



# Corporate governance and financial distress: lessons learned from an unconventional approach

Alberto Tron<sup>1</sup> · Maurizio Dallocchio<sup>1</sup> · Salvatore Ferri<sup>2</sup> · Federico Colantoni<sup>3</sup>

Accepted: 8 June 2022  
© The Author(s) 2022

## Abstract

Using a and a unique set of Italian non-listed Unlikely to Pay (UTP) positions, that consist in the phase that precedes the insolvency but where it is still possible for the company to succeed in restructuring, this paper aims to analyze the relationships between corporate governance characteristics and financial distress status. We compare the performance of corporate governance variables in predicting corporate defaults, using both the Logit and Random Forest models, which previous researchers have deemed to be the most efficient machine learning techniques. Our results show that the use of corporate governance variables – especially with regards to CEO renewal and stability in the composition of the board of directors – increases the accuracy of the Random Forest technique and influences the success of the turnaround process. This paper also confirms the Random Forest technique’s ability to significantly outperform the Logit model in terms of accuracy.

**Keywords** Financial distress · Corporate finance · Corporate governance · Corporate default

JEL G32 · G33 · G34 · K22

- 
- ✉ Alberto Tron  
alberto.tron@unibocconi.it
  - ✉ Maurizio Dallocchio  
maurizio.dallocchio@sdbocconi.it
  - ✉ Salvatore Ferri  
salvatore.ferri@uniparthenope.it
  - ✉ Federico Colantoni  
federico.colantoni@student.unisg.ch

<sup>1</sup> Università Commerciale L. Bocconi, Via Roentgen, 1, 20136 Milan, Italy

<sup>2</sup> Università degli Studi Parthenope, Via A. F. Acton, 38, 80133 Naples, Italy

<sup>3</sup> University of St. Gallen, 50, 9000 Dufourstrasse, St. Gallen, Switzerland

## 1 Introduction

Nowadays, models that can predict the bankruptcy of a company are of interest to various economic entities, such as banks, credit agencies, governments, and financial analysts, not to mention customers and suppliers. Although bankruptcy detection models have been gradually developing since the 1960s (Beaver, 1966; Altman, 1968), the vast majority of them are still based only on accounting and financial variables as explanatory factors (Fijorek & Grotowski, 2012; Attaran et al., 2012). Ernst and Young (2010) suggested that most managers believe that innovation (specifically in management processes) can play a fundamental role in accelerating a company's growth. Therefore, corporate governance variables – such as manager turnovers, which introduce new entrepreneurship skills, knowledge, and external relationships (Paoloni et al., 2020), and the presence of internal or external auditors (Cenciarelli et al., 2018) – can be a crucial factor in increasing the competitive advantage of a corporation (Chen et al., 2014; Agostini & Nosella, 2017). Indeed, from the 1990s onwards, various authors (e.g., Gilson 1990; Daily & Dalton, 1994; Gales & Kesner, 1994; Deng & Wang, 2006; Donker et al., 2009; Bredart, 2016; Manzanque et al., 2016) have begun to investigate the link between financial distress and corporate governance, confirming the significant relationship between corporate governance variables and the predictive power of bankruptcy forecast models. However, despite the increase in the amount of research on corporate governance variables, which covers multiple disciplines, scholarly literature on the topic is still limited and fragmented (Martín-de Castro et al., 2019), especially with regards to the role and impact of corporate governance on companies' turnarounds. Research tends to instead remain focused on the analysis of the impact of financial variables in predicting corporate defaults. Therefore, the aim of this paper is to test the effectiveness of internal corporate governance mechanisms pertaining to the board of directors in anticipating corporate defaults in Italy – in which research on the overall analysis of this issue is still lacking. The study of the inclusion of corporate governance variables is particularly suitable to the Italian context, as this country is typically characterized by non-listed family firms, which tend to operate at higher levels of risk than non-family firms (Contreras et al., 2021) and are normally unable to alter their management (or the original entrepreneur).

Moreover, despite the increasing use of corporate default forecasts in various fields, new, more efficient, and more reliable models for predicting a crisis have not yet been developed. Stemming from Altman (1968), most existing literature has relied on the use of “traditional” statistical models, such as Logit, probit, and linear discriminant analysis. Recently, however, researchers have been increasingly interested in analysing machine learning models' ability to predict bankruptcy situations. Previous literature, such as Barboza et al., (2017), has shown that “new age” models are able to offer more accurate forecasts in comparison to “traditional” methods, like Logit. Therefore, the second contribution of this paper is to reinforce scholarly literature in this field. In doing so, this paper will confirm the abilities of the models that previous literature (Jones et al., 2017; Barboza et al., 2017) has shown to be the most efficient with regards to machine learning techniques, and those that are able offer better predictions than Logit – specifically with regards to the Random Forest

model. The importance of further testing these models' efficiency in predicting the corporate defaults of machine learning techniques is amplified by the importance of their various practical applications, especially for users such as credit agencies, banks, and investment companies. In fact, as suggested by Althey (2018), we can expect these techniques to have a significant impact on the field of economics and finance within a short period of time. The practical applications of these techniques vary. For example, these techniques could be used when solving policy and decision problems (Kleinberg et al., 2015), when predicting loan repayments and attempting to improve credit scoring (Bjorkegren & Grissen, 2019), or when enhancing economic models (Althey, 2018). The increasing use of these techniques (Althey, 2018) occurs due to the fact that they can be easily applied without the need to manipulate the original database, as they are resilient to a series of statistical issues, such as omitted variables, multicollinearity, outliers, and heteroscedasticity.

To this end, we use a unique set of non-listed Italian UTP positions to compare the predictive performance of a Logit and a Random Forest model using the ROC (Receiver Operating Characteristic) curve: a method frequently analyzed in previous literature (Swets et al., 2000; Jones et al., 2015, 2017; Barboza et al., 2017). Banks classify positions that will probably not respect their contractual obligations as UTP positions. This is one of the more innovative elements of this study, as previous researchers have mainly concentrated on NPLs (Non Performing Loans, i.e., borrowings that are already long overdue and so are officially in a distressed situation). We adopt a different perspective in comparison to literature regarding the application of machine learning techniques in predicting corporate defaults. We would like to highlight, from the outset, that the ability to correctly forecast the classification as UTP has significant practical managerial consequences: UTP is a status that anticipates the company's distress and, therefore, it allows managers to find timely and realistic solutions to allow them to face a potential financial crisis. The development of models that can predict these situations and send alert signals is of particular interest to all the stakeholders of a company (Davis & Karim, 2008; Dallochio & Tron, 2020) and for banks (as shown by the guidance to banks on non-performing loans of European Central Bank<sup>1</sup>) and governments (such as in the new Italian Code of Bankruptcy<sup>2</sup>).

As predictive variables, we use the turnover of the members of the board of directors and the number of CEOs in the years prior to the crisis as corporate governance indicators, as in Elloumi & Gueyè (2001), Fahlenbrach et al. (2007), Lin et al., (2020), and Fernando et al., (2020), along with the presence of the statutory and external auditors, as in Bredart (2014) and Cenciarelli et al., (2018). The final results confirm that financial variables (Z-Score) have a primary role in predicting corporate distress situations. However, they also corroborate the importance of corporate governance factors. These conclusions are further confirmed by the fact that, when comparing the models with and without the corporate governance variables, we find

<sup>1</sup> ECB (2016), Guidance to banks on nonperforming loans. Retrieved from: [https://www.bankingsupervision.europa.eu/legalframework/publiccons/pdf/npl/npl\\_summary.en.pdf](https://www.bankingsupervision.europa.eu/legalframework/publiccons/pdf/npl/npl_summary.en.pdf).

<sup>2</sup> D.lgs. n.14, 12 January 2019. Codice della crisi d'impresa e dell'insolvenza. Retrieved from: <https://www.gazzettaufficiale.it/eli/id/2019/02/14/19G00007/sg>.

that the ROC curve is higher for those models that also include corporate governance variables, as in Liang et al., (2016).

## 2 Literature review

Literature on predicting bankruptcy is quite extensive. However, as highlighted by Wang et al., (2014), there is still no mature or definite theory of corporate failure since the “most reliable” method of predicting corporate bankruptcy has not yet been identified. After the first pieces of research on the topic, performed in the United States by Tamari (1964), Beaver (1966), and Altman (1968), several approaches for forecasting bankruptcies have been elaborated upon and improved. In the 1990s, researchers started to develop more sophisticated models by also including new strategic variables and financial/accounting measures (Altman & Saunders, 1997; Amigoni, 1998; Eccles, 1991). The ability of corporate governance factors to predict corporate defaults, as confirmed by various authors (e.g., Gilson 1990; Daily & Dalton, 1994; Gales & Kesner, 1994; Deng & Wang, 2006; Donker et al., 2009; Bredart, 2016; Manzaneque et al., 2016; Fernando et al., 2020), is explained by the postulates of Agency Theory. According to this theory (Donker et al., 2009), managers are more focused on obtaining short-term results to maximize their compensation and rewards, while shareholders tend to choose long-term strategies. Therefore, this leads to an ethical conflict between managers, like the CEO or the sole director, and the shareholders. Literature in this field is mainly based on the analysis of the US companies (Manzaneque et al., 2016). However, corporate governance mechanisms can differ significantly from one country to another, which is one of the reasons why the extension of this analysis to other geographical contexts is necessary in order to corroborate the existing literature.

Our interest is mainly focused on the board of directors, because it is considered to be one of the main internal corporate governance mechanisms (Norwahida et al., 2012) and a good measure of a company’s ability to create and leverage intellectual capital. The board of directors and the top management of a company have a crucial role in the success of a turnaround, as they have the power to implement a series of strategic actions that might prevent or solve a crisis (Porter, 1987; Garzella, 2005; Grant, 2011; Leng et al., 2021). As a consequence, top management can be considered the most important figures when it comes to renewing the structure and the strategy of a company during a corporate turnaround (Lohrke et al., 2004). As highlighted by previous literature (Schiuma et al., 2008), people with an adequate level of competencies and skills can create a virtuous circle, generating new ideas and techniques able to innovate the product/processes of a company, thereby improving performance. Moreover, as supported by Santana et al., (2017), management in declining firms can also replace the usual downsizing responses, highlighting this paper’s potential relevance in the world of start-ups and small firms (Zingales, 2000). The board has the ability to be a key factor in determining the future of a business. However, overconfident CEOs can lead a company to face higher risks of bankruptcy (Leng et al., 2011, 2021). As a consequence, weak corporate governance can increase the likelihood of opportunistic behavior occurring in the management team, which

could lead to a reduction in profitability and the overall value of a company (La Porta et al., 2000), increasing the likelihood of financial distress. Consequently, the interaction between the role of the board and the likelihood of financial distress should be examined. In this regard, previous studies (Goodstein et al., 1994; Yemareck, 1996) have been more focused on analyzing the problems related to the size of the board, revealing that larger boards tend to incentivize opportunistic behaviors by the management of a company. On the contrary, smaller boards and a larger percentage of independent individuals can increase the performance of the company and reduce the likelihood of financial distress (Jensen, 1993; Fernando et al., 2020). Several studies have also analyzed the impact of management turnover on a firm's performance, with the general consensus being that the likelihood of management turnover is negatively related to firm performance (Huson et al., 2004; Warner et al., 1988), discovered a positive relationship between low returns stocks and the probability of the turnover of the CEO, president, or board chairman. Similarly, Kim (1996) empirically demonstrates that firm stock returns have a persistent negative effect on turnover probability. On the contrary, Huson et al., (2004) discover that investors view turnover announcements as good news, presaging performance improvements. However, despite the increasing amount of research on the impact of the turnover of management and board members on the probability of financial distress (Gilson, 1990; Elloumi & Gueyiè, 2001), literature in this area is still limited and is primarily focused on data from the US.

