Incorporating Media News to Predict Financial Distress: Case Study on Chinese Listed Companies

Lifang Zhang[1], Mohammad Zoynul Abedin[2], Zhenkun Liu[3]*

[1] School of Finance, Nanjing University of Finance and Economics, No.3 Wenyuan Road, Qixia District, Nanjing, Jiangsu Province 210023, China

[2] Department of Accounting and Finance, School of Management, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, Wales, the United Kingdom

[3] School of Management, Nanjing University of Posts and Telecommunications, No 66 Xinmofan Road, Gulou District, Nanjing, Jiangsu Province 210023, China

* Corresponding author: Zhenkun Liu

E-mail addresses: Lifang Zhang (zhanglf@nufe.edu.cn);

Mohammad Zoynul Abedin (m.z.abedin@swansea.ac.uk);

Zhenkun Liu (liuzk@njupt.edu.cn)

Funding

Financial interests and non-financial interests

23 **Incorporating Media News to Predict Financial Distress: Case Study**
24 **on Chinese Listed Companies**

25 **Abstract**

26   Financial distress prediction has been a prominent research field for several
27 decades. Accurate prediction of financial distress not only helps to safeguard the
28 interests of investors, but also improves the ability of managers to manage financial
29 risks. Prior studies predominantly rely on accounting metrics derived from financial
30 statements to predict financial distress. Our research takes a step further by
31 incorporating media news to enhance the accuracy of financial distress prediction.
32 Based on the data from Chinese listed companies, 7 classifiers are established to verify
33 the additional value of media news in improving the financial distress prediction
34 performance of models. Experimental results demonstrate that the inclusion of media
35 news in predictive models is effective as it contributes to better performance compared
36 to models that solely rely on accounting features. Moreover, random forest model is a
37 reliable tool in financial distress prediction due to its superior ability to capture complex
38 feature relationships. Evaluation indicators, statistical tests, and Bayesian A/B tests
39 further confirm that the inclusion of media news can significantly improve the
40 identification of financially distressed companies.

41 **Keywords:** Accounting Features, Financial Distress Prediction, Media News

1. Introduction

In today's rapidly evolving commercial landscape, an increasing number of enterprises are faced with the challenge of maintaining their competitive superiority in mature market (Khedmatgozar & Shahnazi, 2018). In the absence of a sturdy mechanism for adjusting to these alterations, corporations are at risk of confronting predicaments such as inadequate cash reserves and disproportionate obligations, ultimately resulting in financial adversity. Financial distress is a critical issue in the realm of commerce, as it can have a detrimental impact on a company's long-term growth and even threaten its survival (Peng et al., 2011). Of particular significance is the observation that financial distress can have adverse ramifications for the welfare of diverse stakeholders, encompassing investors, executives, staff, vendors, and regulatory bodies (Olson et al., 2012). In addition, the occurrence of a significant number of financially distressed firms can lead to a systemic financial crisis, which has the capacity to disrupt the stability of the entire capital market. Considering the detrimental effects of financial distress, it is not unexpected that scholars, professionals, and policymakers alike are actively seeking guidance from dependable financial distress prediction (FDP) models to furnish timely indications of impending distress. The advantages of FDP are manifold, encompassing the facilitation of informed business decisions by managers, the provision of additional investment insights for investors (Yan et al., 2016), and the enhancement of regulatory effectiveness for regulators. As a result, FDP has emerged as an essential pursuit for scholars, professionals, and policymakers across various domains (du Jardin, 2017; Wang et al., 2020; Wu, 2010).

Within the current landscape of research on FDP, there exist two primary avenues for improving prediction accuracy: the acquisition of effective features and the construction of credible models (Chou et al., 2017; Olson et al., 2012). Historically, accounting features have been the primary means of assessing a company's financial health, often obtained by converting its financial data into financial ratios. As such,

70 these features have been extensively employed to uncover a firm's financial condition

71 (Beaver, 1966; Geng et al., 2015). It is worth noting that non-financial data extracted

72 from textual disclosures can also be a valuable resource for FDP, despite the traditional

73 emphasis on accounting metrics (Hajek & Henriques, 2017). In contrast to accounting

74 features, non-accounting features derived from textual disclosures can provide insights

75 into the external factors that impact a company's operations, including cost management,

76 business tactics, and corporate governance. As such, these features can offer a more

77 holistic understanding of a firm's financial status. As a result, scholars have recently

78 directed considerable attention towards textual disclosures, investigating them as a

79 novel and effective means of developing a more robust FDP model (Schumaker et al.,

80 2012; Volkov et al., 2017). For instance, recent studies have examined the predictive

81 effectiveness of textual features to gain a better understanding of FDP, and the results

82 have demonstrated that they can serve as a crucial complement to accounting features

83 in the context of FDP, facilitating more accurate and dependable predictions

84 (Balakrishnan et al., 2010). More specifically, textual disclosures pertaining to the

85 economic and business landscape, as well as the corresponding policies of nations, have

86 been demonstrated to hold considerable importance in the realm of FDP (Doumpos et

87 al., 2017). Additionally, sentiment features, which present the tone of financial

88 documents, have also been found to contribute to the prediction ability (Chan & Chong,

89 2017; Hajek et al., 2014). The extant literature reviewed above collectively suggests

90 that the integration of both accounting and non-accounting features is crucial for

91 accurate financial distress prediction. Specifically, the amalgamation of financial ratios,

92 alongside sentiment and textual characteristics, has been incorporated to construct an

93 enhanced FDP framework that exhibits enhanced efficacy. Moreover, the utilization of

94 diverse features has demonstrated a significant correlation with augmented predictive

95 capabilities (Wang et al., 2018). Therefore, it is recommended that future research on

96 FDP should consider the inclusion of both accounting and non-accounting features,

97 with a particular emphasis on textual disclosures, to achieve a more comprehensive and

98   accurate prediction model.

99   Despite previous studies exploring the use of emotions in financial documents to
100   predict financial distress, there has been a dearth of research on the impact of emotions
101   in media news texts on the ability to predict financial distress. Our study adopts a novel
102   approach by constructing new emotional features based on media news from thousands
103   of sources to support financial distress prediction. We compare the performance of
104   seven models in predicting financial distress based on media news features or
105   accounting features alone. Using evaluation indicators, statistical tests, and Bayesian
106   A/B tests, we demonstrate that emotional features constructed based on media news
107   texts are valuable for predicting financial distress.

108   There are several compelling reasons why predicting financial distress is crucial
109   for making informed investment decisions. Firstly, it enables us to evaluate a company's
110   risk level, thereby allowing us to manage our investment portfolio more effectively. By
111   reducing the risk in our portfolio, we can safeguard our investments. Secondly,
112   predicting financial distress can facilitate more informed investment decisions. If we
113   can anticipate a company's financial distress, we can reduce our investment in that
114   company or sell our stocks when the stock price drops, thereby minimizing our losses.
115   Finally, predicting financial distress can help us assess a company's value more
116   accurately. If a company is facing financial distress, this may impact its stock and
117   market value. Therefore, predicting financial distress can enable us to make more
118   informed investment decisions by providing us with a better understanding of a
119   company's true value.

120   The following sections of this study will provide a comprehensive overview of our
121   research. Section 2 will introduce the relevant work on developing new features to
122   predict financial distress. In Section 3, we will outline the empirical methods used to
123   verify the prediction performance of our models. Section 4 will provide details on the
124   experimental setup, while Section 5 will present the experimental results. Finally, in
125   Section 6, we will elaborate on the conclusions drawn from our study.

2. Related Work

FDP is a complex problem that has garnered significant attention from academics over the past few decades. Through extensive research, it has become clear that acquiring effective features is a practical and effective way to improve FDP performance.

In the domain of FDP, accounting information has been found to exhibit a strong correlation with a company's financial performance (Lin et al., 2011). Accounting information can naturally reflect a company's creditworthiness, repayment ability, and operational status, making it a crucial source of predictive information for estimating financial distress. As a result, financial ratios, which can be computed using accounting information from a company's financial statements, have been effectively utilized in FDP. For example, Beaver (1966) was the pioneer in utilizing financial ratios to predict financial distress and demonstrated their effectiveness. Subsequently, a growing cohort of scholars has adopted accounting attributes in the context of FDP, thereby substantiating their noteworthy contribution to the development of precise prognostic models (Delen et al., 2013; Sun & Li, 2008). Geng et al. (2015) conducted an empirical study on FDP and employed a set of 31 financial ratios to evaluate their effectiveness. These metrics comprised net profit margin of total assets, return on total assets, and cash flow per share, alongside other relevant indicators. The results of their study demonstrated the efficiency of these financial ratios in FDP. Huang et al. (2017) conducted a study on FDP and employed a set of 95 financial ratios to construct predictive models. Through a rigorous selection process, they identified the most significant financial ratios and found that a smaller set of vital financial ratios performed well in FDP. Their study highlights the importance of selecting the most relevant financial ratios for FDP, rather than relying on a large set of ratios that may not be as informative.

To achieve a more precise estimation of a company's operational status, researchers have increasingly turned to non-accounting data, such as textual disclosures,

154   which can provide valuable insights into the management of expenses, the formulation

155   of corporate strategies, and the handling of managerial circumstances. Consequently,

156   the examination of non-accounting attributes in relation to their complementary

157   influence on traditional accounting features in the domain of FDP has garnered

158   considerable attention within the research community. These non-accounting features

159   have been demonstrated to hold substantial influence as determinants of FDP outcomes

160   (Schumaker et al., 2012; Volkov et al., 2017). These studies highlight the importance

161   of incorporating non-accounting features into FDP models to achieve a more

162   comprehensive and accurate prediction of financial distress. For instance, Cecchini et

163   al. (2010) conducted a study on FDP and investigated the role of textual features

164   extracted from financial texts. Through their analysis, they found that combining both

165   textual and accounting features outperformed benchmark methods that utilized only

166   accounting or textual features. Their study highlights the importance of incorporating

167   both types of features into FDP models to achieve a more accurate and comprehensive

168   prediction of financial distress. Doumpos et al. (2017) successfully incorporated non-

169   accounting variables, such as those pertaining to the macroeconomic and business

170   environment, into a predictive model for financial distress. By combining these

171   qualitative features with traditional accounting metrics, the authors were able to achieve

172   favorable outcomes in their analysis. Hájek et al. (2015) augmented traditional

173   accounting features with sentiment-based variables extracted from annual reports,

174   resulting in more accurate predictions and promising outcomes. Recent endeavors to

175   enhance the predictive power of financial distress prediction models have centered on

176   the integration of accounting and non-accounting features. **Table 1** provides an

177   overview of relevant research on developing new features to support financial distress

178   prediction.

