Personal Characteristics of Female Executives and Corporate Solvency

Based on CatBoost Analysis

Chenya Huang*

The University of Sydney Business School, The University of Sydney, Sydney, NSW, Australia

Chua7955@uni.sydney.edu.au

Abstract. Executive characteristics were once a hot topic in corporate management research. However, related studies lacked more detailed analysis, for example, using gender as a control variable to study the impact of female executive characteristics in corporate management. This paper uses the CatBoost algorithm for the first time to explore the predictability of personal characteristics of female executives on corporate solvency, using Chinese listed firms from 2016 to 2021 as a research sample. It was found that (1) among all the variables studied, the personal characteristics of female corporate executives did not highlight a strong predictive power on solvency; (2) among the personal characteristics of female executives, the relative importance of the presence of human resources, management and marketing backgrounds and the presence of both chairman and CEO were higher and had a relatively high degree of prediction on corporate solvency.

Keywords: component; Corporate Governance; Solvency; CatBoost; Factor Analysis

1 Introduction

After the introduction of the Upper Echelon Theory in 1984, the association between top characteristics and firm performance has gradually become one of the research issues of interest to scholars. With the rising status of women in the new era, the association between female executives and firm performance has also become a key area of interest for scholars. Studies have shown that female executives are positively related to the financial performance of companies to a certain extent. However, some scholars believe that women have a significant negative impact on the financial performance of companies due to their energy and ability limitations. Overall, these views are still in a conflicting stage.

H. Mallick et al. (Eds.): ICEMCI 2022, AEBMR 231, pp. 864–873, 2023.
https://doi.org/10.2991/978-94-6463-098-5_99

Company management is often expressed through financial performance, which generally contains profitability, operating capacity and solvency. Solvency refers to a company's ability to use asset realization to repay its debts before they mature. According to the length of time limit, it can be divided into short-term solvency and long-term solvency. The correct analysis, evaluation and forecast of the solvency of an enterprise is helpful for the company management to optimize the company's financing structure and reduce the cost of debt financing.

Currently, most of the analysis of solvency is focused on different solvency indicators. However, there is a certain correlation and overlap between the information they contain. The existence of comprehensive solvency indicators can help reduce the cost of analysis and provide a more intuitive picture of a company's solvency situation. Few studies have been conducted to determine whether the personal characteristics of female executives are predictive of corporate solvency.

To this end, this paper explores the predictive power between individual characteristics of female executives and corporate solvency, using CatBoost model in machine learning by constructing a comprehensive solvency scoring mechanism with female executives in the management team as the main entry point for the research sample of listed companies from 2016 to 2021.

2 Related Works

2.1 The Impact of Female Executives on Financial Performance

The impact of women on business is based on the Role Congruity Theory, which points to the conflicting relationship between women's social and leadership roles. Since its introduction in 2002, the Role Congruity Theory has been widely influential in research related to women leaders and women's career development. Eagly and Karau (2002) suggested that the phenomenon of gender double bind occurs when the relevant temperamental traits of women conflict with the relevant traits required for leadership due to society's stereotypical expectations of different gender roles [3]. The phenomenon puts women in the bondage of meeting both the requirements of leadership and being female. This exposes women to specific dilemmas and will suffer more negative perceptions than men. A report by Catalyst (2007) found that when women lead with stereotypical characteristics, such as being gentle and emotional, they are perceived as incompetent and overly emotional [2]. When women exhibit assertive leadership styles, they are perceived as unfeminine. In fact, in Boulouta's study (2013), female directors were found to be more risk-averse and had lower levels of overconfidence [1]. On the Social Role Theory, it suggests that female roles are more communal than male roles. The decision-making style of female directors reflects a clear concern for publicness and corporate external relations (Nielsen and Huse, 2010) [5]. In terms of financial performance, Ren et al. (2011) found empirically that a higher proportion of female executives promotes corporate performance, and verified the positive moderating effect of female executive human capital on corporate performance [6].