For corporate governance systems, another key aspect for a company is the presence of internal and external auditors. In this sense, literature has shown that the presence of internal and external audit systems can have a significant impact on changes to a company's financial performance and on its probability of default (see Guo et al., 2016 and Cenciarelli et al., 2018, among others). Internal and external auditors can guarantee the quality of the information of the financial reports provided by the company for investors (Bratten et al., 2013), and their role has relevant consequences during a financial crisis (Cenciarelli et al., 2018). In this sense, also the presence of the audit committee can have a significant positive impact in preventing the risk of frauds and irregularities (Beasley et al., 2000). For distressed firms in particular, statutory auditors and external auditors are obliged to judge the ability of the company to operate as a going concern entity for the following 12 months. In this sense, with the European Union's 2015/848/EU regulation, auditors were required to promptly communicate to the top management of a company the presence of indicators of financial distress. In this field, research has shown auditors' ability to anticipate the emergence of a financial crisis (Bhimani et al., 2009). Therefore, their presence helps a company to prevent triggering this event. Research on this issue is still limited, especially in European countries (Cenciarelli et al., 2018).

Until recent times, methodological approaches were based mainly on Logit/probit models when it came to predicting corporate defaults, (Ohlson, 1980; Aziz et al., 1988; Platt & Platt, 1990; Ward, 1994; Back et al., 1996; McGurr & DeVaney, 1998; Kahya & Theodossiou, 1999; Beynon & Peel, 2001; Neophytou et al., 2004; Lin & Piesse, 2004; Westgaard & Van der Wijst, 2001; Foreman, 2003; Brockman & Turtle, 2003; among others) and the discriminant analysis model (Altman, 1968; Edmister 1972; Piesse & Wood, 1992; Altman & Narayanan, 1997; Pompe & Feelders, 1997;

McGurr & DeVaney, 1998; Yang et al., 1999; Altman et al., 2013; among others), whose merits have been studied extensively in literature (Efron, 1975; Ohlson, 1980; Altman et al., 1994; among others). However, Begley et al., (1996) supported the need for the development of new models, since traditional ones based on Altman (1968) and Ohlson's (1980) theories, have shown their weaknesses.

A new generation of corporate default predictions has arisen, including studies using hazard models, which tend to predict corporate defaults better than traditional ones (Shumway, 2001). These techniques are not binary classifiers as they calculate the probability of a corporate default over time (Chava & Jarrow, 2004; Tian et al., 2015; among others). They are therefore more suitable in the long run.

The need to improve the forecast of corporate defaults has recently led to the development of new types of models based on machine learning techniques, such as generalized boosting, AdaBoost, and Random Forest (Jones et al., 2017; Samuel, 1959) proposed the concept of machine learning, defining it as “*a discipline that gives computers the ability to learn without a clear program*”. These techniques are based on the running of a series of processes that continue to improve the classification of observations using their common patterns (Tian et al., 2012), thereby learning from experience.

One of the most famous machine learning techniques is the artificial neural network, which is very useful when it comes to solving complex and non-linear relationships “*by mimicking the structure of the brain and connecting artificial neurons using simple structures*” (Kim et al., 2020). Machine learning techniques have been used in default predictions before the 1990s, and several authors (Yang et al., 1999; Zhao et al., 2014; Geng et al., 2015; Jones et al., 2017) have demonstrated that they have both better prediction performances than Logit/probit and, on average, good levels of efficiency. They reach a good level of reliability, especially in the Italian context, wherein small companies often do not excel with regards to the accuracy of their accounting information (Falavigna, 2012; Zhao et al., 2014; Jones et al., 2017) identified five main issues with the artificial neural network:

- i) poor performance in case of unbalanced data;
- ii) high error rate with a small sample;
- iii) high difficulties in selecting the hidden layer;
- iv) less capacity to handle large numbers of potentially irrelevant inputs and to handle both categorical and continuous data;
- v) lack of interpretability of data.

Another machine learning technique is represented by the support vector machines, which are widely used in various fields, including corporate default predictions. The support vector machines are based on the concept of a separating hyperplane (Jones et al., 2017), which allows us to identify the greatest distance between the most similar observations that are oppositely classified (Cortes & Vapnik, 1995; Noble, 2006). Therefore, when a sample is completely separable in groups, the support vector machines are able to build very accurate models. However, when managing economic-finance data, this is virtually impossible. Therefore, the support vector machines contain a margin of error (Zhou et al., 2014). Support vector machines have

been notably used by researchers in corporate default predictions (Min & Lee, 2005; Yu et al., 2010), and have had a predictive quality similar to (Barboza et al., 2017; Jones et al., 2017) or even better (Shin et al., 2005) than the artificial neural network. Liang et al., (2016) also showed that support vector machines are more reliable in anticipating corporate defaults when using financial ratios and corporate governance indicators. Similar to neural networks, the limitations of support vector machines are their lack of interpretability and difficulties in handling large numbers of potentially irrelevant inputs (Jones et al., 2017).

A third technique applied in machine learning is boosting. This technique is able to find the model that best classifies a sample. The boosting technique, based on the continuous use of different sets of the initial sample, creates various training sets and identifies the one with the lowest error rate (Begley et al., 1996; Hastie et al., 2009). Unlike the Logit model, the boosting technique is resistant to over-fitting and, unlike neural networks and support vector machines, is able to deal with irrelevant inputs. It also tends to be better at handling mixed types of data (Jones et al., 2017). Due to these characteristics, this technique – together with AdaBoost, a derived algorithm – has been able to predict corporate defaults with a remarkable degree of reliability: significantly higher than that of Logit/neural networks/support vector machines (Barboza et al., 2017; Jones et al., 2017; Aliaj et al., 2020).

The Random Forest model is similar to boosting but it is based on the concept of charting decision rules using a tree structure. This technique, created by Breiman (2001), randomly selects a subset of characteristics from each node of the tree, following a bagging technique. Therefore “*a particularly strong predictor in the dataset (along with some moderately strong predictors) will be used by most if not all the trees in the top split*” (Jones et al., 2017). A more detailed and precise description of Random Forest can be found in Booth et al., (2014) and Calderoni et al., (2015). The Random Forest technique has various advantages: (i) it is robust to outliers; (ii) it is robust to missing data; and (iii) it allows for the identification of the importance of each variable in the classification results (Yeh et al., 2014; Jones et al., 2017). Several authors (Olson et al., 2012; Barboza et al., 2017; Jones et al., 2017; Li & Wang, 2018; Aliaj et al., 2020) have shown that the Random Forest model is more interpretable and can achieve more accurate results in predicting corporate defaults than other machine learning algorithms.

Despite the reliability of these techniques in predicting corporate defaults, the amount of empirical research on the topic is still limited. Tsai et al., (2014), using a set of Taiwanese companies, compared the predictability power of three neural network techniques, support vector machines, and Random Forest, demonstrating the ability of Random Forest to perform significantly better. Heo et al. (2014), using a sample of Korean companies, showed that the AdaBoost has more predictive power than other classifiers, especially for large companies. Kim et al. (2014), using publicly traded U.S. restaurants, demonstrated that, from 1988 to 2010, the AdaBoost and Random Forest models have been the best predictors of performance. They had the smallest degree of error overall and in terms of type I error rates, or rather the probability of rejecting the null hypothesis given that it is true. Both Jones et al., (2017) and Barboza et al., (2017), by comparing several statistical methods and using a set of US firms, showed that the Random Forest and the boosting techniques are

the most performing classifiers in several cases. The applications of machine learning techniques in Italy are still limited. Bragoli et al. (2019), using a dataset of Italian companies from 2007 to 2015, showed that machine learning techniques outperform traditional classifiers. Aliaj et al., (2020) using a large sample of Italian companies, demonstrated that, in the Italian context, Random Forest provides the best results, thus corroborating the findings of Barboza et al., (2017). Donato et al. (2020) showed that, using a non-parametric supervised classification algorithm on a random sample of 100 non-listed SMEs, it is possible to sufficiently predict the distress of a firm in advance (4–5 years prior to failure).

From the analysis of the existing literature, it is reasonably evident that boosting and Random Forest are the best techniques for predicting corporate defaults. Moreover, differently from other machine learning techniques, the Random Forest model allows us to easily interpret the results – a fundamental factor for helping executives to improve their businesses. Barboza et al., (2017), despite the increasing number of studies in this field, suggested that “*new studies, exploring different models, contexts and datasets, are relevant, since results regarding the superiority of these models are still inconclusive*”. Therefore, this paper aims to fill the research gap on the ability of Random Forest and Logit models to predict not only the default of a company, but also its classification as UTP. This issue has not yet been adequately covered in academic literature, despite its remarkable impact from a managerial perspective. In fact, it intercepts the signals of a stressed situation and anticipates the potential status of default (Ambrosini & Tron, 2016; Caputo & Tron, 2016). This feature offers an important option for managers, who can implement several corrective actions. After the default and – even worse – after the bankruptcy, the number of adoptable solutions is strictly limited (Tron et al., 2018; Ferri et al., 2020). Furthermore, several papers have revealed the impact of corporate governance indicators (such as composition and the mandated duration of the main governance – and control – bodies, changes in majority shareholders; etc.) on default probability (Elloumi & Gueyi, 2001; Switzer et al., 2018; Lin et al., 2020; Fernando et al., 2020) and on the turnaround outcome (Miglani et al., 2020). However, previous studies have not yet thoroughly investigated the impact of governance variables in predicting the default of a company using new machine learning models, like Random Forest. Only Liang et al., (2016), on the basis of a sample of Taiwanese companies, demonstrated a better performance in predicting defaults, relying on corporate governance and financial variables, rather than using only financial variables. Therefore, as suggested by Barboza et al., (2017), we used machine learning models, also fed by corporate governance variables, to analyze the Italian context, along with its specific features in terms of company size and corporate governance.

### 3 Research methodology

Thanks to the support of a leading bank operating in Italy, and following the example of Dallochio et al., (2020), we collected data on Italian companies classified as UTP in 2014. The database structure allowed us to verify whether or not a company had been placed on the special register of “UTP positions”.

Out of an initial sample of 10,143 companies made available by the bank, we selected 72 that had the following features:

- private;
- set up and registered in Italy;
- out of trouble (“*in bonis*”) before August 2014 and still “*in bonis*” in 2017;
- able to repay interest and/or part of the principal payment;
- originally included in the “special register” of the bank;
- positively concluded a restructuring process at least two years before this study began.

Since the initial sample only included private Italian companies that were able to complete a full turnaround process, two additional samples were identified: firms that defaulted during the period between 2014 and 2016, and firms that did not demonstrate economic and financial distress during the period between 2007 and 2016. To construct these two samples, we followed a two-step process.

Firstly, by identifying their economic sector and using the revenues of the companies included in the first sample as an approximation of the size of a company, we identified, thanks to Bureau Van Dijk Database (AIDA), 34,124 companies (excluding listed companies) as our first sample.

Secondly, still drawing from the same source, we highlighted 76 companies that defaulted between 2014 and 2016, while the “healthy” companies (once again selected from the aforementioned cluster) were identified through a process of pairwise sampling, which allowed us to identify 72 entities. Therefore, the final database was composed of 220 companies: 72 of the original sample (success in restructuring), 72 healthy companies (no economic and/or financial distress between 2007 and 2016), and 76 companies that defaulted between 2014 and 2016. Then, we started collecting economic/financial data through AIDA, and governance data through the data provider CERVED<sup>3</sup>. As previously mentioned, existing literature on machine learning models (Barboza et al., 2017; Jones et al., 2017) has only used economic/financial variables. On top of this, we decided to include corporate governance variables that were used in this paper, along with their sources, can be found in Table 1.

To predict the classification of a company as a UTP position, as control variables, this study used one of the most adopted and easy-to-use tools for assessing default risk, the Z'-Score and Z''-Score. The Z'-Score was created by Altman (1993) as a revision for non-listed companies of the original Z-Score (1968). In analytical terms, the Z'-Score model is based on five factors:

- Liquidity ( $\frac{\text{CurrentAssets} - \text{CurrentLiabilities}}{\text{TotalAssets}}$ ),
- Profitability ( $\frac{\text{RetainedEarnings}}{\text{TotalAssets}}$ ),
- Productivity ( $\frac{\text{EBIT}}{\text{TotalAssets}}$ ),
- Leverage ( $\frac{\text{BookValueofEquity}}{\text{BookValueofTotalLiabilities}}$ ),
- Asset turnover ( $\frac{\text{Sales}}{\text{TotalAssets}}$ ).

<sup>3</sup> CERVED is an Italian listed company, which is a leading data provider for banks and credit agencies.