179 **Table 1**

180 Relevant research on developing new features to support financial distress prediction.

Reference	New features	Validation mechanism	Model number	Parameters Setting	Test method	Main academic significance
(Geng et al., 2015)	Accounting indicators	10 times repeated sub-sampling validation	4	No	No	This study tested the FDP performance of four models based on accounting indicators.
(Liang et al., 2016)	Accounting indicators, and corporate governance indicators	10-fold cross validation	5	Subjective setting	No	This work is to fully examine the power of corporate government indicators combined with accounting indicators for FDP.
(Al-Hadi et al., 2019)	Accounting indicators and corporate social responsibility indicators	Intra sample fitting	6	No	T test	This study examines the association between corporate social responsibility and financial distress
(Hernandez Tinoco & Wilson, 2013)	Accounting indicators, market indicators and macro-economic indicators	Train and holdout	5	Subjective setting	Kolmogorov–Smirnov test	This study investigates the utility of combining accounting indicators, market indicators, and macro-economic indicators to predict financial distress.
(Kallunki & Pyykkö, 2013)	Accounting indicators and the personal characteristics of management	No	4	No	Vuong test	This paper proved that appointing CEOs and directors with past personal payment default entries increases the likelihood of financial distress of the firm.
(Li et al., 2021)	Sentiment features in annual reports	10-fold cross validation	4	No	No	This study constructed corpora based on annual reports, and found that sentiment features in annual reports can independently achieve relatively satisfactory FDP performance.
(Muñoz‐Izquierdo et al., 2020)	Accounting indicators and auditing indicators	No	5	No	Chi‐square test	This paper analyzed the usefulness of combining accounting indicators and auditing indicators in order to predict corporate financial distress.
(Liang et al., 2020)	Accounting indicators, and corporate governance indicators	10-fold cross validation	3	No	Wilcoxon test, and t test	By manually reading annual reports and proxy statements, this paper collected some novel corporate governance indicators.
(Wang et al., 2020)	Accounting indicators, and financial reports	10-fold cross validation	8	No	T test	A novel and robust framework considered both the prediction accuracy and the diversity in ensemble is proposed for the FDP incorporating multiple features.
(Jiang et al., 2022)	Accounting indicators, periodic reports, and current reports	10-fold cross validation	4	No	No	This paper proposed a framework to predict the financial distress of unlisted public firms using current reports.
(Zhao et al., 2022)	Accounting indicators, comments on online stock forums, etc	10-fold cross validation	6	No	No	In this study, comments on online stock forum were firstly used to predict financial distress.
Our study	Accounting indicators, and media news	10-fold cross validation	7	Grid search	T test, and Bayesian A/B test	Our study proves the additivity improvement of media news on the financial distress prediction.

181

8

3. Empirical Methods

Machine learning techniques have found extensive application in various business analysis scenarios (Abedin et al., 2019, 2021, 2023; Efat et al., 2022; Hajek et al., 2022; Lu et al., 2022; Yang et al., 2023). In this study, we leverage these techniques to validate the incremental value of media news in predicting FDP performance. In the following section, we present several empirical models that showcase the impact of media news information on forecasting the financial distress of publicly traded firms.

3.1. Classification and Regression Tree

Classification and regression tree (CART) is a well-known decision tree algorithm first introduced by Gordon et al. (1984). This algorithm operates under the assumption that the decision tree is a binary tree, and it utilizes a technique known as binary recursive partitioning to split each non-leaf node into two branches. These branches are then labeled as "True" or "False" based on the feature values associated with them. The modeling process for CART can be summarized as follows:

Given dataset $\mathbf{X} = \{x_i, y_i\}$, $(i = 1, 2, ..., n)$, where y_i is the category label of the corresponding sample x_i. $y_i \in \{c_1, c_2, \cdots, c_M\}$, where $c_m (m = 1, 2, ..., M)$ represents the category, M is the number of categories. If p_m is the probability of the m-th category, the GINI index can be expressed as follows:

$$GINI(p) = \sum_{m=1}^{M} p_m (1 - p_m) = 1 - \sum_{m=1}^{M} p_m^2 \tag{1}$$

For binary classification problem, $M=2$; thus, Eq. (1) can be expressed as:

$$GINI(p) = 2p(1 - p) \tag{2}$$

According to the preset segmentation value a of feature A_i, the sample $\mathbf{X}$ is divided into $\mathbf{X}_1$ and $\mathbf{X}_2$. Then, we can obtain:

$$GINI(\mathbf{X}, A_i)\big|_a = \frac{|\mathbf{X}_1|}{|\mathbf{X}|} GINI(\mathbf{X}_1) + \frac{|\mathbf{X}_2|}{|\mathbf{X}|} GINI(\mathbf{X}_2) \tag{3}$$

where $GINI(\mathbf{X}, A_i)\big|_a$ represents the GINI index of $\mathbf{X}$ decomposed by $A=a$. The larger the value of $GINI(\mathbf{X}, A_i)\big|_a$, the larger the uncertainty of the sample. Hence, it is

208 imperative to opt for the feature *A* that possesses the lowest GINI index as the splitting
209 feature for the root node. This approach should be applied iteratively on every non-leaf
210 node. Once the GINI index of an internal node surpasses that of its root node, a
211 comprehensive CART model is established.

212 Furthermore, it is crucial to prune the constructed CART in order to decrease the
213 likelihood of overfitting. This is typically achieved by minimizing the overall loss
214 function of the CART. The loss function can be defined as follows:

215 $$C_a(T) = C(T) + a|T| \qquad (4)$$

216 In this context, *T* and |*T*| represent the number of subtrees and their corresponding nodes,
217 respectively. *C*(*T*) denotes the prediction error of the training sample, while *a* serves as
218 a parameter to gauge the model's complexity and its fitting degree to the training sample.
219 To reduce the complexity and enhance prediction performance, CART prunes the
220 feature and its branches from the leaf nodes of the tree. This is done when the loss
221 function of CART decreases after pruning.

222 **3.2. Extremely Randomized Tree**

223 Extremely randomized tree (ERT) is an ensemble technique that involves training
224 numerous random decision trees on the training set and leveraging the average value to
225 enhance prediction accuracy and minimize the risk of overfitting (Geurts et al., 2006).
226 ERT bears a striking resemblance to random forest (RF), as both methods comprise
227 multiple decision trees. However, RF selects the best branching attribute from random
228 subsets, whereas ERT selects the branching value entirely at random, resulting in
229 greater uncertainty than RF. The primary modeling process of ERT can be succinctly
230 summarized as follows:
231 (1) Construct a training set of size *S*;
232 (2) Randomly select *n* learning samples without repetition from the training set *S*;
233 (3) Build a tree based on the entire learning sample. Randomly select feature *f* at
234 each node, and split the node with a random cutting point;
235 (4) Repeat steps 2 and 3 until reaching the leaf node;
236 (5) Aggregate the results of each tree and use majority voting to obtain the class

237   with the highest number of votes.

3.3. Gaussian Naive Bayes Classifier

239   Gaussian naive Bayes Classifier (GNBC) is a type of naive Bayes classifier (John

240   & Langley, 1995). The prior probabilities of GNBC follow a Gaussian distribution and

241   are mainly used to handle continuous variable values. The GNBC assumes that every

242   pair of features is independent of each other, such that given samples $(x_1, x_2, \cdots, x_N)$

243   and a class output variable $\mathbf{y}$, for any x_i, it satisfies:

244
$$P(x_i | y, x_1, x_2, \cdots, x_N) = P(x_i | y) \tag{5}$$

245   According to Bayes rule:

246
$$P(y | x_1, x_2, \cdots, x_N) = \frac{P(y) P(x_1, x_2, \cdots, x_N | y)}{P(x_1, x_2, \cdots, x_N)} \tag{6}$$

247
$$P(y | x_1, x_2, \cdots, x_N) = \frac{P(y) \prod_{i=1}^{N} P(x_i | y)}{P(x_1, x_2, \cdots, x_N)} \tag{7}$$

248   Since $P(x_1, x_2, \cdots, x_N)$ in the above equation is a constant, the following

249   classification rules are used:

250
$$P(y | x_1, x_2, \cdots, x_N) \propto P(y) \prod_{i=1}^{N} P(x_i | y) \tag{8}$$

251
$$\hat{y} = \arg \max_y P(y) \prod_{i=1}^{N} P(x_i | y) \tag{9}$$

252   where $\hat{y}$ represents the predicted value, $P(y)$ is the prior probability of the *y-th*

253   class sample in the training set. GNBC assumes that the probability of the features in

254   naive Bayes follows a gaussian distribution, and the probability density function of

255    $P(x_i | y)$ is:

256
$$P(x_i | y) = \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\delta_y^2} \right) \tag{10}$$

257   where μ_y and δ_y^2 are the mean and variance of the values of *y-th* class sample on the

258   *i-th* feature, which can be estimated by using maximum likelihood estimation method.

3.4. Least Absolute Shrinkage and Selection Operator

260   Least absolute shrinkage and selection operator (LASSO) is a linear regression-

261 based technique that utilizes L_1 regularization to achieve sparsity and feature selection

262 (Tibshirani, 1996). By incorporating an L_1 regularization term into the objective

263 function, LASSO can effectively reduce the weight of certain learned features to zero

264 or near-zero, resulting in feature selection and model sparsity. The objective function

265 of LASSO can be mathematically expressed as follows:

266
$$J_{Lasso}(\beta) = J_{Logistic}(\beta) + \lambda \sum \|\beta\|_1 \tag{11}$$

267
$$J_{Logistic}(\beta) = \sum \left[-y \cdot \log(g(h_\theta(X))) - (1-y) \cdot \log(1 - g(h_\theta(X))) \right] \tag{12}$$

268
$$J_{Lasso} = \sum \left[-y \cdot \log(g(h_\theta(X))) - (1-y) \cdot \log(1 - g(h_\theta(X))) \right] + \lambda \sum \sum \|\beta\|_1 \tag{13}$$

269 where $J_{Logistic}(\beta)$ represents the cross entropy of the logistic regression objective

270 function, $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$ is the regression coefficient, λ is the regularization

271 coefficient, $\lambda > 0$.