2.2 CatBoost model

The presence of a large amount of category-based data in the executive personal characteristic variables leads to the fact that using a regression model in a general sense is not the best choice. The CatBoost model can cope with this problem very well that automatically adopts a special way to handle categorical features. Firstly, it does some statistics on categorical features, calculates the frequency of a category, and then adds hyperparameters to generate new numerical features. It also uses combined category features, which can be exploited to linkage between features, which greatly enriches the feature dimension. The base model of CatBoost uses a symmetric tree, and the way to calculate the leaf-value is different from the traditional boosting algorithm. The traditional boosting algorithm calculates the average, while CatBoost optimizes this aspect by using other algorithms, and these improvements prevent the model from overfitting.

In addition, CatBoost solves the problems of gradient bias and prediction shift, reducing the occurrence of overfitting and improving the accuracy and generalization ability of the algorithm. It replaces the gradient estimation method in the traditional algorithm by using ordered boosting to reduce the bias of gradient estimation and improve the generalization ability of the model. The flow of the ordered boosting algorithm is shown in figure 1.

Algorithm 1: Ordered boosting

input : $\{(\mathbf{x}_k, y_k)\}_{k=1}^{n}, I$;

$\sigma \leftarrow$ random permutation of $[1, n]$;
$M_i \leftarrow 0$ for $i = 1..n$;
for $t \leftarrow 1$ **to** I **do**
  for $i \leftarrow 1$ **to** n **do**
    $r_i \leftarrow y_i - M_{\sigma(i)-1}(\mathbf{x}_i)$;
  for $i \leftarrow 1$ **to** n **do**
    $\Delta M \leftarrow$ $LearnModel((\mathbf{x}_j, r_j) : \sigma(j) \leq i)$;
    $M_i \leftarrow M_i + \Delta M$;
return M_n

Fig. 1. The flow of the ordered boosting algorithm. Source: Gao et al. (2019)

From figure 1, in order to obtain the unbiased gradient estimate, CatBoost trains a separate model Mi for each sample xi, which is obtained by training with a training set that does not contain sample xi. The researchers use Mi to obtain gradient estimates on the samples, and use the gradient to train the base learner and obtain the final model.

3 Experiment

3.1 Experimental analysis

The research for solvency forecasting includes 5 steps, which are data collection, data pre-processing, feature processing, model forecasting, and result analysis. This is shown in the figure 2 below.

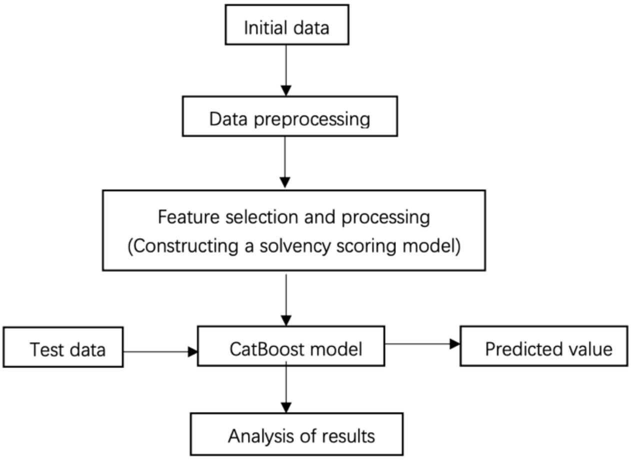

Fig. 2. Steps of Example of the research for solvency forecasting. Source: Self-made

3.2 Data sources and pre-processing

The sample selected in this paper is all Chinese A-share listed companies from 2016 to 2021, excluding ST stocks. The research subjects are executives of listed companies as well as solvency indicators. The data are obtained from Guotaian Economic and Financial Research Database, and the missing data are excluded. The final complete data sample of 15927 is retained.

3.3 Feature selection and processing

Due to the large number of debt service indicators and the correlation or overlap of factors among the indicators, this paper measures whether an enterprise can repay its debts from the comprehensive solvency score. The comprehensive solvency score is derived from a comprehensive scoring model constructed using factor analysis, in which the variables include current ratio, cash ratio, quick ratio, equity multiplier and other long- and short-term solvency indicators.