**Table 1** Variable definition and source

<i>Variable Name</i>	<i>Definition</i>	<i>Source</i>
n. CEO 07–11	Number of people who took the role of CEO in the period 2007–2011	CER- VED
n. CEO 12–14	Number of people who took the role of CEO in the period 2012–2014	CER- VED
n. Sole Directors 07–11	Number of people who took the role of Sole Director in the period 2007–2011	CER- VED
n. Sole Directors 12–14	Number of people who took the role of Sole Director in the period 2012–2014	CER- VED
Pres. BoD 07–11	Number of people who took the role of President of the Board of Directors in the period 2007–2011	CER- VED
Pres. BoD 12–14	Number of people who took the role of President of the Board of Directors in the period 2012–2014	CER- VED
Pres. BoD Avg. Term	Average duration of the role of the President of the Board of Directors	CER- VED
BoD Avg. Term	Average duration of the role of Board Member	CER- VED
Statutory Auditors	Dummy variable if there is the Board of Statutory Auditors	CER- VED
External Auditors	Dummy variable if there are External Auditors	CER- VED

The reliability of  $Z'$ -Score in measuring the health of small and medium enterprises in the Italian context has been discussed by several authors (Madonna & Cestari, 2015; Paoloni & Celli, 2018; Dallochio et al., 2020). For emerging countries and non-manufacturing companies, Altman elaborated upon the  $Z''$ -Score (1995), using a correction factor of 3.25 and deleting the asset turnover. In this case, the  $Z''$ -Score proved to be reliable in the Italian context (Altman et al., 2013). Due to the reliability in the Italian context of  $Z'$  and  $Z''$ , and in line with existing literature, we chose to use the two scores as predictive variables. Furthermore, since the  $Z$ -Scores allowed us to anticipate the emergence of a crisis, thanks to their ability to recognise the relationship between the potential corporate default and the accounting indicators in the years before insolvency, the management team can implement a coherent strategy for preventing the crisis (Altman & Le Fleur, 1985).

Their values were obtained by, once again, relying on AIDA. We collected data from 2007 until the “financial distress moment” (identified as the date of inclusion in the special register of the UTP’s position), as reported by the bank agents responsible for distressed debt positions, through interviews and surveys conducted by the authors. In particular, for each specific type of  $Z$ -Score, we calculated the average score across two periods: (i) before the financial distress (2007–2011); (ii) during the financial distress (2012–2014). Therefore, an overview over the financial variables that we use in this paper and their source can be found in Table 2.

Companies with missing data were finally discarded, allowing us to compare the results of the Logit, which suffered from missing data, and the Random Forest. Therefore, the final sample had 112 constituents: 54 successfully restructured companies, 13 defaulted companies, and 45 healthy companies. We controlled the weighting of every industry represented in the sample: no sector had a weighting higher than about

**Table 2** Variable definition and source

<i>Variable Name</i>	<i>Definition</i>	<i>Source</i>
Z'-Score 07–11	Average Z' -Score score in the period 2007–2011	AIDA
Z'-Score 12–14	Average Z' -Score score in the period 2012–2014	AIDA
Z''-Score 07–11	Average Z'' -Score score in the period 2007–2011	AIDA
Z''-Score 12–14	Average Z'' -Score score in the period 2012–2014	AIDA

7% for restructured and healthy companies. In Table 3, the descriptive statistics of corporate governance variables and Z-Score are reported.

The criteria adopted for the construction of the control samples appear consistent with the research perspective and reflect the actual health status of the companies analyzed (healthy, restructured, failed). We then constructed a dummy variable to approximate the state of the company (0=healthy company, 1=recovered or defaulted enterprise). We proceeded to analyze the correlations between the variables included in Table 4, which allowed us to obtain remarkable results. A relatively higher number of CEOs and board members demonstrated a positive relationship with the dummy variable, while their average term in office has a negative relationship. As expected, we identified a negative relationship between the Z-Score and the probability of default for all considered periods.

After developing the correlation analysis, we applied the two following models:

- Logit model.
- Random Forest Model.

Both of the models were run using, firstly, the Z'-Score and, secondly, the Z''-Score.

We selected the Random Forest technique, as several authors (i.e., Barboza et al., 2017) have shown it to be the best machine learning classifier. This model, robust to the presence of outliers or missing data, is able to identify the importance of each variable in the classification results. The model corrects the decision trees' habit of overfitting to their training set (Friedman, 2001; Schapire & Freund, 2012); an event that happens when the model too closely fits the training set. Estimated using Stata<sup>4</sup>, the Random Forest tree depth was set to 1,000, the number of predictors for each tree was set to 3, the bootstrap sample size was set to 1,000, and the minimum number of cases for parent node was set to 2. Following the example of Hastie et al., (2009), the classifiers were trained and tested on each dataset using repeated 70/30 random allocations between training and test samples. The training sample included, therefore, 78 companies, of which 47 were restructured or defaulted, and the remaining 31 were healthy. Given the small size of our sample, as a robustness check, we employed a 10th K-fold cross validation approach (see Hastie et al., 2009).

By comparing the two models, it is clear to see that both models have their own advantages and disadvantages. The Logit model is particularly suitable for the eco-

<sup>4</sup> Stata is a general-purpose statistical software created by StataCorp. More information can be found on the company's website: <https://www.stata.com/>.

**Table 3** Descriptive statistics by company status

<i>Restructured Companies</i>	No.	Mean	Std. Dev.	P1	P99
Z'-Score 07–11	54	1.33	1.34	-0.15	9.51
Z'-Score 12–14	54	1.00	0.97	-0.94	4.95
Z''-Score 07–11	54	3.97	3.35	0.71	23.76
Z''-Score 12–14	54	3.09	2.78	-5.27	12.53
n. CEO 07–11	54	0.80	1.38	0	7
n. CEO 12–14	54	0.56	0.92	0	4
n. Sole Directors 07–11	54	0.20	0.49	0	2
n. Sole Directors 12–14	54	0.04	0.19	0	1
Pres. BoD 07–11	54	0.54	0.69	0	2
Pres. BoD 12–14	54	0.37	0.59	0	2
Pres. BoD Avg. Term	54	8.02	5.16	1	23
BoD Avg. Term	54	7.18	3.85	1	18
Statutory Auditors	54	0.87	0.33	1	0
External Auditors	54	0.75	0.43	1	0
<b><i>Healthy Companies</i></b>	<b>No.</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>P1</b>	<b>P99</b>
Z'-Score 07–11	45	2.53	1.37	0.42	7.00
Z'-Score 12–14	45	2.75	1.55	0.19	7.84
Z''-Score 07–11	45	6.16	3.07	0.40	18.79
Z''-Score 12–14	45	6.68	3.51	0.01	21.23
N. CEO 07–11	45	0.60	0.98	0	3
N. CEO 12–14	45	0.30	0.69	0	3
N. Sole Directors 07–11	45	0.07	0.25	0	1
N. Sole Directors 12–14	45	0.07	0.25	0	1
Pres. BoD 07–11	45	0.44	0.76	0	2
Pres. BoD 12–14	45	0.29	0.63	0	3
Pres. BoD Avg. Term	45	10.84	7.36	3.09	26
BoD Avg. Term	45	10.73	5.42	1.00	31
Statutory Auditors	45	0.84	0.36	1	0
External Auditors	45	0.75	0.43	1	0
<b><i>Defaulted Companies</i></b>	<b>No.</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>P1</b>	<b>P99</b>
Z'-Score 07–11	13	0.96	1.00	-0.51	3.08
Z'-Score 12–14	13	0.05	1.67	-2.52	3.76
Z''-Score 07–11	13	3.97	3.00	-1.08	9.06
Z''-Score 12–14	13	1.76	4.44	-4.44	10.41
n. CEO 07–11	13	0.54	0.97	0	3
n. CEO 12–14	13	0.69	0.95	0	3
n. Sole Directors 07–11	13	0.46	0.66	0	2
n. Sole Directors 12–14	13	0.31	0.63	0	2
Pres. BoD 07–11	13	1.08	0.95	0	2
Pres. BoD 12–14	13	0.62	0.51	0	1
Pres. BoD Avg. Term	13	5.04	5.42	1	20
BoD Avg. Term	13	5.16	5.40	1	20.5
Statutory Auditors	13	0.38	0.51	1	0
External Auditors	13	0.38	0.51	1	0

conomic and financial field, since this model is appropriate for predicting binary events and does not require the independent variables to have equal variance in each group, or even be normally distributed (Hilbe, 2015). Moreover, the Logit model is less sub-

**Table 4** Correlation analysis

	Dummy	Z'-Score 07-11	Z'-Score 12-14	Z'-Score 07-11	Z'-Score 12-14	N. CEO 07-11	N. CEO 12-14	N. Sole Direc- tors 07-11	N. Sole Direc- tors 12-14	Pres. BoD 07-11	Pres. BoD 12-14	BoD Avg. Term	N. BoD 07-11	N. BoD 12-14	Statutory Auditors	Ex- ternal Audi- tors
Dummy	1.00															
Z'-Score	-0.4312	1.00														
07-11																
Z'-Score	-0.5805	0.8294	1.00													
12-14																
Z'-Score	-0.3216	0.8539	0.6429	1.00												
07-11																
Z'-Score	-0.4976	0.7343	0.8723	0.7965	1.00											
12-14																
N. CEO	0.0620	-0.0786	-0.0377	-0.0819	-0.0469	1.00										
07-11																
N. CEO	0.1702	-0.0453	-0.1052	-0.0378	-0.1304	0.1119	1.00									
12-14																
N. Sole Direc- tors	0.2049	-0.0134	-0.1567	0.0462	-0.1311	-0.0474	0.0169	1.00								
07-11																
N. Sole Direc- tors	0.0370	0.0017	-0.0757	-0.0522	-0.1342	-0.1321	-0.1459	0.0917	1.00							
12-14																

Table 4 (continued)

Dummy	Z'-Score 07-11	Z'-Score 12-14	Z'-Score 07-11	Z'-Score 12-14	N. CEO 07-11	N. CEO 12-14	N. Sole Direc- tors 07-11	N. Sole Direc- tors 12-14	Pres. BoD 07-11	Pres. BoD 12-14	Pres. BoD Avg. Term	N. BoD 07-11	N. BoD 12-14	Statutory Auditors	Ex- ternal Auditors	
Pres. BoD 07-11	0.1265	-0.2128	-0.1504	-0.1752	0.2086	0.1762	-0.0065	0.1132	1.00							
Pres. BoD 12-14	0.1059	-0.0940	-0.1181	-0.1080	0.2170	0.2825	0.1228	-0.1132	-0.0403	1.00						
Pres. BoD	-0.3803	0.2813	0.1979	0.2764	-0.1705	-0.1354	-0.2867	-0.3012	-0.3928	-0.2695	1.00					
Avg. Term	-0.2612	0.2506	0.1916	0.2637	-0.0964	-0.1716	-0.2515	-0.2556	-0.5294	-0.3526	0.7301	1.00				
N. BoD 07-11	0.2529	-0.2259	-0.2215	-0.2305	0.4203	0.2207	-0.0447	-0.0001	0.5146	0.2185	-0.4460	-0.3308	1.00			
N. BoD 12-14	0.2731	-0.1588	-0.1959	-0.2128	0.1638	0.3393	-0.0128	-0.1193	0.1387	0.4677	-0.3541	-0.2517	0.4134	1.00		
Statu- tory Auditors	-0.0843	0.1795	0.1939	0.0790	0.0024	0.1122	-0.1040	-0.1656	-0.1065	0.0396	0.1254	0.1614	0.0116	0.0925	1.00	
Ex- ternal Auditors	-0.0749	-0.0664	-0.0184	-0.1237	-0.1111	0.2049	-0.0033	-0.1010	-0.0280	0.1292	-0.0756	-0.1261	0.1917	0.1395	0.1848	1.00

ject to overfitting (Hilbe, 2015). On the contrary, the random forest model also has several advantages in that it is robust to outliers and to missing data (Lantz, 2019). This allows this model to be particularly suitable for analyzing small databases with an optimum level of generalization (Lantz, 2019).