3.5. Linear Discriminant Analysis

273   Linear discriminant analysis (LDA) is a powerful tool for tackling classification

274 problems, particularly those involving multiple classes. Its effectiveness lies in its

275 ability to identify a linear combination of features that optimally separates different

276 classes, while simultaneously minimizing the variation within each class.

277   The main modeling process of LDA is as follows:

278   (1) Compute the mean vector for each class:

279
$$\mu_j = \frac{1}{N_j} \sum_{x \in X_j} x \tag{14}$$

280 where j=1, 2, X_j represents the j-th class samples, N_j represents the number of

281 samples in the j-th class, and μ_j represents the mean of the j-th class samples.

282   (2) Compute the covariance matrix of the samples:

283
$$\sum_j = \sum_{x \in X_j} (x - \mu_j)(x - \mu_j)^T \tag{15}$$

284   Then, define the within-class scatter matrix and between-class scatter matrix as

285 follows:

286
$$\mathbf{S}_w = \sum_1 + \sum_2 = \sum_{x \in X_1} (x - \mu_1)(x - \mu_1)^T + \sum_{x \in X_2} (x - \mu_2)(x - \mu_2)^T \tag{16}$$

287 $$\mathbf{S}_b = \left(\mu_1 - \mu_2\right)\left(\mu_1 - \mu_2\right)^T \tag{17}$$

288   (3) Define the objective function as maximizing the Fisher's function *J*:

289 $$\arg\max_w J\left(w\right) = \frac{w^T S_b w}{w^T S_w w} \tag{18}$$

290   where w is an arbitrary m-dimensional vector, m is the number of features, the

291   maximum value of J is the largest eigenvalue of $S_w^{-1}S_b$. Solving Eq. (18), we can

292   obtain $w = S_w^{-1}\left(\mu_1 - \mu_2\right)$, which determines the optimal projection direction. Project the

293   test set samples X_t onto the optimal projection direction to obtain the predictions $\hat{y}$:

294 $$\hat{y} = w^{\mathrm{T}} X_t \tag{19}$$

3.6. Multilayer Perceptron

296   The multilayer perceptron (MLP) is a type of feedforward artificial neural network

297   (ANN) (Jain et al., 1996). ANNs are composed of a series of output connections, each

298   with its own weight. The inputs are then multiplied by their respective weights, and the

299   resulting values are summed with the bias term. The output is then passed through an

300   activation function, which limits the range of the output value. The mathematical

301   expression for this process is as follows:

302 $$y_k = \varphi\left(\sum_{j=1}^{m} w_{kj} X_j + b_k\right) \tag{20}$$

303   where y_k is the output vector, k is the number of nodes in the hidden layer, X_j is the

304   input vector; w_{kj} and b_k represent the weight and the bias; $\varphi(\cdot)$ is the activation

305   function.

306   MLP, typically consist of an input layer, multiple hidden layers, and an output layer,

307   with all layers being fully connected. If the input layer contains M neurons, the output

308   of the first hidden layer can be expressed as:

309 $$O_{1,k_1} = \varphi\left(\sum_{m=1}^{M} W_{m,k_1} X_m\right),\ k_1 = 1,2,...,K \tag{21}$$

310   where W_{m,k_1} the weight between the input layer and the neurons in the hidden layer.

311   This output can be used as the input of the next hidden layer. The output of any neuron

312   in another hidden layer is:

313
$$O_{h,k_h} = \varphi\left(\sum_{k_{h-1}=1}^{K_{h-1}} W_{k_{h-1},k_h} O_{h-1,k_{h-1}}\right), \ k_1 = 1, 2, ..., K, \ h = 2, ..., N \tag{22}$$

314 where W_{k_{h-1},k_h} is the weight between the neurons in layer *h*-1 and layer *h* of the hidden

315 layer, K_h is the number of neurons in the input layer and the *hth* hidden layer, and *N* is

316 the number of hidden layers. The output layer neuron is obtained by weighted sum of

317 the last hidden layer, and its expression is:

318
$$y_l = \sum_{k_N=1}^{K_N} W_{k_N,l} O_{N,k_N}, \ l = 1, 2, ..., L \tag{23}$$

319 where $W_{k_N,l}$ is the weight between the *N-th* hidden layer neurons and the output layer,

320 and *L* is the number of neurons in the output layer.

321 *3.7. Random Forest (RF)*

322 RF is an ensemble learning algorithm proposed by Breiman (2001), which

323 involves creating an ensemble model of multiple decision trees. The algorithm works

324 by training multiple decision trees on randomly generated training sets. The category

325 of a test sample is then determined by the majority vote of each decision tree. The

326 modeling process of RF can be summarized as follows:

327 (1) Suppose the dataset is $\mathbf{X} = \{x_i, y_i\}, \ (i = 1, \ 2, \ ..., \ n)$, the number of weak

328 learners is *M*, the output results of the weak learners are *f(x)*, and the output results of

329 the strong learner assembled with the weak learners are *H(x)*.

330 (2) When training the *m-th* (*m*=1, 2, …, *M*) weak learner, *n* samples are randomly

331 drawn with replacement from $\mathbf{X}$ to form a new training set $\mathbf{X}_m$. This process is repeated

332 *M* times to obtain *M* training sets.

333 (3) Based on the *M* new training sets, *M* weak learners are trained. At each node

334 of the weak learner, a random feature subset is selected, and the optimal feature for

335 splitting the weak learner node is selected based on the maximum information gain

336 criterion. Then the weak learner's prediction result is calculated.

337 (4) The prediction results of *M* weak learners are summarized, and the category

338 with the highest number of votes is selected as the output of the strong learner.

339 *3.8. Bayesian A/B Test*

340 Bayesian A/B test is commonly used to compare the effectiveness of two or more

341 different versions of marketing strategies (Broemeling, 2009). Compared to traditional

342 statistical tests, Bayesian A/B test has several unique advantages and characteristics.

343 1. Easy interpretation of results: Bayesian A/B test typically presents results in the

344 form of probabilities, making them easier to understand and interpret. For example, we

345 can say that version A is 80% more effective than version B, rather than simply

346 reporting a *p*-value.

347 2. No need for fixed sample size: Bayesian A/B test allows for observation of

348 results during the experiment and stopping the experiment as needed. This means that

349 we do not need to-determine a fixed sample size, saving time and resources.

350 3. Ability to handle multiple experimental groups: Bayesian A/B test can easily be

351 extended to cases with multiple experimental groups, not just two. This allows us to

352 test multiple versions simultaneously, leading to faster identification of the optimal

353 solution.

354 In Bayesian A/B test, we are typically interested in the difference in effects between

355 experimental groups. To obtain the distribution of this difference, we can calculate the

356 difference between posterior distributions. This difference distribution can help us

357 answer probability questions about the difference in effects between experimental

358 groups, such as the probability that version A is more effective than version B

359 (Hoffmann et al., 2022).

360 **4. Experimental Details**

361 In this section, we introduce the data sets and the hyperparameter setting of

362 empirical methods.

363 *4.1. Datasets*

364 Extensive research on Chinese firms has explored the correlation between

365 financial distress and poor profitability. In China, companies listed as "Special

366 Treatment" are those that have encountered financial difficulties or other issues, such

367 as consecutive losses, securities regulation violations, or other risks. This designation

368 often results in decreased investor confidence, as these companies are viewed as high-

369 risk investments. Additionally, they may face higher financing costs and more stringent

370 regulatory requirements. Ding et al. (2008) found that Chinese companies marked as

371 such, which exhibited signs of deterioration in the year prior, had a high likelihood of

372 experiencing financial distress in the following year. By predicting which companies

373 will receive this label, firms can receive an early warning about potential bankruptcy

374 and take steps to improve their profitability. Therefore, we have chosen to use these

375 labeled companies as examples of financially distressed firms that may face bankruptcy

376 in the future.

377   We gather accounting metrics for publicly traded companies using the China Stock

378 Market & Accounting Research Database (CSMAR)[1]. Additionally, we utilize the

379 intelligent public opinion analysis database within CSMAR to extract the number of

380 comments and emotional indicators from various media sources regarding these

381 companies. Our accounting indicators cover eleven aspects of each company per year,

382 including solvency, ratio structure, development ability, risk level, dividend distribution,

383 operating ability, per share indicators, disclosed financial indicators, cash flow analysis,

384 relative value indicators, and profitability. As for media news indicators, CSMAR's

385 intelligent public opinion analysis database has collected news articles on listed

386 companies from thousands of sources, which we then summarize into average

387 evaluation indicators for each company. These indicators include monthly average

388 evaluation quantity and average evaluation sentiment score. The evaluation emotion

389 score is derived from the overall emotion level of all news in the database, with a higher

390 score indicating a more optimistic evaluation. In our study, we classified 235637 media

391 news items in 2018, 221056 media news items in 2019, 619963 media news items in

392 2020, and 2156731 media news items in 2021 according to the evaluated listed

393 company in the news. We then summarized this data into the media news features of

394 each company in each dataset, namely the monthly average number of media news

395 items and the average evaluation sentiment score.

---

[1] https://www.gtarsc.com

396   In our study, we utilized accounting indicators and media news indicators from the

397   previous year to forecast whether a company would enter a special processing state in

398   the upcoming year[2]. To achieve this, we established four datasets. These datasets were

399   designed to predict the company status in 2019 using the company indicators from 2018,

400   predict the company status in 2020 using the company indicators from 2019, predict

401   the company status in 2021 using the company indicators from 2020, and predict the

402   company status in 2022 using the company indicators from 2021.

403   To assess the effectiveness of media news in predicting financial distress, we

404   initially trained 7 classifiers using 426 accounting features. Subsequently, we trained 7

405   classifiers again using a total of 428 features, including 2 media news features, to

406   compare their FDP performance. To ensure a fair comparison between the two sets of

407   classifiers, we employed the ten-fold cross-validation method. This involved dividing

408   the dataset into 10 groups and sequentially using each group as the test set while using

409   the remaining samples as the training set. The details of our experiments are listed in

410   **Fig. 1**.