3.4 Constructing a comprehensive solvency scoring model

In order to present the comprehensive solvency of enterprises more intuitively, it is necessary to reduce the dimensionality of indicators. Factor analysis method can identify representative factors among many variables and reduce the number of variables.

3.5 Indicator test

After performing KMO and Bartlett tests on the data, the results showed that the value of KMO was 0.757, which was greater than the standard value of 0.5; the significance of Bartlett test was 0.000, which was less than the significance level of 0.05, rejecting the null hypothesis of Bartlett's sphericity test and considered suitable for factor analysis.

Table 1. KMO & Bartlett test

KMO measure of sampling adequacy		0.75
Bartletts test of sphericity	Chi-squared approximation	688225.93
	Degrees of Freedom	7
	Significance	0.00

Source: Self-made

3.6 Total variance explanation

When selecting the common factors, the selection criteria are eigenvalues greater than or equal to 1 and a cumulative contribution of 80%. As can be seen from the figure below, the first five principal components meet the criteria and explain 80.492% of the variance, reflecting 80.492% of the information of the variables in the sample. This indicates that the 13 solvency indicators can be simplified to 5 principal components without losing most of the information.

Table 2. Total variance explanation (1)

	Initial Eigenvalues			Sum of squares of extracted loads		
Component	Total	Percentage of variance	Cumulative%	Total	Percentage of variance	Cumulative%
1	4.029	30.991	30.991	4.029	30.991	30.991
2	2.899	22.303	53.294	2.899	22.303	53.294
3	1.445	11.112	64.405	1.445	11.112	64.405
4	1.091	8.394	72.800	1.091	8.394	72.800
5	1.000	7.692	80.492	1.000	7.692	80.492
6	.983	7.563	88.055			
7	.853	6.560	94.615			
8	.293	2.257	96.872			
9	.155	1.192	98.064			
10	.144	1.111	99.175			
11	.092	.710	99.886			

12	.015	.114	100.000			
13	7.323E-9	5.633E-8	100.000			

Source: Self-made

Table 3. Total variance explanation (2)

Component	Sum of squared rotational loads		
	Total	Percentage of variance	Cumulative%
	3.527	27.129	27.129
1	2.904	22.338	49.466
2	1.838	14.142	63.609
3	1.195	9.191	72.800
4	1.000	7.692	80.492

Source: Self-made

3.7 Factor loading

Table 4. Rotated component matrix

	Component				
	1	2	3	4	5
X1	.965				
X2	.964				
X9	.892				
X3	.860				
X8		.992			
X7		.992			
X11		.966			
X4			.946		
X13			.925		
X10				.773	
X6				.688	
X5					
X12					.998

Source: Self-made

3.8 Factor score

Table 5. Component matrix

	Component				
	1	2	3	4	5
X2	.928				
X1	.925				

X9	.900				
X3	.853				
X8		.988			
X7		.988			
X11		.963			
X4	.535		.799		
X13	.596		.753		
X10				.774	
X6				.644	
X5					
X12					.998

Source: Self-made

Assume that the linear combination model of the public factor F represented by the variable x is:

$$F_i = \beta_{i1}x_1 + \beta_{i2}x_2 + \cdots + \beta_{ij}x_j$$

The formula is called the factor score model, and the factor coefficients in the component matrix are substituted into the model to obtain the factor score function.

The factor calculation formula is:

$$F_1 = 0.928x_1 + 0.925x_2 + 0.900x_3 + 0.853x_4 + 0.535x_8 + 0.596x_9$$
$$F_2 = 0.998x_5 + 0.988x_6 + 0.963x_7$$
$$F_3 = 0.799x_8 + 0.753x_9$$
$$F_4 = 0.774x_{10} + 0.644x_{11}$$
$$F_5 = 0.998x_{13}$$

Based on the factor scores and variance contribution, the final combined factor model is:

$$F = 0.385F_1 + 0.277F_2 + 0.138F_3 + 0.104F_4 + 0.096F_5$$

The characteristics of female executives selected include age, education, voice, and relevant background. Among them, voice refers to whether the executives are both CEO and chairman. Relevant backgrounds are divided into six main areas, which are production, R&D and design background; human resources, management and marketing background; finance, finance and legal background; overseas background; academic background and financial background.