The Logit model tends to require less computation and is easily interpretable when compared to the Random Forest model, as it is a linear model. However, the Random Forest model obtains a higher predicting performance compared to the Logit model (Jones et al., 2017) and does not require any adjustments to the databases used, since it does not require variables to be scaled or normalized (Lantz, 2019). Moreover, the Logit model is exposed to several limitations which do not affect the Random Forest model: heteroskedasticity, serial correlation, non-normality of error terms, and it is not suitable for nonlinear relationships.

In order to compare the predictive performance of the models, we decided to use the ROC curves: a method commonly used in previous literature (Swets et al., 2000; Jones et al., 2015, 2017; Barboza et al., 2017). The ROC curve, which plots the true positive rate (sensitivity) relative to the false positive rate (1 – specificity), would have an AUC (area under the curve) of exactly 0.5 in case of random guess. Therefore, every classifier should reach a value higher than 0.5. As suggested by Jones et al., (2017), a value higher than 0.9 signals a strong classifier, while a value between 0.8 and 0.9 is indicative of a good or useful classifier.

## 4 Results

Firstly, we ran the Logit model using the  $Z'$ -Score (Model 1) and the  $Z''$ -Score (Model 2) on the training sample.

In the two Logit models, corporate governance variables are generally not statistically significant in anticipating the default of a company. However, the turnover of the board of directors – which can often occur unexpectedly as a result of the actions of the shareholder that caused the crisis – is statically significant at 5% in both models and has a negative coefficient.

In both models, the  $Z$ -Scores are statically significant but with differing impacts in the case of the  $Z$ -Score relative to the period 2007–2011 or 2012–2014. As expected, the 12–14  $Z$ -Score is statically significant at 1% in both models, with a negative coefficient. Therefore, the models confirm the ability of  $Z$ -Score to anticipate the emergence of a crisis two years earlier than the occurrence of a company's classification as UTP. Furthermore, the sign of the coefficient confirms the negative (expected) relationship between the  $Z$ -Score and the dummy dependent variable, since the higher the  $Z$ -Score, the lower the probability of default. Interestingly, the impact of  $Z'$ -Score in predicting the status of a company is higher than  $Z''$ -Score. The 07–11  $Z$ -Score is also statically significant at 5% in both Model 1 and Model 2, but the sign of the coefficient is positive, which is an unexpected result. This coefficient could be a consequence of the use of the average of  $Z$ -Scores over the last five years, which is also a limitation of this work. Secondly, this could be due to the characteristics of the sample, which could therefore be influenced by outliers - a factor that does not influence the Random Forest technique.

**Table 5** Logit results

<i>Dependent variable</i>	Dummy	Dummy
<b>Models</b>	<b>(1)</b>	<b>(2)</b>
<i>Independent variables</i>		
Z Score' 07–11	1.564** (0.698)	
Z Score' 12–14	-3.104*** (0.954)	
Z Score'' 07–11		0.465** (0.234)
Z Score'' 12–14		-0.903*** (0.288)
Number CEO 07–11	-0.300 (0.367)	-0.330 (0.328)
Number CEO 12–14	0.543 (0.432)	0.405 (0.391)
Number Sole Directors 07–11	-1.317 (0.981)	-0.941 (0.903)
Number Sole Directors 12–14	-2.086 (1.979)	-2.199* (1.334)
Number Pres. BoD 07–11	-0.001 (0.698)	0.010 (0.607)
Number Pres. BoD 12–14	0.386 (0.920)	0.428 (0.812)
BoD Average Term	-0.431** (0.186)	-0.379** (0.151)
BoD Pres Average Term	0.064 (0.111)	0.045 (0.085)
Statutory Auditors	0.261 (0.991)	0.227 (0.915)
External Auditors	-1.549 (-1.068)	-1.842* (0.984)
Costant	6.987*** -2.516	6.504*** -2.226
Observations	78	78
***, **, and * indicates statistically significant levels of 1%, 5% and 10%, respectively.		
Notes: In Model 1 the Z'-Score was used while in Model 2 the Z''-Score was used		

The results of Tables 5 and 6 confirm the overall good ability of the Logit model to correctly classify the status of a company.

Secondly, we ran the Random Forest model using the Z'-Score (Model 3) and the Z''-Score (Model 4) on the training sample. The main results are shown in Table 7. The results show that Model 3, based on Z'-Score, is better than Model 4 at classifying the status of a company. However, the two models have high degrees of accuracy for both the training and test samples.

To interpret the model performance, we used the relative variables importance (RVIs), which reports the number of times on average that a variable is used in the decision trees of the model (Hastie et al., 2009; Friedman & Meulman, 2003). A RVI

**Table 6** *Logit Classification*

Models	(1)	(2)	
Sensitivity	Pr(+D)	91.49%	91.49%
Specificity	Pr( --D)	74.19%	77.42%
Positive predictive value	Pr( D +)	84.31%	86.00%
Negative predictive value	Pr( - D -)	85.19%	85.71%
False +rate for true -D	Pr(+ ~ D)	25.81%	22.58%
False - rate for true D	Pr( - D)	8.51%	8.51%
False +rate for classified	Pr(- D +)	15.69%	14.00%
False - rate for classified	Pr( D -)	14.81%	14.29%
<b>Correctly classified</b>		<b>84.62%</b>	<b>85.90%</b>

Notes:<sup>a</sup> In Model 1 the Z'-Score was used while in Model 2 the Z''-Score was used

**Table 7** Random forest

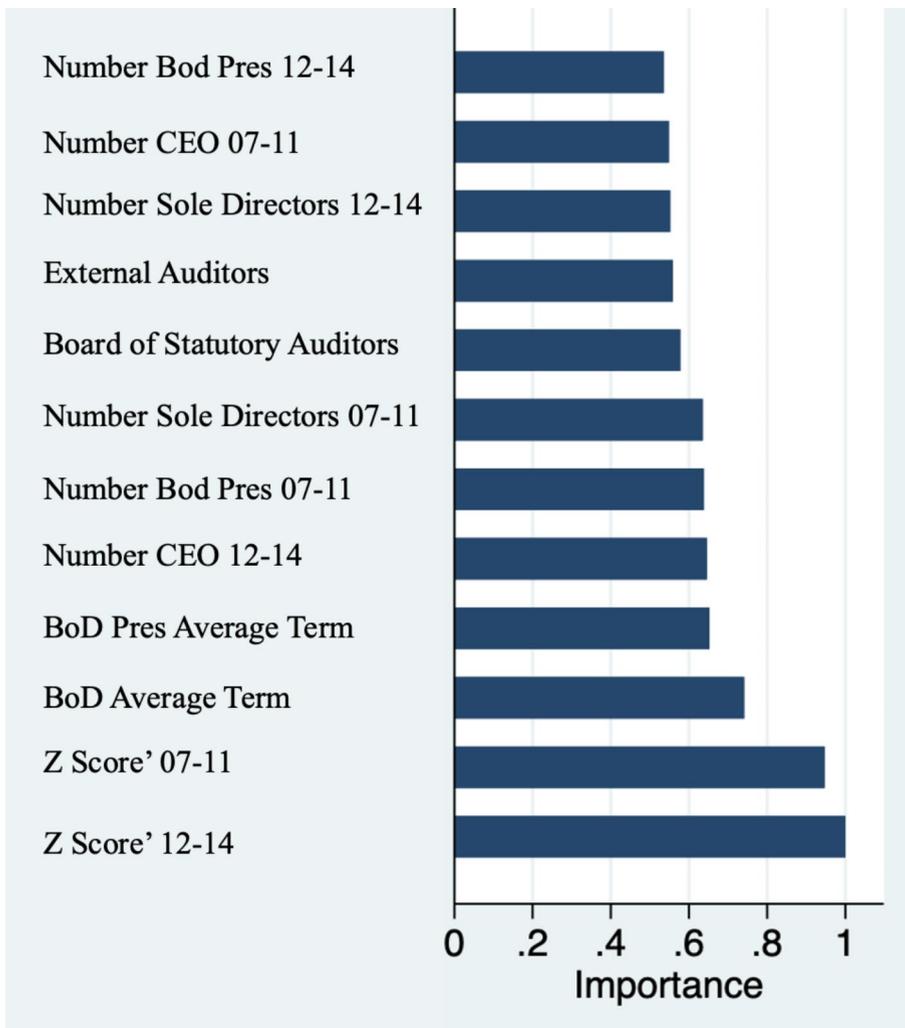
Models <sup>a</sup>	(3)	(4)
Observations	78	78
OOB Error	0.205	0.282
Correct Class	110	106
Incorrect Class	2	6
Error Rate	0.017	0.054

Notes:<sup>a</sup> In Model 3 the Z'-Score was used while in Model 4 the Z''-Score was used

bigger than 0 implies that the variable is used in the decision trees of the model and, therefore, contributes to improving the prediction capabilities of the model itself. RVIs are reported in Fig. 1 (Model 3) and Fig. 2 (Model 4) on a scale from 0 to 1.

Generally, all variables contribute to the overall predictive accuracy of the models. However, the strength of the RVIs differs significantly across variables. The results show that financial variables are still the most important indicators for predicting the financial default of a company. The 12–14 Z-Score is the most important variable in both models, confirming the results obtained in the Logit framework. The 07–11 Z-Score is the second most used variable in both models. Nevertheless, in Model 4 its importance is aligned to the 12–14 Z-Score, suggesting that the Random Forest model is able to predict the emergence of a crisis, giving the company time to find an appropriate solution. In opposition to Logit results, the governance variables demonstrate a good overall contribution in both models, which confirms the crucial role of corporate governance variables, as their RVI is higher than 0.5 for all variables. The turnover of the board of directors' members is rated third in terms of its importance, while the number of sole directors and chairmen of the board in the years prior to the crisis are the least relevant variables. Similarly, the presence of internal and external auditors does not seem to significantly affect the probability of default. On the contrary, the number of CEOs before the crisis is one of the most impactful variables, highlighting the importance of the relationship between the stability of the board of directors and, especially, of the person responsible for managing the company and the probability of an economic/financial crisis.

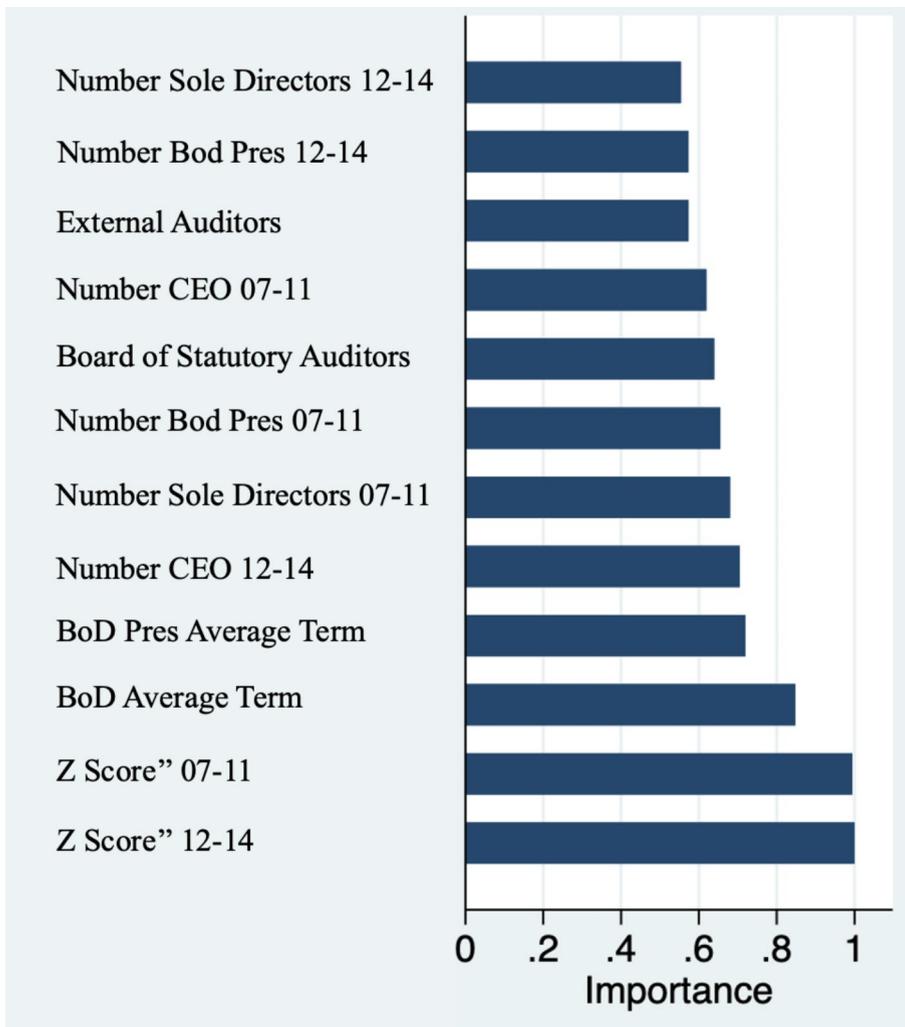
The RVIs show the difference between the Logit and the Random Forest models. Logit, highlights only three significant variables, while Random Forest shows that all of the variables contribute to improving the performance of the model. Jones



**Fig. 1** RVI's Model 3. *Notes: The Model 3 is based on the use of the Random Forest combines with the Z'-Score*

(2017) shows that Random Forest allows for the inclusion of many variables which are also highly correlated. In the case of Logit, this could lead to multicollinearity and overfitting.