411
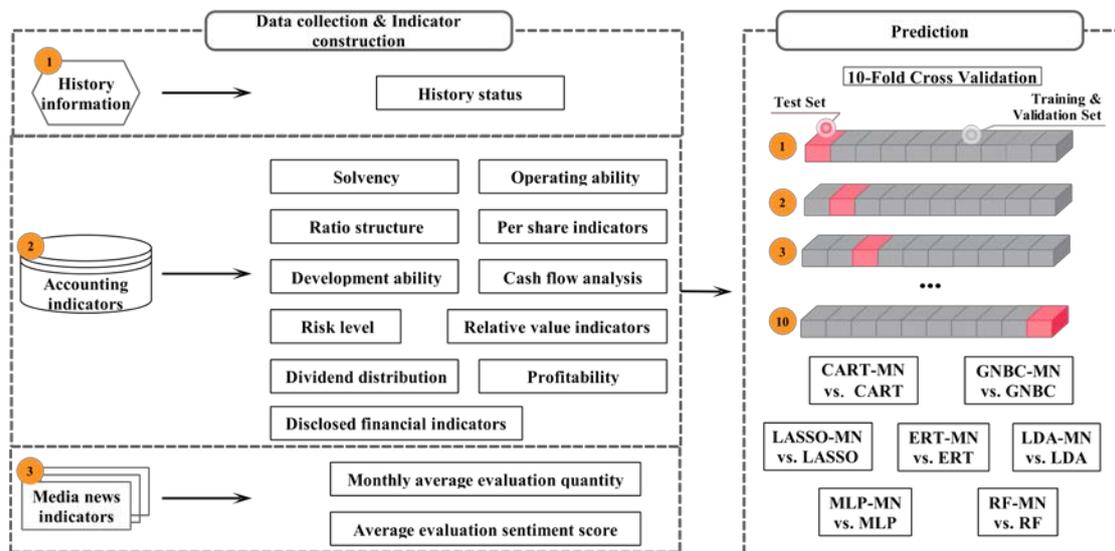
412

413   **Fig.1.** The details of our experiments.

---

[2] The meaning of features is listed in **Appendix**.

414 *4.2. Parameters Setting*

415   In our research, we carefully selected seven classifiers from various categories,

416 namely decision tree (CART), Bayesian classifier (GNBC), linear classifier (LASSO),

417 discriminant analysis method (LDA), neural network (MLP), and two ensemble

418 classifiers (ERT and RF). We referred to these classifiers as CART, GNBC, LASSO,

419 LDA, MLP, ERT, and RF when we trained them using accounting features. However,

420 when we trained them using both accounting and media news features, we referred to

421 them as CART-MN, GNBC-MN, LASSO-MN, LDA-MN, MLP-MN, ERT-MN, and

422 RF-MN. To ensure a fair comparison between the two groups of classifiers, we utilized

423 grid search to determine their hyperparameters. The hyperparameters of the classifiers

424 are presented in **Table 2**.

425

426 **Table 2**
427 Setting of candidate hyperparameters in all classifiers.

Models	Parameters	Candidate Values
CART & CART-MN	*min_samples_leaf*	[1, 25, 50, 75, 100]
	max_features	[0.01, 0.25, 0.5, 0.75, 1]
	min_samples_split	[2, 25, 50, 75, 100]
	max_depth	[3, 5, 7, 9]
ERT &ERT-MN	*n_estimators*	100
	min_samples_split	[2, 25, 50, 75, 100]
	max_depth	[3, 5, 7, 9]
	min_samples_leaf	[1, 25, 50, 75, 100]
GNBC & GNBC-MN	-	-
LASSO & LASSO-MN	The reciprocal of the coefficient of the regularization term	[0.0001, 0.001, 0.01, 0.25, 0.5, 0.75, 1]
LDA & LDA-MN	-	-
MLP & MLP-MN	*max_iter*	100
	hidden_layer_sizes	[1, 25, 50, 75, 100]
	learning_rate_init	[0.01, 0.25, 0.5, 0.75, 1]
	batch_size	[128, 256, 512, 1024]
RF & RF-MN	*n_estimators*	100
	min_samples_split	[2, 25, 50, 75, 100]
	max_depth	[3, 5, 7, 9]
	min_samples_leaf	[1, 25, 50, 75, 100]

5. Experimental Results

In this section, we aim to showcase the enhanced predictive capabilities of FDP models by incorporating media news features. To achieve this, we compare the performance of the classifiers in two sets from various angles, highlighting the additive value of media news features.

5.1. Comparison of Prediction Performance based on Evaluation Indicators

In this section, we utilize evaluation metrics to compare the predictive performance of various classifiers. As shown in **Table 3**, the two ensemble classifiers, ERT and RF, exhibit the higher prediction accuracy across four datasets. This suggests that when conducting FDP, ensemble classifiers tend to perform better. Additionally, classifiers that consider both accounting indicators and media news demonstrate superior FDP performance compared to those that solely consider accounting indicators. It is worth noting that MLP, as a neural network, exhibits lower prediction accuracy when media news is not considered. However, upon incorporating media news features, the FDP performance of MLP-MN is only marginally inferior to ERT-MN and RF-MN.

To summarize, both ERT and RF demonstrated high accuracy in predicting outcomes, regardless of whether media news was considered or not. However, MLP only achieved high accuracy when media news was considered. By comparing the FDP performance of different classifiers with and without media news, we can conclude that incorporating media news features can enhance the FDP performance of classifiers. Based on the evaluation indicators, it can be inferred that media news information can aid investors and decision-makers in comprehending the financial status of listed companies.

Table 3
Evaluation indicators of all classifiers on four datasets.

Models	Dataset 1	Dataset 2	Dataset 3	Dataset 4

CART	0.8936 (0.0679)	0.9431 (0.0268)	0.9209 (0.0518)	0.8935 (0.0426)
CART-MN	0.9204 (0.0440)	0.9606 (0.0259)	0.9408 (0.0387)	0.9087 (0.0468)
ERT	0.9465 (0.0309)	0.9740 (0.0211)	0.9579 (0.0274)	0.9318 (0.0539)
ERT-MN	0.9571 (0.0247)	**0.9780** (0.0185)	**0.9679** (0.0210)	0.9398 (0.0473)
GNBC	0.5229 (0.0567)	0.7444 (0.0171)	0.6856 (0.0566)	0.5737 (0.0518)
GNBC-MN	0.5232 (0.0565)	0.7456 (0.0181)	0.6859 (0.0567)	0.5738 (0.0518)
LASSO	0.9310 (0.0352)	0.9701 (0.0168)	0.9365 (0.0353)	0.9069 (0.0496)
LASSO-MN	0.9320 (0.0354)	0.9706 (0.0168)	0.9393 (0.0354)	0.9095 (0.0503)
LDA	0.9003 (0.0642)	0.9330 (0.0644)	0.8609 (0.0859)	0.8569 (0.0768)
LDA-MN	0.9005 (0.0647)	0.9334 (0.0644)	0.8610 (0.0861)	0.8571 (0.0766)
MLP	0.8104 (0.2154)	0.7350 (0.2473)	0.6670 (0.2121)	0.8087 (0.2148)
MLP-MN	0.9512 (0.0241)	0.9768 (0.0126)	0.9414 (0.0426)	0.9400 (0.0371)
RF	0.9411 (0.0398)	0.9728 (0.0303)	0.9564 (0.0324)	0.9315 (0.0473)
RF-MN	**0.9581** (0.0277)	0.9780 (0.0264)	0.9650 (0.0282)	**0.9447** (0.0359)

5.2. Comparison of Prediction Performance based on Statistical Test

In our study, we not only evaluate indicators, but also conduct statistical testing to demonstrate the superiority of classifiers that consider media news over those that do not. Our focus is on determining whether the use of media news can enhance FDP accuracy, and we compare pairs of classifiers using t-tests. For instance, we compare CART (0.8936, 0.9431, 0.9209, and 0.8935) vs. CART-MN (0.9204, 0.9606, 0.9408, and 0.9087) by using the mean evaluation indicator values obtained from four datasets. The statistical testing reveals that almost all pairs of comparison have a significant difference at 95% confidence level, except for GNBC-MN vs. GNBC, which has a significant difference at 85% confidence level, and LASSO-MN vs. LASSO, which has a significant difference at 90% confidence level. We attribute the weak advantage of GNBC-MN over GNBC to the poor predictive ability of GNBC, as evidenced by the evaluation indicator values in **Table 3**. Furthermore, we find that the combination of media news and accounting indicators is crucial to improving FDP performance, as demonstrated by the significant difference in FDP performance between LASSO-MN and LASSO at a 90% confidence level. The limited improvement of LASSO-MN is due to the inability of linear methods to construct combined features.

Table 4

Statistical test results between two groups of classifiers on four datasets.

Control Model	Benchmark Model	*p*-value	*t*-statistic
CART-MN	CART	0.0042	7.9163***
ERT-MN	ERT	0.0120	5.4668***
GNBC-MN	GNBC	0.1493	1.9292*
LASSO-MN	LASSO	0.0573	3.0076**
LDA-MN	LDA	0.0374	3.5762***
MLP-MN	MLP	0.0119	5.4883***
RF-MN	RF	0.0238	4.2539***

Note: *, **, *** denote the H0 is rejected at 85%, 90%, and 95% confidence level.

5.3. Comparison of Prediction Performance based on Bayesian A/B Test

In this section, we utilize Bayesian A/B testing to compare the FDP performance of various classifiers. To begin, we will calculate the performance rank of each classifier, with higher prediction accuracy resulting in a higher rank. Once the ranks have been determined, we will input them into the Bayesian A/B test to calculate the probability of each classifier being the optimal choice, as well as the expected loss associated with using it as the optimal classifier. **Table 5** displays the results of the Bayesian A/B test conducted between two groups of classifiers across four datasets. It is worth noting that the results were consistent across all groups, as the classifiers that took media news into consideration outperformed those that did not across all four datasets. This consistency can be attributed to the fact that the same ranks were inputted into the Bayesian A/B test.

Table 5

Bayesian A/B test results between two groups of classifiers on four datasets.

