods

Either qualitative or quantitative featuring method has its own advantages and disadvantages. Pure subjective construction of feature system is usually appropriate for the exploratory stage of certain direction of FDP research when no similar literatures can be referred to. Feature system constructed by this type of qualitative method has strong subjectivity, and is sometimes lack of persuasiveness. Qualitative selection based on previous researches chooses representative indicators used by previous researches, and the scientificity of feature set is supported by the results of previous researches. It also has certain degree of subjectivity when adding or deleting indicators selectively to adapt to different backgrounds, and its applicability to other research background is not enough.

Quantitative feature selection is less affected by subjectivity, and it often can produce a feature set for FDP with greater information content and improve modeling efficiency. It is a mainstream method in current FDP research. However, it also may eventually produce an efficient final feature set that is hard to be explained or understood, and it is often unsuitable for constructing pure qualitative attribute system that is necessary for specific FDP method like group decision making.

5.4.2. Classifying featuring methods of FDP into filter and wrapper

Generally, there are two types of featuring methods: filter and wrapper. Filter approaches make use of human’s experiential expertise or statistical information of the dataset to carry out feature selection and are independent of the base classifier. Wrapper approaches, on the other hand, use classification accuracies or criteria derived from the classifier to rank the discriminative power of all (or part) of the possible feature subsets and select the subset that produces the best performance [83].

Therefore, qualitative selection of FDP features belongs to filter method, because features are chosen without quantitative analysis before certain kind of classification/prediction algorithm is used to construct FDP model. However, quantitative featuring methods may belong to either filter or wrapper. For example, the feature selection/extraction methods like t-test, correlation matrix, PCA...
and FA are all filter approaches. For stepwise regression like stepwise MDA or stepwise Logit, it belongs to wrapper when MDA or Logit is also used as the classification tool. However, it is a filter method when the variables selected by them are used as input of other classification algorithms. For the hybrid FDP method in which one algorithm is applied to choose features for another classification algorithm, filter feature selection is applied if feature selection and model construction are two independent stages, while it is a wrapper when feature selection and model construction are integrated together and a performance measure derived from the model is used to determine the result of feature selection.

6. Conclusions and future research topics in FDP

FDP is a hot exploratory issue in accounting, finance, business, and engineering areas. It is gradually forming its own theoretical system to be an independent subject. It has important practical significance on improving awareness of financial risk, preventing corporate financial distress, and avoiding bankruptcy liquidation. This review classified research in this subject into four categories: namely, definition of financial distress concept, FDP modeling methods, sampling methods, and featuring methods. Although abundant articles have been published in this field, yet there are still some valuable topics which need further explored and studies. The main issues that need to be further investigated are summarized as follows:

6.1. Developing simple but accurate predictive models

The current research objective of PDF is to find more accurate predictive models, since even a cautious reduction of predictive error rates of predictive models is able to help people reduce the loss in business significantly. In order to achieve this objective, several state-of-the-art techniques are tried to be integrated to construct an accurate model. However, by reviewing the achievements of fulfilling the objective of investigating more accurate FDP models, we find that lots of very complex models were developed. In order to find more accurate predictive FDP models, we may neglect the fact that FDP is a practice-oriented subject. The reduction on predictive error rate is not significant and valuable unless the model is applied in forecasting financial distress in real world. However, the complexity of models prevents them from being adopted by industrial organizations and users. Since the use of FDP as a managerial tool relates to financial income of an organization, the users commonly will not apply a technique in their work unless they understand the inside mechanism of the technique clearly. Thus, they need simple FDP models, which hold the characteristics of ease of interpretation, explanation, and understanding. Simple models do not surly belong to inaccurate models. Gepp et al. [38] provided some evidence that simpler models are able to produce better predictive powers.

6.2. Developing dynamic modeling methods for FDP

In real-world FDP practice, it is a critical problem how to consider the existence of FDCD and update FDP models effectively when they become outdated with time going on. Current literatures on FDP almost focus on algorithmic exploration that is on the assumption that the data set for current FDP modeling is already constructed, and dynamic variation of modeling data with gradual emergence of new financial distress samples is almost neglected. Although some pilot work on this topic has already been done, more effective methods for treating concept drift and dynamic FDP modeling with time process needs further attention and deep research.

6.3. Developing FDP models based on decision implementations

Most current researches of FDP were carried out from the perspective of quantitative prediction, and far less attention were paid to the importance of expertise and non-financial information. Generally speaking, before the outbreak of enterprise financial distress, not only quantitative financial indicators have abnormal signs, but also there exist many hidden peril of management that cannot be quantified, which are often the source of outbreak of financial distress. Therefore, it is necessary to break the traditional concentration on quantitative FDP models that are based on financial indicators, and to do further study on how to use expertise and non-financial information to predict financial distress. As an important branch in operational research and management science, group decision-making methods have been applied to solve many managerial problems and help improve management efficiency. However, its application to FDP is not paid enough attention. In fact, financial distress early warning based on group decision-making methods can provide important supplement for quantitative FDP models, and help managers find the reasons for enterprise financial distress, only on which effective managerial measures can be taken to prevent further deterioration of financial condition. Besides, more non-financial indicators should be considered in group decision-making FDP methods as well as traditional quantitative FDP models. The mechanism of avoiding financial distress from the enterprise operation also needs further investigation. The competition between firms happens between the entire supply chains. One firm falling into failure in a supply chain may force the other firms in the chain to fail. Thus, how to avoid financial distress from the view of supply chain management needs to be further investigated.

6.4. Developing imbalanced predictive models for FDP

The real-world problem of FDP consists of imbalanced datasets. The traditional use of pairing strategy to produce balanced data sets is a method for handling imbalanced samples, namely: the use of expertise of human beings to balance the real-world imbalanced data sets. With the development of imbalanced data processing, more techniques have demonstrated themselves useful in FDP modeling. These techniques chiefly include data level solution for modeling with imbalanced samples, and model level solution for modeling with imbalanced samples. The cost-sensitive learning technique is also a potential alternative for FDP imbalanced modeling. This topic needs further investigation, since it fits the real-world characteristics of FDP.

6.5. Developing financial distress early warning system

The previous researches have not formed the overall framework of financial distress early warning system, resulting in the situation which is full of concrete algorithmic research but lack of framework design of financial distress early warning information system. This is not only unfavorable to the formation and improvement of financial distress early warning system, but also limits the application of existing specific FDP methods to the practice of enterprise. Therefore, it is necessary to study on the logic system of financial distress early warning.

Acknowledgements

This research is partially supported by the National Natural Science Foundation of China (Nos. 71371171, and 71171179), the Humanity and Social Science Foundation of Ministry of Education of China (No. 13VJC830140), the Zhejiang Provincial Natural Science Foundation (LY13G010001, and LR13G010001), the Zhejiang
Provincial Philosophy and Social Science Foundation – Zhijiang Provincial Qianjiang Talent Foundation of China. The authors gratefully thank anonymous referees for their useful comments and editors for their work.

References