Despite confirming the fact that financial variables are still the most important indicator in predicting the financial default of a company, the results also confirm the impact of corporate governance variables on the probability of bankruptcy. These results corroborate the importance of the members of the board of directors' turnover, as in Elloumi & Gueyiè (2001), suggesting that stability in the composition of the board of directors can positively impact the performance of a company. Furthermore, we also ran the models excluding corporate governance variables. In this framework,



**Fig. 2** RVI's Model 4. *Notes: The Model 4 is based on the use of the Random Forest combines with the Z''-Score*

we found that the ROC curve was higher for models that also included governance variables, as shown in Liang et al., (2016). Therefore, the Random Forest results reveal the importance of governance variables, especially with regards to the turnover of members on the board of directors and the number of CEOs in the year before the crisis. These results corroborate the theses of Gilson (1990), Fahlenbrach et al. (2007), and Fernando et al., (2020).

The results of the Logit and Random Forest models also confirm the Z''-Score and Z''-Score's ability to predict the emergence of a crisis, in line with previous literature (Dallocchio et al., 2020; Paoloni & Celli, 2018; Madonna & Cestari, 2016; Altman et al., 2013) researching the Italian context.

**Table 8** ROC

ROC - ALL SAMPLE				
Model	Obs	Area	Std. Err.	[95% Conf. Interval]
1 - Logit Z'-Score	112	0.9075	0.0266	0.85527
2 - Logit Z''-Score	112	0.8637	0.0346	0.79593
3 - RF Z'-Score	112	0.9849	0.0151	0.95523
4 - RF Z''-Score	112	0.9728	0.0152	0.94305
ROC - TEST SAMPLE				
Model	Obs	Area	Std. Err.	[95% Conf. Interval]
1 - Logit Z'-Score	34	0.8429	0.0702	0.70528 0.98044
2 - Logit Z''-Score	34	0.7821	0.0829	0.61958 0.94470
3 - RF Z'-Score	34	0.9357	0.0645	0.80935 1.00000
4 - RF Z''-Score	34	0.8357	0.0798	0.67928 0.99214

Furthermore, the Random Forest model corroborates the thesis that not only the Z-Score but also corporate governance variables can predict both the default of a company and its potential classification as UTP, which is of course a precursor. From a managerial standpoint, this is a key feature, because it allows a company to adopt an appropriate restructuring process well before the overall situation collapses. Anticipating the emergence of a crisis is becoming increasingly important in many countries, in which banking systems are suffering with the heavy burden of NPLs and correlated capital losses. This is also the case for Italy (along with many other continental European countries), which uses the new Bankruptcy Code, wherein the government requires companies and their managers to adopt restructuring procedures before events of credit. In this framework, controlling bodies or supervisory boards are also responsible for intercepting signs of economic and financial discomfort in time. Having said that, the option of using simple models, like the Random Forest model, could have a profound impact on the activities of both managing and controlling entities.

As previously mentioned, we have compared the predictive performance of the models using the ROC curves. The ROC curve data is shown in Tables 8, while the curves are shown in Fig. 3 (for all samples) and in Fig. 4 (for the test sample). The Random Forest shows a high degree of accuracy in the training phase. However, this does not imply high levels of reliability for the model, as its robustness and credibility actually depend on its ability to correctly predict the outcomes for the test sample. Nevertheless, both the overall results and the test results displayed in Table 8 show that the “new” machine learning classifier Random Forest significantly outperforms the traditional Logit model, both using Z'-Score and Z''-Score. The ROC area using the Random Forest and the Z'-Score is 0.9357 in the test sample: a value that signals a strong classifier. These results confirm those found in existing literature (Olson et al., 2012; Tsai et al., 2014; Barboza et al., 2017; Jones et al., 2017). Furthermore, these results confirm the suitability of Random Forest for the Italian context, along with its peculiarities, as in the case of Aliaj et al., (2020).

It is worth noting that Random Forest shows the predicting ability of the Z-Score, contrary to the case of Barboza et al., (2017). However, Barboza et al., (2017) used the original the Z-Score, while we used the Z'- and the Z''-Scores; the first of which, in particular, was shown to have high levels of reliability. We also included governance variables, confirming the theories of the aforementioned study, showing that

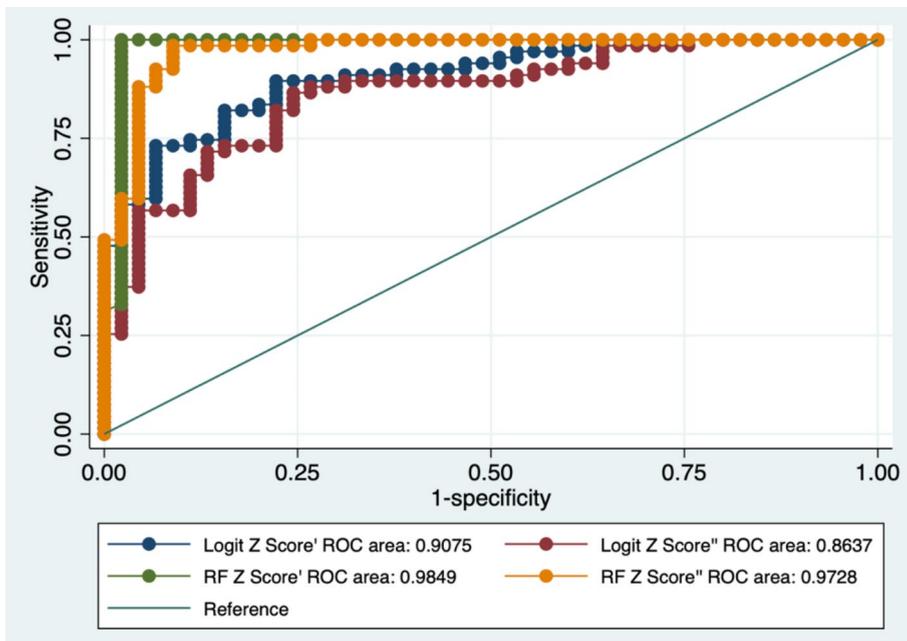


Fig. 3 ROC Curve ALL SAMPLE

Random Forest allows us to include different indicators and, therefore, predict bankruptcy cases more efficiently.

Despite this fact, the ROC area varies drastically as a function of the inclusion of  $Z'$ - or  $Z''$ -Score. The ROC curve in the test sample is 0.9357 in the case of Model 3, and it is 0.8393 in Model 4. It reaches 0.8321 in Model 1 and 0.7750 in Model 2. The results show that both Logit and Random Forest benefit from improved predicting power when  $Z'$ -Score is considered. This result could be linked to the fact that  $Z''$ -Score was mainly built for use in emerging markets. This implies that the ratio of Sales/Total Assets, which is not included in the  $Z''$ -Score, could represent a key factor when anticipating the emergence of a crisis.

However, without exception, in all tests the Random Forest revealed better outcomes than those of the Logit model, due to its ability to better use the information contained in the corporate governance variables.

## 5 Robustness tests

To check the robustness of our results, different tests were carried out. Firstly, we assumed a 10th K-fold cross validation approach. The K-fold cross validation, as suggested by Hastie et al., (2009), is particularly useful when working with a small sample size, as in our case. This technique is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. Therefore, it is used on limited samples in order to estimate how the model is expected to perform

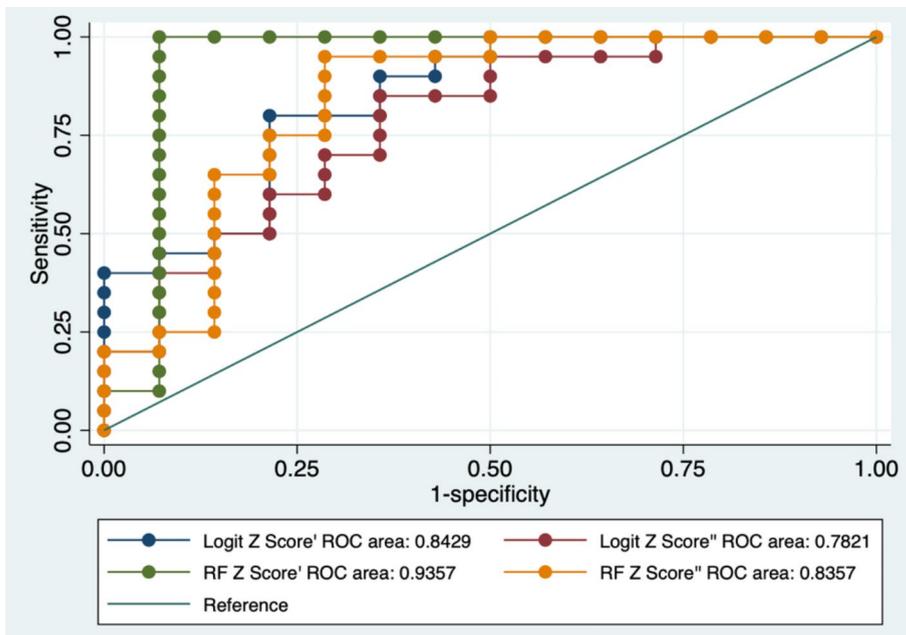


Fig. 4 ROC Curve TEST SAMPLE

Table 9 K-Fold cross validation

Model	1 - Logit Z'-Score		2 - Logit Z''-Score		3 - RF Z'-Score		4 - RF Z''-Score	
	Full	Test	Full	Test	Full	Test	Full	Test
Sensitivity	89.55%	86.57%	88.06%	79.10%	100.00%	88.06%	100.00%	86.57%
Specificity	71.11%	68.89%	71.11%	57.79%	100.00%	77.78%	100.00%	64.44%
Positive predictive value	82.19%	80.56%	81.94%	73.61%	100.00%	85.51%	100.00%	78.38%
Negative predictive value	82.05%	77.50%	80.00%	65.00%	100.00%	81.40%	100.00%	76.32%
False+rate for true~D	28.89%	31.11%	28.89%	42.22%	0.00%	22.22%	0.00%	35.56%
False - rate for true D	10.45%	13.43%	11.94%	20.90%	0.00%	11.94%	0.00%	13.43%
False+rate for classified	17.81%	19.44%	18.06%	26.39%	0.00%	14.49%	0.00%	21.62%
False - rate for classified -	17.95%	22.50%	20.00%	35.00%	0.00%	18.60%	0.00%	23.68%
Correctly classified	82.14%	79.46%	81.25%	70.54%	100.00%	83.93%	100.00%	77.68%
ROC area	0.9240	0.8305	0.8793	0.7891	1.000	0.8693	1.000	0.7904
p-value for Full vs. Test ROC areas		0.0000		0.0000		0.0010		0.0000

in general when used to make predictions on data not used during the training of the model. We used this technique on both the Logit and the Random Forest model. The results are shown in Table 9.