Model	CART	CART-MN
Probability Being Best	0.43%	99.57%
Expected Loss	0.6650	0.0004
Model	ERT	ERT-MN
Probability Being Best	0.43%	99.57%
Expected Loss	0.6650	0.0004
Model	GNBC	GNBC-MN
Probability Being Best	0.43%	99.57%

Expected Loss	0.6650	0.0004
Model	**LASSO**	**LASSO-MN**
Probability Being Best	0.43%	99.57%
Expected Loss	0.6650	0.0004
Model	**LDA**	**LDA-MN**
Probability Being Best	0.43%	99.57%
Expected Loss	0.6650	0.0004
Model	**MLP**	**MLP-MN**
Probability Being Best	0.43%	99.57%
Expected Loss	0.6650	0.0004
Model	**RF**	**RF-MN**
Probability Being Best	0.43%	99.57%
Expected Loss	0.6650	0.0004

491

492   Simultaneously, we conducted Bayesian A/B tests on all classifiers and presented

493 the results in **Table 6**. It is evident that classifiers that take media news into account

494 have a greater likelihood of becoming the optimal classifier than those that do not.

495 Furthermore, all classifiers that consider media news as the optimal classifier have

496 lower expected losses than those that do not. Notably, the Bayesian A/B test results

497 based on all classifiers indicate that RF-MN has the highest probability of becoming

498 the optimal classifier and the lowest expected loss when used as the optimal classifier.

499 From a Bayesian testing perspective, RF-MN can be a viable solution for decision-

500 makers to consider when conducting FDP.

501

502 **Table 6**

503 Bayesian A/B test results of all classifiers on four datasets.

Models	Probability Being Best	Expected Loss
CART	0.64%	2.5907
CART-MN	1.77%	2.0869
ERT	12.10%	1.3149
ERT-MN	25.90%	0.9308
GNBC	0.17%	3.4833
GNBC-MN	0.23%	3.2582
LASSO	1.48%	2.1445
LASSO-MN	3.14%	1.8678
LDA	0.36%	2.8630

LDA-MN	0.53%	2.6474
MLP	0.18%	3.3775
MLP-MN	15.35%	1.2065
RF	6.92%	1.5458
RF-MN	**31.27%**	**0.8260**

6. Conclusion

Financial distress has emerged as a pressing issue due to its detrimental impact on economic agents and the commercial environment. Accurately predicting financial distress can safeguard the interests of investors and enhance the ability of company managers to manage financial risk. Despite its significance, media news has often been overlooked in FDP. Extracting semantic features from media news, which reflect public opinions and emotions towards listed companies, is a challenging yet meaningful task. To demonstrate the additive contribution of media news to FDP performance, we employed several classifiers. Our evaluation of prediction performance through various indicators, statistical tests, and Bayesian A/B tests confirms the effectiveness of media news in FDP. Notably, the semantic features in media news, particularly those related to beneficial effects and technological novelty, are particularly prominent. Therefore, we recommend that listed companies implement policies to cultivate a positive media image.

Prior research has primarily focused on developing complex models to enhance prediction performance in FDP, with little attention paid to the potential impact of media news on FDP performance. This study aims to address this gap and enrich the research content of FDP. Notably, this is the first study to incorporate media news into FDP, with the aim of producing more effective prediction results by incorporating public opinions on listed companies. These insights can guide individual managers in taking measures to prevent financial distress, thereby enriching the theories of risk management.

Enterprises FDP can help investors reduce investment risk. Based on the FDP results, investors can adjust their investment amount and investment ratio. Predicting

528 financial distress can also help investors increase their investment return. If investors

529 can predict financial distress in advance, they can timely withdraw their investment

530 before the company's financial distress occurs, and obtain higher investment returns.

531 Predicting financial distress can also help investors improve the rationality of their

532 investment decisions. For example, investors can conduct more in-depth analysis of the

533 company's financial situation, or evaluate the investment value of the company. These

534 measures can help investors improve the rationality of their investment decisions,

535 reduce investment risks, and increase investment returns.

536 Predicting financial distress of companies can help market regulators improve the

537 efficiency of market supervision. Market regulators need to monitor the financial

538 condition of companies to prevent the risk of financial distress. Predicting financial

539 distress of companies can help market regulators enhance market transparency. If

540 market regulators can predict the occurrence of financial distress in advance, they can

541 timely disclose the financial condition of companies and improve market transparency.

542 There are two shortcomings of our proposed model. One is that the prediction

543 performance of the proposed model has not been verified in other fields, which should

544 be further improved in future research. The other is that the number of the accounting

545 features and the media news features are limited. In future, we will consider exploring

546 more features that can reflect the company's financial situation and accurately predict

547 the company's financial situation.

548 **List of abbreviations**

FDP	Financial distress prediction
ERT	Extremely randomized tree
GNBC	Gaussian naive bayes classifier
LDA	Linear discriminant analysis
ANN	Artificial neural network
CART	Classification and regression tree
RF	Random forest
LASSO	Least absolute shrinkage and selection operator
MLP	Multilayer perceptron
ERT-MN	Extremely randomized tree trained based on accounting and media news features
GNBC-MN	Gaussian naive bayes classifier trained based on accounting and media news features
LDA-MN	Linear discriminant analysis trained based on accounting and media news features

CART-MN	Classification and regression tree trained based on accounting and media news features
RF-MN	Random forest trained based on accounting and media news features
LASSO-MN	Least absolute shrinkage and selection operator trained based on accounting and media news features
MLP-MN	Multilayer perceptron trained based on accounting and media news features
CSMAR	China Stock Market & Accounting Research Database

549 **Funding**

550   This research has not any fundings.

551 **Financial interests and non-financial interests**

552   The authors have no relevant financial or non-financial interests to disclose.

553

554 **Reference**

555 Abedin, M. Z., Guotai, C., Moula, F., Azad, A. S. M. S., & Khan, M. S. U. (2019).
556   Topological applications of multilayer perceptrons and support vector machines
557   in financial decision support systems. *International Journal of Finance &*
558   *Economics*, *24*(1), 474–507. https://doi.org/10.1002/ijfe.1675

559 Abedin, M. Z., Hajek, P., Sharif, T., Satu, M. S., & Khan, M. I. (2023). Modelling
560   bank customer behaviour using feature engineering and classification techniques.
561   *Research in International Business and Finance*, *65*, 101913.
562   https://doi.org/10.1016/j.ribaf.2023.101913

563 Abedin, M. Z., Moon, M. H., Hassan, M. K., & Hajek, P. (2021). Deep learning-based
564   exchange rate prediction during the COVID-19 pandemic. *Annals of Operations*
565   *Research*. https://doi.org/10.1007/s10479-021-04420-6

566 Al-Hadi, A., Chatterjee, B., Yaftian, A., Taylor, G., & Monzur Hasan, M. (2019).
567   Corporate social responsibility performance, financial distress and firm life
568   cycle: evidence from Australia. *Accounting & Finance*, *59*(2), 961–989.
569   https://doi.org/10.1111/acfi.12277

570 Balakrishnan, R., Qiu, X. Y., & Srinivasan, P. (2010). On the predictive ability of
571   narrative disclosures in annual reports. *European Journal of Operational*
572   *Research*, *202*(3), 789–801. https://doi.org/10.1016/j.ejor.2009.06.023

573 Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of*
574   *Accounting Research*, *4*, 71–111. https://doi.org/10.2307/2490171

575 Breiman, L. (2001). Random forests. *Machine Learning*.
576   https://doi.org/10.1023/A:1010933404324

577 Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and
578   Regression Trees. *Biometrics*, *40*(3), 874. https://doi.org/10.2307/2530946

579 Broemeling, L. D. (2009). Bayesian methods for measures of agreement. In *Bayesian*
580   *Methods for Measures of Agreement*. https://doi.org/10.1111/j.1467-
581   985x.2009.00624_2.x

582 Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Making words work:
583   Using financial text as a predictor of financial events. *Decision Support Systems*,
584   *50*(1), 164–175. https://doi.org/10.1016/j.dss.2010.07.012

585 Chan, S. W. K., & Chong, M. W. C. (2017). Sentiment analysis in financial texts.
586   *Decision Support Systems*, *94*, 53–64. https://doi.org/10.1016/j.dss.2016.10.006

587 Chou, C.-H., Hsieh, S.-C., & Qiu, C.-J. (2017). Hybrid genetic algorithm and fuzzy
588   clustering for bankruptcy prediction. *Applied Soft Computing*, *56*, 298–316.
589   https://doi.org/10.1016/j.asoc.2017.03.014

590 Delen, D., Kuzey, C., & Uyar, A. (2013). Measuring firm performance using financial
591   ratios: A decision tree approach. *Expert Systems with Applications*, *40*(10),
592   3970–3983. https://doi.org/10.1016/j.eswa.2013.01.012

593 Ding, Y., Song, X., & Zen, Y. (2008). Forecasting financial condition of Chinese
594   listed companies based on support vector machine. *Expert Systems with*

595   *Applications*, *34*(4), 3081–3089. https://doi.org/10.1016/j.eswa.2007.06.037

596   Doumpos, M., Andriosopoulos, K., Galariotis, E., Makridou, G., & Zopounidis, C.

597   (2017). Corporate failure prediction in the European energy sector: A

598   multicriteria approach and the effect of country characteristics. *European*

599   *Journal of Operational Research*, *262*(1), 347–360.

600   https://doi.org/10.1016/j.ejor.2017.04.024

601   du Jardin, P. (2017). Dynamics of firm financial evolution and bankruptcy prediction.

602   *Expert Systems with Applications*, *75*, 25–43.

603   https://doi.org/10.1016/j.eswa.2017.01.016

604   Efat, M. I. A., Hajek, P., Abedin, M. Z., Azad, R. U., Jaber, M. Al, Aditya, S., &

605   Hassan, M. K. (2022). Deep-learning model using hybrid adaptive trend

606   estimated series for modelling and forecasting sales. *Annals of Operations*

607   *Research*. https://doi.org/10.1007/s10479-022-04838-6

608   Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical

609   study of listed Chinese companies using data mining. *European Journal of*

610   *Operational Research*. https://doi.org/10.1016/j.ejor.2014.08.016

611   Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine*

612   *Learning*. https://doi.org/10.1007/s10994-006-6226-1

613   Hajek, P., Abedin, M. Z., & Sivarajah, U. (2022). Fraud Detection in Mobile Payment

614   Systems using an XGBoost-based Framework. *Information Systems Frontiers*.

615   https://doi.org/10.1007/s10796-022-10346-6

616   Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent

617   detection of financial statement fraud – A comparative study of machine learning

618   methods. *Knowledge-Based Systems*, *128*, 139–152.