3.9 Model Predictions

The processed data and features are fed into the model for prediction, and the dataset will be divided into training, validation and test sets in the ratio of 7:2:1. The prediction effectiveness of the model is measured by R2 with RMSE and compared with Random Forest and XGBoost. The comparison results are shown in the following table:

Table 6. comparison of three models

Model	R^2	RMSE
CatBoost	0.9998	3.4435
Random Forest	0.4080	14.3046
XGBoost	0.4762	7.3332

Source: Self-made

As can be seen from the table 6, CatBoost is significantly better than the other two models in terms of both model reliability and accuracy.

3.10 Analysis of results

Table 7. The relative importance of the characteristics of female executives

	Feature	Feature Importance
1	LnAsset	63.6750
2	Ass-lia	11.0257
7	Funcback2	2.6560
5	IsDuality	0.6730
8	Funcback3	0.1868
10	Academic	0.1501
6	Funcback1	0.1453
3	LnAge	0.1146
9	OverseaBack	0.0927
4	IsUnder	0.0115

Source: Self-made

The table 7 demonstrates the relative importance of the variables for predicting the solvency of the firm. Among all variables, the relative importance of firm size and gearing ratio is higher. Among the executive characteristics, the relative importance of having a background in human resources, management and marketing (FunBack2) and being both chairman and CEO (IsDuality) is high. Other variables such as whether they have overseas background and age have lower relative importance. funBack2 reflects executives' ability to manage the company internally and externally and their sensitivity to market changes. IsDuality, which ranks second in the importance of personal characteristics, reflects the extent of the executive's voice, from which it can be determined whether female executives are valued in corporate management.

4 Conclusion

This paper selects and processes the relevant research variables for the sample of listed companies from 2016 to 2021, and for the first time uses the CatBoost algorithm to investigate whether the personal characteristics of female executives have a certain predictive power on corporate solvency, and achieves the relative research results. After the study, it was found that (1) among all the variables studied, the personal characteristics of female executives of the company did not highlight a strong predictive power on solvency; (2) among the personal characteristics of female executives, the relative importance of whether they have human resources, management and marketing backgrounds and whether they are also the chairman and CEO was higher and had a relatively high degree of prediction on the solvency of the company.

In comparison with other machine learning models, CatBoost performs well in all metrics, indicating that CatBoost is indeed more suitable for data containing a large number of category features and has good application in predicting company financial performance and business management. The limitation of this paper is that although CatBoost has prevented the model from overfitting to some extent, it may still be overfitting based on the final results. This may be related to the overlap between some of the variables and the factors in the composite debt service score. In addition, whether the relatively important executive characteristics have a positive or negative effect on the prediction of corporate solvency needs further study.

References

1. Boulouta. (2013). Hidden Connections: The Link Between Board Gender Diversity and Corporate Social Performance. Journal of Business Ethics, 113(2), 185–197. https://doi.org/10.1007/s10551-012-1293-7
2. Catalyst. (2007). The Double-Bind Dilemma for Women in Leadership: Damned if You Do, Doomed if You Don't. IBM Corporation.
3. Eagly, & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. Psychological Review, 109(3), 573–598. https://doi.org/10.1037//0033-295X.109.3.573
4. Gao, K., Chen, H., Zhang, X., Ren, X., Chen, J., & Chen, X. (2019). A novel material removal prediction method based on acoustic sensing and ensemble XGBoost learning algorithm for robotic belt grinding of Inconel 718. International Journal of Advanced Manufacturing Technology, 105(1-4), 217–232. https://doi.org/10.1007/s00170-019-04170-7
5. Nielsen, & Huse, M. (2010). The Contribution of Women on Boards of Directors: Going beyond the Surface. Corporate Governance: an International Review, 18(2), 136–148. https://doi.org/10.1111/j.1467-8683.2010.00784.x
6. Ren, T., & Wang, Z. (2011). Female participation in TMT and firm performance: evidence from Chinese private enterprises. Nankai Business Review International, 2(2), 140–157. https://doi.org/10.1108/20408741111139918