The results did not show a significant change in AUC performance for Logit models, but the Random Forest model suffered from a 5% reduction, especially in the case of Model 4. However, the results confirm the superiority of Random Forest and Z'-Score (Model 3), which is still a strong classifier. This technique again confirms the superiority of the Z'-Score for both approaches, and the role of corporate governance variables in the Random Forest model. Moreover, this thesis is also corroborated by the fact the ROC area in the test sample is superior in the case of Model 1 (Logit – Z'-Score) in comparison to that of Model 4 (Random Forest – Z''-Score).

Secondly, as we used a dummy variable for three different possible statuses related to the company (healthy, restructured, or ceased), we ran the model excluding defaulted companies in order to have only two possible causes for the dummy. The new sample, therefore, only includes 99 companies. The results of the Logit are reported in Table 10. The 12–14 Z-Score is again statically significant at 1% in the case of Model 5, and at 5% in the case of Model 6. However, in this case, governance variables and the 07–11 Z-Score are statically significant only in Model 5.

The new ROC curves are shown in Tables 11 and are drawn in Figs. 5 and 6. Previous results are also confirmed in this case. The Random Forest model again demonstrates a higher predicting power than the Logit model in all cases and samples. In this case, the Z'-Score and corporate governance variables in the Random Forest also prove to be better discriminant indicators for anticipating the emergence of a crisis. This conclusion is also reinforced by the fact that the ROC area is higher for Model 5 than Model 8, as in the case of the K-Fold test.

## 6 Conclusions

This paper extends prior empirical research on financial distress and corporate governance in a geographical context such as Italy. Italian companies' corporate governance system characteristics are more likely to heighten agency problems and, therefore, they could contribute to worsening situations of financial distress. The final results corroborate the thesis on the central role of corporate governance on crisis management – a topic that is relatively new in comparison to studies on financially sound enterprises. This research has produced evidence regarding the importance of corporate governance variables – especially those linked to the board of directors and the top management of a company – in anticipating the emergence of bankruptcies, while the presence of auditors seems to be less relevant. Although financial variables are the most crucial factors in all models, the Random Forest model shows that corporate governance variables play a primary role; especially the renewal of the CEO and the average term of the board of directors. These results suggest that the stability of the CEO, the composition of the board of directors, and the person responsible for managing the company can deeply affect the probability of an economic/financial crisis. This thesis is also corroborated by the fact that the exclusion of these variables deeply affects the performance of the models, reducing their capabilities when predicting bankruptcies. These findings have important implications for family-owned firms and for banks with regards to improving the performance of their credit models.

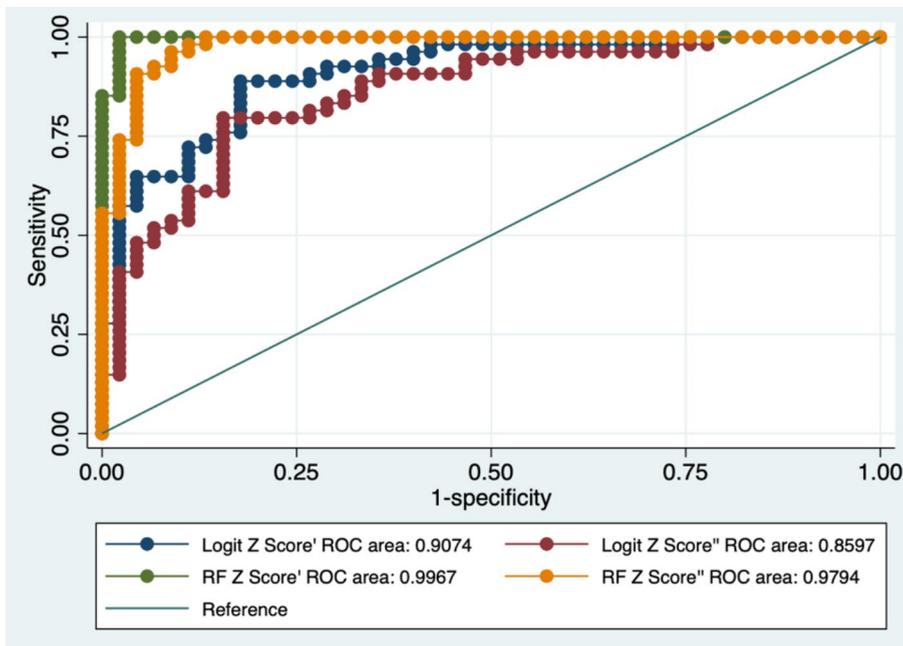
**Table 10** Logit new

<i>Dependent variable</i>	Dummy	Dummy
<b>Models</b>	<b>(5)</b>	<b>(6)</b>
<i>Independent variables</i>		
Z Score' 07–11	1.558** (0.731)	
Z Score' 12–14	-3.349*** (1.083)	
Z Score'' 07–11		0.351 (0.267)
Z Score'' 12–14		-0.714** (0.307)
Number CEO 07–11	0.013 (0.320)	-0.044 (0.294)
Number CEO 12–14	0.294 (0.358)	0.239 (0.336)
Number Sole Directors 07–11	-0.191 (1.202)	0.413 (1.183)
Number Sole Directors 12–14	3.978 (2.892)	-0.615 (1.723)
Number Pres. BoD 07–11	-0.072 (0.636)	-0.035 (0.549)
Number Pres. BoD 12–14	-0.113 (0.867)	-0.129 (0.731)
BoD Average Term	-0.186 (0.119)	-0.203** (0.097)
BoD Pres Average Term	-0.035 (0.110)	0.000 (0.077)
Statutory Auditors	1.641 (1.002)	0.728 (0.859)
External Auditors	-0.944 (1.034)	-0.950 (0.957)
Costant	4.703** -2.097	3.919** -1.813
Observations	77	77
***, **, and * indicates statistically significant levels of 1%, 5% and 10%, respectively.		
Notes: In Model 5 the Z'-Score was used while in Model 6 the Z''-Score was used		

Using a unique set of UTP Italian companies, we compare the Logit and Random Forest models' ability to predict bankruptcies. Despite their low diffusion and use, we confirm that the Random Forest outperforms the Logit model in predicting corporate defaults. Our findings and suggestions for corporate default predictions are as topical as ever. Firstly, machine learning techniques have been proven to be particularly effective in forecasting corporate defaults. Secondly, they have remarkable practical applications for various business operators, such as credit agencies and banks. These techniques, which are relatively easy to implement, are stable predictors and are resilient to a series of statistical issues, like omitted variables, multicollinearity, outliers, and heteroscedasticity. Moreover, the RVIs of the Random Forest technique

**Table 11** ROC new

ROC - ALL SAMPLE				
Model	Obs	Area	Std. Err.	[95% Conf. Interval]
5 - Logit Z'-Score	99	0.9074	0.0294	0.84984 0.96498
6 - Logit Z''-Score	99	0.8597	0.0374	0.78638 0.93296
7 - RF Z'-Score	99	0.9967	0.0035	0.98991 1.00000
8 - RF Z''-Score	99	0.9794	0.0122	0.95545 1.00000
ROC - TEST SAMPLE				
Model	Obs	Area	Std. Err.	[95% Conf. Interval]
5 - Logit Z'-Score	22	0.8803	0.0735	0.73624 1.00000
6 - Logit Z''-Score	22	0.7949	0.0999	0.59915 0.99059
7 - RF Z'-Score	22	0.9829	0.0205	0.94274 1.00000
8 - RF Z''-Score	22	0.8376	0.0869	0.66725 1.00000



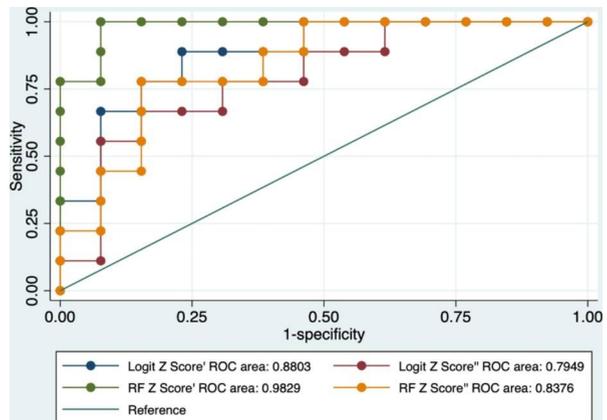
**Fig. 5** New ROC Curve ALL SAMPLE

also allow us to interpret the importance of the variables; a fundamental factor helping managers with key activities targeted towards the prevention of crises.

Furthermore, the results of the models show the predictive power of Z'- and Z''-Scores in the Italian context specifically, corroborating the theses of previous researchers. In addition, we also confirm that the Z'-Score is a better indicator in anticipating potential corporate disease.

The main limitation of our work is, of course, the sample, which only includes Italian non listed UTP companies. This fact surely impacts upon our results. However, we also consider this feature an additional contribution of our research. As a matter of fact, we demonstrate the reliability of these models in predicting the classification of

**Fig. 6** New ROC Curve TEST SAMPLE



UTP for private firms. This is a factor of pivotal importance, because it anticipates the default and can act as a stimulus when fixing emerging problems. Future researchers, given the ability of these models to manage a great quantity of data, should include more countries, considering both listed and private companies and despite the low availability of data, possibly also the date a company was classified as UTP. Secondly, because overfitting is not a significant issue for the main machine learning techniques, future researchers could include more predictive variables, especially those linked to sustainability, given the limited amount of existing literature on the matter (Elloumi & Gueyiè, 2001; Fahlenbrach et al., 2007; Ricci et al., 2020).

**Funding** Open access funding provided by Università Commerciale Luigi Bocconi within the CRUI-CARE Agreement.

**Declarations of interest** None.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agostini, L., & Nosella, A. (2017). Enhancing radical innovation performance through intellectual capital components. *Journal of Intellectual Capital*, 18(4), 789–806. <https://doi.org/10.1108/JIC-10-2016-0103>
- Aliaj, T., Anagnostopoulos, A., & Piersanti, S. (2020). Firms Default Prediction with Machine Learning. In V. Bitetta, I. Bordino, A. Ferretti, F. Gullo, & S. Pascolutti, Ponti G. (Eds.) (Eds.), *Mining Data for Financial Applications. MIDAS 2019* (11985 vol.). Cham: Springer. Lecture Notes in Computer Science [https://doi.org/10.1007/978-3-030-37720-5\\_4](https://doi.org/10.1007/978-3-030-37720-5_4)