619   https://doi.org/10.1016/j.knosys.2017.05.001

620   Hajek, P., Olej, V., & Myskova, R. (2014). Forecasting corporate financial

621   performance using sentiment in annual reports for stakeholders' decision-

622   making. *Technological and Economic Development of Economy*, *20*(4), 721–

623   738. https://doi.org/10.3846/20294913.2014.979456

624   Hájek, P., Olej, V., & Myšková, R. (2015). *Predicting Financial Distress of Banks*

625   *Using Random Subspace Ensembles of Support Vector Machines BT   -*

626   *Artificial Intelligence Perspectives and Applications* (R. Silhavy, R. Senkerik, Z.

627   K. Oplatkova, Z. Prokopova, & P. Silhavy (eds.); pp. 131–140). Springer

628   International Publishing.

629   Hernandez Tinoco, M., & Wilson, N. (2013). Financial distress and bankruptcy

630   prediction among listed companies using accounting, market and

631   macroeconomic variables. *International Review of Financial Analysis*, *30*, 394–

632   419. https://doi.org/10.1016/j.irfa.2013.02.013

633   Hoffmann, T., Hofman, A., & Wagenmakers, E. J. (2022). Bayesian Tests of Two

634   Proportions: A Tutorial with R and JASP. *Methodology*.

635   https://doi.org/10.5964/meth.9263

636   Huang, J., Wang, H., & Kochenberger, G. (2017). Distressed Chinese firm prediction

637   with discretized data. *Management Decision*, *55*(5), 786–807.

638   https://doi.org/10.1108/MD-08-2016-0546

639 Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A

640   tutorial. In *Computer*. https://doi.org/10.1109/2.485891

641 Jiang, C., Lyu, X., Yuan, Y., Wang, Z., & Ding, Y. (2022). Mining semantic features

642   in current reports for financial distress prediction: Empirical evidence from

643   unlisted public firms in China. *International Journal of Forecasting*, *38*(3),

644   1086–1099. https://doi.org/10.1016/j.ijforecast.2021.06.011

645 John, G. H., & Langley, P. (1995). Estimating Continuous Distributions in Bayesian

646   Classifiers George. *Proceedings of the Eleventh Conference on Uncertainty in*

647   *Artificial Intelligence.*

648 Kallunki, J.-P., & Pyykkö, E. (2013). Do defaulting CEOs and directors increase the

649   likelihood of financial distress of the firm? *Review of Accounting Studies*, *18*(1),

650   228–260. https://doi.org/10.1007/s11142-012-9203-x

651 Khedmatgozar, H. R., & Shahnazi, A. (2018). The role of dimensions of perceived

652   risk in adoption of corporate internet banking by customers in Iran. *Electronic*

653   *Commerce Research*, *18*(2), 389–412. https://doi.org/10.1007/s10660-017-9253-

654   z

655 Li, S., Shi, W., Wang, J., & Zhou, H. (2021). A Deep Learning-Based Approach to

656   Constructing a Domain Sentiment Lexicon: a Case Study in Financial Distress

657   Prediction. *Information Processing & Management*, *58*(5), 102673.

658   https://doi.org/10.1016/j.ipm.2021.102673

659 Liang, D., Lu, C.-C., Tsai, C.-F., & Shih, G.-A. (2016). Financial ratios and corporate

660   governance indicators in bankruptcy prediction: A comprehensive study.

661   *European Journal of Operational Research*, *252*(2), 561–572.

662   https://doi.org/10.1016/j.ejor.2016.01.012

663 Liang, D., Tsai, C.-F., Lu, H.-Y. (Richard), & Chang, L.-S. (2020). Combining

664   corporate governance indicators with stacking ensembles for financial distress

665   prediction. *Journal of Business Research*, *120*, 137–146.

666   https://doi.org/10.1016/j.jbusres.2020.07.052

667 Lin, F., Liang, D., & Chen, E. (2011). Financial ratio selection for business crisis

668   prediction. *Expert Systems with Applications*, *38*(12), 15094–15102.

669   https://doi.org/10.1016/j.eswa.2011.05.035

670 Lu, Y., Yang, L., Shi, B., Li, J., & Abedin, M. Z. (2022). A novel framework of credit

671   risk feature selection for SMEs during industry 4.0. *Annals of Operations*

672   *Research*. https://doi.org/10.1007/s10479-022-04849-3

673 Muñoz-Izquierdo, N., Laitinen, E. K., Camacho-Miñano, M., & Pascual-Ezama, D.

674   (2020). Does audit report information improve financial distress prediction over

675   Altman's traditional Z-Score model? *Journal of International Financial*

676   *Management & Accounting*, *31*(1), 65–97. https://doi.org/10.1111/jifm.12110

677 Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining

678   methods for bankruptcy prediction. *Decision Support Systems*, *52*(2), 464–473.

679 https://doi.org/10.1016/j.dss.2011.10.007

680 Peng, Y., Wang, G., Kou, G., & Shi, Y. (2011). An empirical study of classification

681 algorithm evaluation for financial risk prediction. *Applied Soft Computing*, *11*(2),

682 2906–2915. https://doi.org/10.1016/j.asoc.2010.11.028

683 Schumaker, R. P., Zhang, Y., Huang, C.-N., & Chen, H. (2012). Evaluating sentiment

684 in financial news articles. *Decision Support Systems*, *53*(3), 458–464.

685 https://doi.org/10.1016/j.dss.2012.03.001

686 Sun, J., & Li, H. (2008). Listed companies' financial distress prediction based on

687 weighted majority voting combination of multiple classifiers. *Expert Systems*

688 *with Applications*, *35*(3), 818–827. https://doi.org/10.1016/j.eswa.2007.07.045

689 Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of*

690 *the Royal Statistical Society: Series B (Methodological)*.

691 https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

692 Volkov, A., Benoit, D. F., & Van den Poel, D. (2017). Incorporating sequential

693 information in bankruptcy prediction with predictors based on Markov for

694 discrimination. *Decision Support Systems*, *98*, 59–68.

695 https://doi.org/10.1016/j.dss.2017.04.008

696 Wang, G., Chen, G., & Chu, Y. (2018). A new random subspace method

697 incorporating sentiment and textual information for financial distress prediction.

698 *Electronic Commerce Research and Applications*, *29*, 30–49.

699 https://doi.org/10.1016/j.elerap.2018.03.004

700 Wang, G., Ma, J., Chen, G., & Yang, Y. (2020). Financial distress prediction:

701 Regularized sparse-based Random Subspace with ER aggregation rule

702 incorporating textual disclosures. *Applied Soft Computing*, *90*, 106152.

703 https://doi.org/10.1016/j.asoc.2020.106152

704 Wu, W.-W. (2010). Beyond business failure prediction. *Expert Systems with*

705 *Applications*, *37*(3), 2371–2376. https://doi.org/10.1016/j.eswa.2009.07.056

706 Yan, Y., Lv, Z., & Hu, B. (2016). Building Investor Trust in the P2P Lending

707 Platform with a Focus on Chinese P2P Lending Platforms. *2016 International*

708 *Conference on Identification, Information and Knowledge in the Internet of*

709 *Things (IIKI)*, 470–474. https://doi.org/10.1109/IIKI.2016.15

710 Yang, C., Abedin, M. Z., Zhang, H., Weng, F., & Hajek, P. (2023). An interpretable

711 system for predicting the impact of COVID-19 government interventions on

712 stock market sectors. *Annals of Operations Research*.

713 https://doi.org/10.1007/s10479-023-05311-8

714 Zhao, S., Xu, K., Wang, Z., Liang, C., Lu, W., & Chen, B. (2022). Financial distress

715 prediction by combining sentiment tone features. *Economic Modelling*, *106*,

716 105709. https://doi.org/10.1016/j.econmod.2021.105709

717

Media News Indicators	Feature 1	Monthly Average Number of Media News Items
	Feature 2	Average Evaluation Sentiment Score
History Status	Feature 3	Financial Status of the Company in the Previous Year
Accounting Indicators (debt paying ability)	Feature 4	Current Ratio
	Feature 5	Quick Ratio
	Feature 6	Conservative Quick Ratio
	Feature 7	Cash to Current Ratio
	Feature 8	Working Capital to Loan Ratio
	Feature 9	Working Capital
	Feature 10	Interest Cover A
	Feature 11	Interest Cover B
	Feature 12	Net Cash Flow from Operating Activities/Current Liabilities
	Feature 13	Cash Flow Interest Cover
	Feature 14	Cash Flow to Debt Ratio
	Feature 15	Debt Assets Ratio
	Feature 16	Long Debt to Asset
	Feature 17	Tangible Debt Assets Ratio
	Feature 18	Tangible Asset Interest Bearing Debt Ratio
	Feature 19	Equity Multiplier
	Feature 20	Debt Equity Ratio
	Feature 21	Equity-to-Debt Ratio
	Feature 22	Debt-to-Long Capital Ratio
	Feature 23	Long Debt to Equity
	Feature 24	Long Debt to Working Capital
	Feature 25	Earning before Interest, Taxes, Depreciation and Amortization to Total Liabilities
	Feature 26	Net Cash Flow from Operations to Total Liabilities
	Feature 27	Net Cash Flow from Operations to Interest-Bearing Debt
	Feature 28	Debt-to-Equity Market Ratio
	Feature 29	Debt Tangible Equity Ratio
	Feature 30	Fixed Charge Coverage Ratio
	Feature 31	Equity Multiplier 2
Accounting Indicators (disclosed financial indicators)	Feature 32	Extraordinary Item
	Feature 33	Net Profit Attributable to Shareholders of Listed Companies Excluding Extraordinary Item
	Feature 34	Weighted Average Return on Equity
	Feature 35	Weighted Average Return on Equity Excluding Extraordinary Item
	Feature 36	Basic Earnings Per Share Excluding Extraordinary Item
	Feature 37	Per Share Net Cash Flow from Operations
	Feature 38	Net Asset Per Share Attributable to Shareholders of Listed Companies