- Althey, S. (2018). *The Economics of Artificial Intelligence: An Agenda* (forthcoming). University of Chicago Press, Retrieved from: <https://www.gsb.stanford.edu/faculty-research/publications/impact-machine-learning-economics> (last access: 04/30/2022)
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Altman, E. I. (1993). *Corporate financial distress and bankruptcy* (2nd ed.). New York, NY: Wiley & Sons
- Altman, E. I., & Le Fleur, J. K. (1985). I modelli di previsione delle insolvenze: Le loro applicazioni alla gestione d'impresa. *Finanza, Marketing, Produzione*, 4, 77–93. Retrieved from: <http://www.biblio.liuc.it/scripts/essper/ricerca.asp?tipo=scheda&codice=10037013> (last access: 04/30/2022)
- Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (the Italian Experience). *Journal of Banking & Finance*, 18(3), 505–529. [https://doi.org/10.1016/0378-4266\(94\)90007-8](https://doi.org/10.1016/0378-4266(94)90007-8)
- Altman, E. I., & Narayanan, P. (1997). An international survey of business failure classification models. *Financial Markets. Institutions and Instruments*, 6(2), 1–57. <https://doi.org/10.1111/1468-0416.00010>
- Altman, E. I., & Saunders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking & Finance*, 21(11–12), 1721–1742. [https://doi.org/10.1016/S0378-4266\(97\)00036-8](https://doi.org/10.1016/S0378-4266(97)00036-8)
- Altman, E. I., Danovi, A., & Falini, A. (2013). Z-score model's application to Italian companies subject to extraordinary administration. *Journal of Applied Finance*, 23(1), 1–10. Retrieved from: <http://people.stern.nyu.edu/ealtman/BOZZA%20ARTICOLO%2017.pdf> (last access: 04/30/2022)
- Ambrosini, S., & Tron, A. (2016). *Piani di ristrutturazione dei debiti e ruolo dell'attestatore. Principi di attestazione e riforma del 2015*. Bologna: Zanichelli Editore
- Ambrosini, S., Andreani, G., & Tron, A. (2013). *Crisi d'impresa e restructuring. Aspetti economico-aziendali, giuridici e fiscali alla luce delle applicazioni e interpretazioni delle norme introdotte dalla legge 134/2012*, Il Sole 24 Ore, Milano
- Amigoni, F. (1998). *Manuale di organizzazione*. Turin, Italy: UTET
- Attaran, N., Hosseini, S., Hossein, S., & Karami, G. (2012). Bankruptcy prediction, accounting variables and economic development: empirical evidence from Iran. *International Business Research*, 5(8), 147–152
- Aziz, A., Emanuel, D. C., & Lawson, G. H. (1988). Bankruptcy prediction – An investigation of cash flow based models. *Journal of Management Studies*, 25(5), 419–437. <https://doi.org/10.1111/j.1467-6486.1988.tb00708.x>
- Back, B., Laitinen, T., Sere, K., & van Wezel, M. (1996). *Choosing bankruptcy predictors using discriminant analysis, Logit analysis, and genetic algorithms* (Turke Centre for Computer Science, Technical Report No. 40). Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.28.4162> (last access: 04/30/2022)
- Barboza, F., Kimura, H., & Altman, E. I. (2017). Machine learning models and bankruptcy prediction. *Expert System with Applications*, 83(C), 405–417. <https://doi.org/10.1016/j.eswa.2017.04.006>
- Beasley, M. S., Carcello, J. V., & Hermanson, D. R. (2000). *Fraudulent financial reporting: 1987–1997. An analysis of U.S. public companies*. New York: Committee of Sponsoring Organizations of the Treadway Commission (COSO)
- Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71–111. <https://doi.org/10.2307/2490171>
- Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies*, 1(4), 267–284. <https://doi.org/10.1007/BF00570833>
- Beyonon, M. J., & Peel, M. J. (2001). Variable precision rough set theory and data discretisation: An application to corporate failure prediction. *Omega*, 29(6), 561–576. [https://doi.org/10.1016/S0305-0483\(01\)00045-7](https://doi.org/10.1016/S0305-0483(01)00045-7)
- Bhimani, A., Gulamhussen, M., & Lopes, S. (2009). The effectiveness of the auditor's going-concern evaluation as an external governance mechanism: Evidence from loan defaults. *The International Journal of Accounting*, 44(3), 239–255. <https://doi.org/10.1016/j.intacc.2009.06.002>
- Booth, A., Gerding, E., & McGroarty, F. (2014). Automated trading with performance weighted Random Forests and seasonality. *Expert Systems with Applications*, 41(8), 3651–3661. <https://doi.org/10.1016/j.eswa.2013.12.009>
- Bredart, X. (2016). Financial Distress and Corporate Governance: The Impact of Board Configuration. *International Business Research*, 7(3), 72–80. <https://doi.org/10.5539/ibr.v7n3p72>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>

- Bragoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., Mezzogori, D., & Zammori, F. (Eds.). (2019). Machine Learning models for bankruptcy prediction in Italy: do industrial variables count?, *Working Paper N. 19/3 Dipartimento Di Matematica Per Le Scienze, Economiche, Finanziarie Ed Attuariali, Vita e Pensiero*, Milan, 3–41. <http://hdl.handle.net/10807/143387>
- Bjorkegren, D., & Grissen, D. (2019). Behavior revealed in mobile phone usage predicts loan repayment. *The World Bank Economic Review*, 34(3), 618–634. <https://doi.org/10.1093/wber/lhz006>
- Bratten, B., Gaynor, L. M., McDaniel, L., Montague, N. R., & Sierra, G. E. (2013). The audit of fair values and other estimates: The effects of underlying environmental, task, and auditor-specific factors. *Auditing: A Journal of Practice & Theory*, 32(sp1), 7–44
- Bredart, X. (2014). Financial Distress and Corporate Governance: The Impact of Board Configuration. *International Business Research*, 7(3), DOI: <https://doi.org/10.5539/ibr.v7n3p72>
- Brockman, P., & Turtle, H. J. (2003). A barrier option framework for corporate security valuation. *Journal of Financial Economics*, 67(3), 511–529. [https://doi.org/10.1016/S0304-405X\(02\)00260-X](https://doi.org/10.1016/S0304-405X(02)00260-X)
- Calderoni, L., Ferrara, M., Franco, A., & Maio, D. (2015). Indoor localization in a hospital environment using Random Forest classifiers. *Expert Systems with Applications*, 42(1), 125–134. <https://doi.org/10.1016/j.eswa.2014.07.042>
- Caputo, A., & Tron, A. (2016). The attestation of corporate turnaround plans in Italy: operating problems and possible solutions. *International Journal of Critical Accounting*, 8(1), 0–44. DOI: <https://doi.org/10.1504/IJCA.2016.076097>
- Cenciarelli, V., Greco, G., & Allegrini, M. (2018). External audit and bankruptcy prediction. *Journal of Management and Governance*, 22, 863–890. <https://doi.org/10.1007/s10997-018-9406-z>
- Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance*, 8(4), 537–569. <https://doi.org/10.1093/rof/8.4.537>
- Chen, C. J., Liu, T. C., Chu, M. A., & Hsiao, Y. C. (2014). Intellectual capital and new product development. *Journal of Engineering and Technology Management*, 33, 154–173
- Contreras, O., Arias, J., & Maquieira, C. (2021). Risk taking behavior in Chilean listed family firms: a socioemotional wealth approach. *International Entrepreneurship and Management Journal*, 17, 165–184. <https://doi.org/10.1007/s11365-019-00628-y>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Daily, C., & Dalton, D. (1994). Corporate governance & the bankrupt firm: an empirical assessment. *Strategic Management Journal*, 15(8), 643–654. <https://doi.org/10.1002/smj.4250150806>
- Dalocchio, M., & Tron, A. (2020). *Le operazioni di ristrutturazione del debito in Italia: note economico-aziendali a margine di una rassegna diacronica delle principali ricerche empiriche svolte negli ultimi anni*, in AA.VV., Contributi in onore di Gaetano Maria Golinelli, 358–392, Rogiosi Editore, Napoli
- Dalocchio, M., Ferri, S., Tron, A., & Vizzaccaro, M. (2020). The use of Z-Score to predict UTP loans. *Corporate Ownership & Control*, 18(1), 163–178. <https://doi.org/10.22495/cocv18i1art13>
- Davis, E. P., & Karim, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, 4(2), 89–120. <https://doi.org/10.1016/j.jfs.2007.12.004>
- Deng, X., & Wang, Z. (2006). Ownership structure and financial distress: Evidence from public-listed companies in China. *International Journal of Management*, 23, 486–502. <https://doi.org/10.2753/CES1097-1475390501>
- Donato, F., & Nieddu, L. (2020). Corporate failure: Bankruptcy prediction for Italian SMEs based on a longitudinal case study from 2000 to 2011. *Corporate Ownership & Control*, 17(3), 27–33. <https://doi.org/10.22495/cocv17i3art2>
- Donker, H., Santen, B., & Zahir, S. (2009). Ownership structure and the likelihood of financial distress in the Netherlands. *Applied Financial Economics*, 19, 1687–1696. <https://doi.org/10.1080/09603100802599647>
- Eccles, R. G. (1991). The performance measurement manifesto. *Harvard Business Review*, 69(1). Retrieved from: <https://hbr.org/1991/01/the-performance-measurement-manifesto> (last access: 04/30/2022)
- Edmister, R. O. (1972). An empirical test of financial ratios analysis for small business failure prediction. *The Journal of Finance and Quantitative Analysis*, 7(2), 1477–1493. <https://doi.org/10.2307/2329929>
- Efron, B. (1975). The Efficiency of Logistic Regression Compared to Normal Discriminant Analysis. *Journal of the American Statistical Association*, 70(352), 892–898. DOI: <https://doi.org/10.1080/01621459.1975.10480319>
- Elloumi, F., & Gueyie, J. P. (2001). Financial distress and corporate governance: An empirical analysis. *Corporate Governance*, 1(1), 15–23. <https://doi.org/10.1108/14720700110389548>

- Ernst & Young (2010). *Ambitious, adept and agile: how global entrepreneurs are changing the world – agility and cultural diversity run circles around established firms*
- Fahlenbrach, R., & Stulz, R. M. (2007). Managerial ownership dynamics and firm value. *NBER Working Paper No. 13202*. <https://doi.org/10.3386/w13202>
- Falavigna, G. (2012). Financial ratings with scarce information: A neural network approach. *Expert Systems with Applications*, 39(2), 1784–1792. <https://doi.org/10.1016/j.eswa.2011.08.074>
- Fernando, J. M. R., Li, L., & Hou Yang, G. (2020). Corporate governance and correlation in corporate defaults. *Corporate Governance International Review*, 28, 188–206. <https://doi.org/10.1111/corg.12306>
- Ferri, S., Tron, A., Fiume, R., & Della Corte, G. (2020). The relation between cash flows and economic performance in the digital age: an empirical analysis. *Corporate Ownership & Control*, 17(3), 84–91. <https://doi.org/10.22495/cocv17i3art6>
- Fijorek, K., & Grotowski, M. (2012). Bankruptcy prediction: some results from a large sample of Polish companies. *International Business Research*, 5(9), 70–77. <https://doi.org/10.5539/ibr.v5n9p70>
- Foreman, R. D. (2003). A logistic analysis of bankruptcy within the US local telecommunications industry. *Journal of Economics and Business*, 55(2), 135–166. [https://doi.org/10.1016/S0148-6195\(02\)00133-9](https://doi.org/10.1016/S0148-6195(02)00133-9)
- Friedman, J. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, V9(5), 1189–232. Retrieved from: <http://luthuli.cs.uiuc.edu/~daf/courses/Optimization/Papers/2699986.pdf> (last access: 04/30/2022)
- Friedman, J., & Meulman, J. (2003). Multiple Additive Regression Trees with Application in Epidemiology. *Statistics in Medicine*, 22(99), 1365–1381. <https://doi.org/10.1002/sim.1501>
- Gales, L. M., & Kesner, I. F. (1994). An analysis of board of director size & composition in bankrupt organizations. *Journal of Business Research*, 30(3), 271–282. [https://doi.org/10.1016/0148-2963\(94\)90057-4](https://doi.org/10.1016/0148-2963(94)90057-4)
- Garzella, S. (2005). *Il sistema d'azienda e la valorizzazione delle "potenzialità inesprese", Una "visione" strategica per il risanamento*. Turin: Giappichelli
- Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research*, 241(1), 236–247. <https://doi.org/10.1016/j.ejor.2014.08.016>
- Gilson, S. (1990). Bankruptcy, boards, banks, & blockholders. *Journal of Financial Economics*, 27(2), 355–387. [https://doi.org/10.1016/0304-405X\(90\)90060-D](https://doi.org/10.1016/0304-405X(90)90060-D)
- Goodstein, J., Gautam, K., & Boeker, W. (1994). The effects of board size and diversity on strategic change. *Strategic Management Journal*, 15(3), 241–250. <https://doi.org/10.1002/smj.4250150305>
- Grant, R. M. (2011). *L'analisi strategica per le decisioni aziendali*. Bologna: Il Mulino
- Guo, J., Huang, P., Zhang, Y., & Zhou, N. (2016). The Effect of Employee Treatment Policies on Internal Control Weaknesses and Financial Restatements. *The Accounting Review*, 9(4), 1167–1194
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (New York: Springer). Retrieved from: <https://web.stanford.edu/~hastie/Papers/ESLII.pdf> (last access: 04/30/2022)
- Heo, J., & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing*, 24, 494–499. <https://doi.org/10.1016/j.asoc.2014.08.009>
- Hilbe, J. (2015). *Practical guide to logistic regression*. New York: CRC Press
- Huson, M. R., Malatesta, P. H., & Parrino, R. (2004). Managerial succession and firm performance. *Journal of Financial Economics*, 74(2), 237–275. <https://doi.org/10.1016/j.jfineco.2003.08.002>
- Jensen, M. (1993). The modern industry revolution, exit, and the failure of internal control systems. *Journal of Finance*, 48(3), 831–880. <https://doi.org/10.1111/j.1540-6261.1993.tb04022.x>
- Jones, S., & Hensher, D. (2004). Predicting Firm Financial Distress: A Mixed Logit Model. *The Accounting Review*, 9(4), 1011 – 1038. Retrieved from: <http://www.jstor.org/stable/4093084> (last access: 04/30/2022)
- Jones, S. (2017). Corporate bankruptcy prediction: a high dimensional analysis. *Review of Accounting Studies*, 22(3), 1366–1422. <https://doi.org/10.1007/s11142-017-9407-1>
- Jones, S., Johnstone, D., & Wilson, R. (2015). An Empirical Evaluation of the Performance of Binary Classifiers in the Prediction of Credit Ratings Changes. *Journal of Banking and Finance*, 56, 72–85. Retrieved from: <http://www.sciencedirect.com/science/article/pii/S0378426615000333> (last access: 04/30/2022)
- Jones, S., Johnstone, D., & Wilson, R. (2017). Predicting Corporate Bankruptcy: An Evaluation of Alternative Statistical Frameworks. *Journal of Business Finance & Accounting*, 44(1), 3–34. DOI: <https://doi.org/10.1111/jbfa.12218>