	Feature 39	Basic Earnings Per Share
	Feature 40	Fully Diluted Earnings Per Share
Accounting Indicators (ratio structure)	Feature 41	Current Asset Ratio
	Feature 42	Cash Asset Ratio
	Feature 43	Accounts Receivable Ratio
	Feature 44	Working Capital to Current Asset
	Feature 45	Working Capital Ratio
	Feature 46	Working Capital to Net Asset
	Feature 47	Non-current Asset Ratio
	Feature 48	Fixed Asset Ratio
	Feature 49	Intangible Asset Ratio
	Feature 50	Tangible Asset Ratio
	Feature 51	Owner's Equity Ratio
	Feature 52	Retained Earnings to Asset Ratio
	Feature 53	Long Asset Fit Asset
	Feature 54	Equity to Fixed Asset Ratio
	Feature 55	Current Liability Ratio
	Feature 56	Operating Liability Ratio
	Feature 57	Financial Liability Ratio
	Feature 58	Non-Current Liability Ratio
	Feature 59	Proportion of Equity Attributable to Parent Company Owners
	Feature 60	Proportion of Equity Attributable to Minority Shareholders
	Feature 61	Main Operating Profit Ratio
	Feature 62	Financial Activities Profit Ratio
	Feature 63	Operating Profit Ratio
	Feature 64	Non-operating Revenue Ratio
	Feature 65	Turnover Tax Rate
	Feature 66	Consolidated Tax Rate A
	Feature 67	Consolidated Tax Rate B
	Feature 68	Income Tax Rate
	Feature 69	Proportion of Net Profit of Parent Company
	Feature 70	Proportion of Net Profit Attributable to Minority Shareholders
	Feature 71	Proportion of Consolidated Income to Net Profit
	Feature 72	Proportion of Other Consolidated Income
	Feature 73	Proportion of Comprehensive Income Attributable of Parent Company
	Feature 74	Proportion of Consolidated Income Attributable of Minority Shareholders
	Feature 75	Proportion of Owner's Equity to Investment Capital of Parent Company
Accounting Indicators (business capability)	Feature 76	Accounts Receivable to Revenue Ratio
	Feature 77	Accounts Receivable Turning Ratio A
	Feature 78	Accounts Receivable Turning Ratio B
	Feature 79	Accounts Receivable Turning Ratio C
	Feature 80	Accounts Receivable Turning Ratio D

Feature 81 Accounts Receivable Turning Ratio TTM

Feature 82 Accounts Receivable Turning Days A

Feature 83 Accounts Receivable Turning Days B

Feature 84 Accounts Receivable Turning Days C

Feature 85 Accounts Receivable Turning Days TTM

Feature 86 Inventory to Revenue Ratio

Feature 87 Inventory Turning Ratio A

Feature 88 Inventory Turning Ratio B

Feature 89 Inventory Turning Ratio C

Feature 90 Inventory Turning Ratio D

Feature 91 Inventory Turning Ratio TTM

Feature 92 Inventory Turning Days A

Feature 93 Inventory Turning Days B

Feature 94 Inventory Turning Days C

Feature 95 Inventory Turning Days TTM

Feature 96 Operation Cycle A

Feature 97 Operation Cycle B

Feature 98 Operation Cycle C

Feature 99 Operation Cycle TTM

Feature 100 Accrued Payable Rate A

Feature 101 Accrued Payable Rate B

Feature 102 Accrued Payable Rate C

Feature 103 Accrued Payable Rate D (%)

Feature 104 Accrued Payable Rate TTM (%)

Feature 105 Working Capital Turnover Rate A

Feature 106 Working Capital Turnover Rate B

Feature 107 Working Capital Turnover Rate C

Feature 108 Working Capital Turnover Rate D

Feature 109 Working Capital Turnover Rate TTM

Feature 110 Cash Equivalents Turnover Ratio A

Feature 111 Cash Equivalents Turnover Ratio B

Feature 112 Cash Equivalents Turnover Ratio C

Feature 113 Cash Equivalents Turnover Ratio D

Feature 114 Cash Equivalents Turnover Ratio TTM

Feature 115 Current Assets to Revenue Ratio

Feature 116 Current Assets Turnover A

Feature 117 Current Assets Turnover B

Feature 118 Current Assets Turnover C

Feature 119 Current Assets Turnover D

Feature 120 Current Assets Turnover TTM

Feature 121 Fixed Assets to Revenue Ratio

Feature 122 Fixed Asset Rate A

	Feature 123 Fixed Asset Rate B
	Feature 124 Fixed Asset Rate C
	Feature 125 Fixed Asset Rate D
	Feature 126 Fixed Asset Rate TTM
	Feature 127 Non-Current Assets Turnover A
	Feature 128 Non-Current Assets Turnover B
	Feature 129 Non-Current Assets Turnover C
	Feature 130 Non-Current Assets Turnover D
	Feature 131 Non-Current Assets Turnover TTM
	Feature 132 Capital Intensity
	Feature 133 Total Asset Rate A
	Feature 134 Total Asset Rate B
	Feature 135 Total Asset Rate C
	Feature 136 Total Asset Rate D
	Feature 137 Total Asset Rate TTM
	Feature 138 Equity Rate A
	Feature 139 Equity Rate B
	Feature 140 Equity Rate C
	Feature 141 Equity Rate D
	Feature 142 Equity Rate TTM
	Feature 143 Assets Profit Ratio A
	Feature 144 Assets Profit Ratio B
	Feature 145 Assets Profit Ratio C
	Feature 146 Assets Profit Ratio TTM
	Feature 147 Return on Assets (ROA) A
	Feature 148 Return on Assets (ROA) B
	Feature 149 Return on Assets (ROA) C
	Feature 150 Return on Assets (ROA) TTM
	Feature 151 Current Assets Net Margin A
	Feature 152 Current Assets Net Margin B
Accounting	Feature 153 Current Assets Net Margin C
Indicators	Feature 154 Current Assets Net Margin TTM
(profitability)	Feature 155 Fixed Assets Net Margin A
	Feature 156 Fixed Assets Net Margin B
	Feature 157 Fixed Assets Net Margin C
	Feature 158 Fixed Assets Net Margin TTM
	Feature 159 Return on Equity (ROE) A
	Feature 160 Return on Equity (ROE) B
	Feature 161 Return on Equity (ROE) C
	Feature 162 Return on Equity (ROE) TTM
	Feature 163 Earnings before Interest and Taxes (EBIT)
	Feature 164 EBIT TTM

Feature 165 Earnings before Interest and After Taxes

Feature 166 Earnings before Interest, Taxes, Depreciation, and Amortization (EBITDA)

Feature 167 EBITDA TTM

Feature 168 Net Profit to Gross Profit

Feature 169 Gross Profit to EBIT

Feature 170 EBIT to Total Assets

Feature 171 Return on Invested Capital (ROIC)

Feature 172 Long Rate of Return on Capital

Feature 173 Gross Profit Ratio

Feature 174 Gross Profit Ratio TTM

Feature 175 Operating Cost Ratio

Feature 176 Operating Cost Ratio TTM

Feature 177 Operating Profit Margin

Feature 178 Operating Profit Margin TTM

Feature 179 Net Profit Ratio

Feature 180 Net Profit Ratio TTM

Feature 181 Total Operating Cost Ratio

Feature 182 Total Operating Cost Ratio TTM

Feature 183 Sales Expense Ratio

Feature 184 Sales Expense Ratio TTM

Feature 185 Administrative Expense Ratio

Feature 186 Administrative Expense Ratio TTM

Feature 187 Financial Expense Ratio

Feature 188 Financial Expense Ratio TTM

Feature 189 Research and Development Expense Ratio

Feature 190 Selling Expense Ratio

Feature 191 Selling Expense Ratio TTM

Feature 192 Total Profit Cost Ratio

Feature 193 Total Profit Cost Ratio TTM

Feature 194 Asset Impairment Loss to Operating Revenue Ratio

Feature 195 Asset Impairment Loss to Operating Revenue Ratio TTM

Feature 196 EBITDA Margin

Feature 197 EBITDA Margin TTM

Feature 198 EBIT Margin

Feature 199 EBIT Margin TTM

Feature 200 Cash to Profit Ratio

Feature 201 Cash to Profit Ratio TTM

Feature 202 Return on Equity (ROE) of Parent Company A

Feature 203 Return on Equity (ROE) of Parent Company B

Feature 204 Return on Equity (ROE) of Parent Company C

Feature 205 Return on Equity (ROE) of Parent Company TTM

Feature 206 Consolidated Income of Parent Company A

	Feature 207 Comprehensive Income of Parent Company B
	Feature 208 Comprehensive Income of Parent Company C
	Feature 209 Comprehensive Income of Parent Company TTM
	Feature 210 Return on Investment (ROI)
Accounting Indicators (cash flow analysis)	Feature 211 Net Profit Cash Cover
	Feature 212 Net Profit Cash Cover TTM
	Feature 213 Operating Revenue Cash Cover
	Feature 214 Operating Revenue Cash Cover TTM
	Feature 215 Net Cash Content of Operating Revenue
	Feature 216 Net Cash Content of Operating Revenue TTM
	Feature 217 Net Cash Content of Operating Profit
	Feature 218 Net Cash Content of Operating Profit TTM
	Feature 219 Net Cash Flow from Financing Activities to Creditors
	Feature 220 Net Cash Flow from Financing Activities to Creditors TTM
	Feature 221 Net Cash Flow from Financing Activities to Shareholders
	Feature 222 Net Cash Flow from Financing Activities to Shareholders TTM
	Feature 223 Depreciation and Amortization
	Feature 224 Depreciation and Amortization TTM
	Feature 225 Corporate Cash Flow 1
	Feature 226 Corporate Cash Flow 2
	Feature 227 Corporate Cash Flow TTM 1
	Feature 228 Corporate Cash Flow TTM 2
	Feature 229 Equity Cash Flow 1
	Feature 230 Equity Cash Flow 2
	Feature 231 Equity Cash Flow TTM 1
	Feature 232 Equity Cash Flow TTM 2
	Feature 233 Free Cash Flow of Firm (original)
	Feature 234 Free Cash Flow of Equity (original)
	Feature 235 Overall Cash Recovery Rate
	Feature 236 Working Index
	Feature 237 Capital Expenditure to Depreciation and Amortization Ratio
	Feature 238 Cash Adequacy Ratio
	Feature 239 Cash to Reinvest
	Feature 240 Cash to Meet Invest Needs
	Feature 241 Equity Free Cash Flow
	Feature 242 Firm Free Cash Flow
Accounting Indicators (risk level)	Feature 243 Financial Leverage
	Feature 244 Operating Leverage
	Feature 245 Comprehensive Leverage
Accounting Indicators	Feature 246 Equity Grow Ratio A
	Feature 247 Equity Grow Ratio B
	Feature 248 Capital Preservation and Appreciation Rate of Parent Company

(development capability)	Feature 249 Capital Accumulation Rate A
	Feature 250 Capital Accumulation Rate B
	Feature 251 Capital Accumulation Rate of Parent Company
	Feature 252 Fixed Asset Grow Rate A
	Feature 253 Fixed Asset Grow Rate B
	Feature 254 Total Assets Grow Rate A
	Feature 255 Total Assets Grow Rate B
	Feature 256 Grow Rate of ROE A
	Feature 257 Grow Rate of ROE B
	Feature 258 Grow Rate of Basic Earnings Per Share A
	Feature 259 Grow Rate of Basic Earnings Per Share B
	Feature 260 Grow Rate of Fully Diluted Earnings Per Share A
	Feature 261 Grow Rate of Fully Diluted Earnings Per Share B
	Feature 262 Net Profit Grow Rate A
	Feature 263 Net Profit Grow Rate B
	Feature 264 Total Profit Grow Rate A
	Feature 265 Total Profit Grow Rate B
	Feature 266 Operating Profit Grow Rate A
	Feature 267 Operating Profit Grow Rate B
	Feature 268 Net Profit Grow Rate of Parent Company
	Feature 269 Consolidated Income Grow Rate
	Feature 270 Consolidated Income Grow Rate of Parent Company
	Feature 271 Operating Revenue Grow Rate A
	Feature 272 Operating Revenue Grow Rate B
	Feature 273 Operating Revenue Grow Rate
	Feature 274 Operating Cost Grow Rate
	Feature 275 Sales Expense Grow Rate
	Feature 276 Administrative Expense Grow Rate
	Feature 277 Accruals
	Feature 278 Operating Cash Per Share Growth Rate A
	Feature 279 Operating Cash Per Share Growth Rate B
	Feature 280 Operation Cash Grow Rate A
	Feature 281 Operation Cash Grow Rate B
	Feature 282 Invest Cash Grow Rate A
	Feature 283 Invest Cash Grow Rate B
	Feature 284 Financing Cash Grow Rate A
	Feature 285 Financing Cash Grow Rate B
	Feature 286 Sustainable Grow Rate
	Feature 287 Grow Rate of Owner's Equity A
	Feature 288 Grow Rate of Owner's Equity B
	Feature 289 Grow Rate of Net Asset Per Share B
	Feature 290 Grow Rate of Net Asset Per Share A

	Feature 291 Grow Rate of Owner's Equity C
	Feature 292 Grow Rate of Net Asset Per Share C
	Feature 293 Growth of ROE C
	Feature 294 Net Profit Grow Rate C
	Feature 295 Operating Revenue Grow Rate C
	Feature 296 Total Assets Grow Rate C
	Feature 297 Total Profit Grow Rate C
	Feature 298 Fixed Asset Grow Rate C
	Feature 299 Grow Rate of Basic Earnings Per Share C
	Feature 300 Operating Profit Grow Rate C
	Feature 301 Sustainable Grow Rate 2
	Feature 302 Sustainable Grow Rate 3
Accounting Indicators (index per share)	Feature 303 Earnings Per Share 1
	Feature 304 Earnings Per Share TTM1
	Feature 305 Earnings Per Share 2
	Feature 306 Earnings Per Share TTM2
	Feature 307 Earnings Per Share 3
	Feature 308 Earnings Per Share TTM3
	Feature 309 Earnings Per Share 4
	Feature 310 Earnings Per Share TTM4
	Feature 311 Consolidated Income Per Share 1
	Feature 312 Consolidated Income Per Share TTM1
	Feature 313 Consolidated Income Per Share 2
	Feature 314 Consolidated Income Per Share TTM2
	Feature 315 Earnings Per Share of Parent Company 1
	Feature 316 Earnings Per Share of Parent Company TTM1
	Feature 317 Consolidated Income Per Share of Parent Company 1
	Feature 318 Consolidated Income Per Share of Parent Company TTM1
	Feature 319 Operating Revenue Per Share 1
	Feature 320 Operating Revenue Per Share TTM1
	Feature 321 Prime Operating Revenue Per Share 1
	Feature 322 Prime Operating Revenue Per Share TTM1
	Feature 323 Earnings Per Share before Interest and Taxes 1
	Feature 324 Earnings Per Share before Interest and Taxes TTM1
	Feature 325 Earnings Per Share before Interest, Taxes, Depreciation and Amortization 1
	Feature 326 Earnings Per Share before Interest, Taxes, Depreciation and Amortization TTM1
	Feature 327 Operation Profit Per Share 1
	Feature 328 Operation Profit Per Share TTM1
	Feature 329 Net Asset Per Share 1
	Feature 330 Tangible Asset Per Share 1
	Feature 331 Liability Per Share 1

Feature 332 Capital Reserve Per Share 1

Feature 333 Surplus Reserve Per Share 1

Feature 334 Undivided Profit Per Share 1

Feature 335 Retained Earnings Per Share 1

Feature 336 Net Asset Per Share of Parent Company 1

Feature 337 Per Share Cash Flow from Operations 1

Feature 338 Per Share Cash Flow from Operations TTM1

Feature 339 Per Share Cash Flow from Investing Activities 1

Feature 340 Per Share Cash Flow from Investing Activities TTM1

Feature 341 Per Share Cash Flow from Financing Activities 1

Feature 342 Per Share Cash Flow from Financing Activities TTM1

Feature 343 Per Share Free Cash Flow of Firm 1

Feature 344 Per Share Free Cash Flow of Firm TTM1

Feature 345 Per Share Free Cash Flow of Equity 1

Feature 346 Per Share Free Cash Flow of Equity TTM1

Feature 347 Depreciation and Amortization Per Share 1

Feature 348 Depreciation and Amortization Per Share TTM1

Feature 349 Free Cash Flow of Firm Per Share (original)

Feature 350 Free Cash Flow of Equity Per Share (original)

Feature 351 Cash Flow Per Share 1

Feature 352 Cash Flow Per Share TTM1

Feature 353 Cash Flow Per Share 2

Feature 354 Cash Flow Per Share TTM2

Feature 355 Earnings Per Share of Parent Company 2

Feature 356 Earnings Per Share of Parent Company TTM2

Feature 357 Consolidated Income Per Share of Parent Company 2

Feature 358 Consolidated Income Per Share of Parent Company TTM2

Feature 359 Operating Revenue Per Share 2

Feature 360 Operating Revenue Per Share TTM2

Feature 361 Prime Operating Revenue Per Share 2

Feature 362 Prime Operating Revenue Per Share TTM2

Feature 363 Earnings Per Share before Interest and Taxes 2

Feature 364 Earnings Per Share before Interest and Taxes TTM2

Feature 365 Earnings Per Share before Interest, Taxes, Depreciation and Amortization 2

Feature 366 Earnings Per Share before Interest, Taxes, Depreciation and Amortization TTM2

Feature 367 Operation Profit Per Share 2

Feature 368 Operation Profit Per Share TTM2

Feature 369 Net Asset Per Share 2

Feature 370 Tangible Asset Per Share 2

Feature 371 Liability Per Share 2

Feature 372 Capital Reserve Per Share 2

	Feature 373 Surplus Reserve Per Share 2
	Feature 374 Undivided Profit Per Share 2
	Feature 375 Retained Earnings Per Share 2
	Feature 376 Net Asset Per Share of Parent Company 2
	Feature 377 Per Share Cash Flow from Operations 2
	Feature 378 Per Share Cash Flow from Operations TTM2
	Feature 379 Per Share Cash Flow from Investing Activities 2
	Feature 380 Per Share Cash Flow from Investing Activities TTM2
	Feature 381 Per Share Cash Flow from Financing Activities 2
	Feature 382 Per Share Cash Flow from Financing Activities TTM2
	Feature 383 Per Share Free Cash Flow of Firm 2
	Feature 384 Per Share Free Cash Flow of Firm TTM2
	Feature 385 Per Share Free Cash Flow of Equity 2
	Feature 386 Per Share Free Cash Flow of Equity TTM2
	Feature 387 Depreciation and Amortization Per Share 2
	Feature 388 Depreciation and Amortization Per Share TTM2
	Feature 389 Per Share Free Cash Flow of Firm 3
	Feature 390 Per Share Free Cash Flow of Equity 3
	Feature 391 Price Earning Ratio 1
	Feature 392 Price Earning Ratio 2
	Feature 393 Price Earning Ratio TTM
	Feature 394 Price to Sales Ratio 1
	Feature 395 Price to Sales Ratio 2
	Feature 396 Price to Sales Ratio TTM
	Feature 397 Price Cash Flow Ratio 1
	Feature 398 Price Cash Flow Ratio 2
	Feature 399 Price Cash Flow Ratio TTM
	Feature 400 Price to Book ratio
	Feature 401 Market-to-Tangible-Book Ratio
Accounting Indicators (relative value indicators)	Feature 402 Price Earning Ratio of Parent Company 1
	Feature 403 Price Earning Ratio of Parent Company 2
	Feature 404 Price Earning Ratio of Parent Company TTM
	Feature 405 Price to Book ratio of Parent Company
	Feature 406 Market Value A
	Feature 407 Market Value B
	Feature 408 Tobin's Q Ratio A
	Feature 409 Tobin's Q Ratio B
	Feature 410 Tobin's Q Ratio C
	Feature 411 Tobin's Q Ratio D
	Feature 412 Book-to-Market Ratio A
	Feature 413 Book-to-Market Ratio B
	Feature 414 Ratio of Dividend

	Feature 415 Earnings Per Share (EPS) for Common Stock A
	Feature 416 EPS for Common Stock B
	Feature 417 Enterprise Value Multiple
	Feature 418 Enterprise Value Multiple TTM (Trailing Twelve Months)
Accounting Indicators (dividend distribution)	Feature 419 Cash Dividend Per Share before Tax
	Feature 420 Cash Dividend Per Share after Tax
	Feature 421 Dividend Distribution Ratio
	Feature 422 Change in Dividend Per Share
	Feature 423 Percentage Change in Dividend Per Share
	Feature 424 Cash Dividend Cover
	Feature 425 Dividend Multiplier
	Feature 426 Retained Earnings Ratio
	Feature 427 Dividend Distribution Ratio 2
	Feature 428 Dividend Distribution Ratio 3

719