- Kahya, E., & Theodossiou, P. (1999). Predicting corporate financial distress: A time-series CUSUM methodology. *Review of Quantitative Finance and Accounting*, 13(4), 323–345. <https://doi.org/10.1023/A:1008326706404>
- Kim, Y. (1996). Long-term firm performance and chief executive turnover: an empirical study of the dynamics. *Journal of Law Economics and Organization*, 12(2), 480–496
- Kim, H., Cho, H., & Ryu, D. (2020). Corporate Default Predictions Using Machine Learning. *Literature Review Sustainability*, 12(16), 6325. <https://doi.org/10.3390/su12166325>
- Kim, S. Y., & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and ada-boosted decision tree models. *Economic Modelling*, 36, 354–362. <https://doi.org/10.1016/j.econmod.2013.10.005>
- Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, 105(5), 491–495. <https://doi.org/10.1257/aer.p20151023>
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (2000). Investor protection and corporate governance. *Journal of Financial Economics*, 58(2), 3–27. [https://doi.org/10.1016/S0304-405X\(00\)00065-9](https://doi.org/10.1016/S0304-405X(00)00065-9)
- Lantz, B. (2019). *Machine Learning with R*. Birmingham: Packt
- Leng, J., Ozkan, A., & Trzeciakiewicz, A. (2011). Can managerial (over) confidence lead firms to bankruptcy? Evidence from the UK. EFMA 2018 Annual Meeting
- Leng, J., Ozkan, A., Ozkan, N., & Trzeciakiewicz, A. (2021). CEO overconfidence and the probability of corporate failure: evidence from the United Kingdom. *The European Journal of Finance*, 27(12), 1210–1234
- Li, Y. C., & Wang, Y. F. (2018). Machine Learning Methods of Bankruptcy Prediction Using Accounting Ratios. *Open Journal of Business and Management*, 6(1), 1–20. <https://doi.org/10.4236/ojbm.2018.61001>
- Liang, D., Lu, C. C., Tsai, C. F., & Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, 252(2), 561–572. <https://doi.org/10.1016/j.ejor.2016.01.012>
- Lin, B., Liu, C., Tan, K. J. K., & Zhou, Q. (2020). CEO turnover and bankrupt firms' emergence. *Journal of Business Finance & Accounting*, 47(9–10), 1238–1267. <https://doi.org/10.1111/jbfa.12482>
- Lin, L., & Piesse, J. (2004). The identification of corporate distress: A conditional probability analysis approach. *Applied Financial Economics*, 14, 73–82. Retrieved from: [https://www.researchgate.net/publication/313178035\\_The\\_identification\\_of\\_corporate\\_distress\\_in\\_UK\\_industrials\\_a\\_conditional\\_probability\\_analysis\\_approach](https://www.researchgate.net/publication/313178035_The_identification_of_corporate_distress_in_UK_industrials_a_conditional_probability_analysis_approach) (last access: 04/30/2022)
- Lohrke, F. T., Bedeian, A. G., & Palmer, T. B. (2004). The role of top management teams in formulating and implementing turnaround strategies: A review and research agenda. *International Journal of Management Reviews*, 5(6(2)), 63–90. <https://doi.org/10.1111/j.1460-8545.2004.00097.x>
- Madonna, S., & Cestari, G. (2015). The accuracy of bankruptcy prediction models: a comparative analysis of multivariate discriminant models in Italian context. *European Scientific Journal*, 11(34), 1–24. Retrieved from: <https://iris.unica.it/retrieve/handle/11584/190111/363422/Madonna%20%26%20Cestari%20ESJ.pdf> (last access: 04/30/2022)
- Manzanegue, M., Priego, A., & Merino, E. (2016). Corporate governance effect on financial distress likelihood: Evidence from Spain. *Spanish Accounting Review*, 19(1), 111–121. <https://doi.org/10.1016/j.rcsar.2015.04.001>
- Martín-de Castro, G., Díez-Vial, I., & Delgado-Verde, M. (2019). Intellectual capital and the firm: evolution and research trends. *Journal of Intellectual Capital*, 20(4), 555–580. <https://doi.org/10.1108/JIC-12-2018-0221>
- McGurr, P. T., & DeVaney, S. A. (1998). Predicting business failure of retail firms: An analysis using mixed industry models. *Journal of Business Research*, 43(3), 169–176. [https://doi.org/10.1016/S0148-2963\(97\)00222-1](https://doi.org/10.1016/S0148-2963(97)00222-1)
- Migliani, S., Ahmed, K., & Henry, D. (2020). Corporate governance and turnaround: Evidence from Australia. *Australian Journal of Management*, 45(4), 1–30. <https://doi.org/10.1177/0312896220902225>
- Min, J., & Lee, Y. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4), 603–614. <https://doi.org/10.1016/j.eswa.2004.12.008>
- Neophytou, E., Charitou, A., & Charalambous, C. (2004). Predicting corporate failure: Empirical evidence for the UK. *European Accounting Review*, 13(3), 465–497. <https://doi.org/10.1080/0963818042000216811>

- Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, 24(12), 1565–1567. <https://doi.org/10.1038/nbt1206-1565>
- Norwahida, S., Shin, O., & Shaari, M. (2012). Do the board of director's characteristics affect firm performance? Evidence from Malaysian public listed companies. *International Business Research*, 5(9), 120–127. <https://doi.org/10.5539/ibr.v5n9p120>
- Ohlsion, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. <https://doi.org/10.2307/2490395>
- Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52(2), 464–473. <https://doi.org/10.1016/j.dss.2011.10.007>
- Paoloni, M., & Celli, M. (2018). Crisi delle PMI e strumenti di warning. Un test di verifica nel settore manifatturiero. *Management Control*, 2, 85–106. <https://doi.org/10.3280/MACO2018-002005>
- Paoloni, N., Mattei, G., Strologo, D., A., & Celli, M. (2020). The present and future of intellectual capital in the healthcare sector: A systematic literature review. *Journal of Intellectual Capital*, 21(3), 357–379. <https://doi.org/10.1108/JIC-10-2019-0237>
- Piesse, J., & Wood, D. (1992). Issues in assessing MDA models of corporate failure: A research note. *The British Accounting Review*, 24(1), 33–42. [https://doi.org/10.1016/S0890-8389\(05\)80065-4](https://doi.org/10.1016/S0890-8389(05)80065-4)
- Platt, H. D., & Platt, M. B. (1990). Development of a class of stable predictive variables: The case of bankruptcy prediction. *Journal of Business Finance & Accounting*, 17(1), 31–51. <https://doi.org/10.1111/j.1468-5957.1990.tb00548.x>
- Pompe, P. P. M., & Feelders, A. J. (1997). Using machine learning, neural networks, and statistics to predict corporate bankruptcy. *Computer-Aided Civil and Infrastructure Engineering*, 12(4), 267–276. <https://doi.org/10.1111/0885-9507.00062>
- Porter, M. E. (1987). *Il vantaggio competitivo*. Milano: Edizioni Comunità
- Ricci, F., Scafarto, V., Ferri, S., & Tron, A. (2020). Value relevance of digitalization: The moderating role of corporate sustainability. An empirical study of Italian listed companies. *Journal of Cleaner Production*, 276(10), 123282. <https://doi.org/10.1016/j.jclepro.2020.123282>
- Salchenberger, L. M., Cinar, E. M., & Lash, N. A. (1992). Neural networks: A new tool for predicting thrift failures. *Decision Sciences*, 23(4), 899–916. <https://doi.org/10.1111/j.1540-5915.1992.tb00425.x>
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), 210–229. DOI: <https://doi.org/10.1147/rd.33.0210>
- Santana, M., Valle, R., & Galan, J. L. (2017). Turnaround strategies for companies in crisis: Watch out the causes of decline before firing people. *BRQ Business Research Quarterly*, 20, 206–211. <https://doi.org/10.1016/j.brq.2017.01.003>
- Schapiro, R., & Freund, Y. (2012). *Boosting: Foundations and Algorithms* (Cambridge, MA: MIT Press). Retrieved from: [https://doc.lagout.org/science/0\\_Computer%20Science/2\\_Algorithms/Boosting\\_%20Foundations%20and%20Algorithms%20%5BSchapiro%20%26%20Freund%202012-05-18%5D.pdf](https://doc.lagout.org/science/0_Computer%20Science/2_Algorithms/Boosting_%20Foundations%20and%20Algorithms%20%5BSchapiro%20%26%20Freund%202012-05-18%5D.pdf) (last access: 04/30/2022)
- Schiama, G., Lerro, A., & Sanitate, D. (2008). The intellectual capital dimensions of Ducati's turnaround: exploring knowledge assets grounding a change management program. *International Journal of Innovation Management*, 12(2), 161–193. <https://doi.org/10.1142/S1363919608001935>
- Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1), 127–135. <https://doi.org/10.1016/j.eswa.2004.08.009>
- Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1), 101–124. <https://doi.org/10.1086/209665>
- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better Decisions Through Science. *Scientific American*, 283, 82 – 7. Retrieved from: <https://www.scientificamerican.com/article/better-decisions-through-science/> (last access: 04/30/2022)
- Switzer, L., Wang, J., & Tu, Q. (2018). Corporate Governance and Default Risk in Financial Firms over the Post Financial Crisis Period: International Evidence. *Journal of International Financial Markets Institutions and Money*, 52, 196–210. <https://doi.org/10.1016/j.intfin.2017.09.023>
- Tamari, M. (1964). Financial ratio as a means of forecasting bankruptcy. *Economic Review*, 21, 15–45. Retrieved from <https://www.econbiz.de/Record/financial-ratios-as-a-means-of-forecasting-bankruptcy-tamari/10002892255>
- Tian, Y., Shi, Y., & Liu, X. (2012). Recent advances on support vector machines research. *Technological and Economic Development of Economy*, 18(1), 5–33. <https://doi.org/10.3846/20294913.2012.661205>

- Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking and Finance*, 52, 89–100. <https://doi.org/10.1016/j.jbankfin.2014.12.003>
- Tron, A. (2020). *Corporate Financial Distress: Restructuring and Turnaround*. London: Emerald Publishing
- Tron, A., Valenza, G., & Caputo, A. (2018). Corporate crisis management in Italy: execution, monitoring and performance analysis of recovery business and financial plans. *International Journal of Foresight and Innovation Policy*, 13(1), 114–136. <https://doi.org/10.1504/IJFIP.2018.095861>
- Tsai, C. F., Hsu, Y. F., & Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24, 977–984. <https://doi.org/10.1016/j.asoc.2014.08.047>
- Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41(5), 2353–2361. <https://doi.org/10.1016/j.eswa.2013.09.033>
- Ward, T. J. (1994). An empirical study of the incremental predictive ability of Beaver's naive operating flow measure using four state ordinal models of financial distress. *Journal of Business Finance & Accounting*, 21(4), 547–561. <https://doi.org/10.1111/j.1468-5957.1994.tb00335.x>
- Warner, J. B., Watts, R. L., & Wruck, K. H. (1988). Stock prices and top management changes. *Journal of Financial Economics*, 20(2), 461–492
- Westgaard, S., & Van Der Wijst, N. (2001). Default probabilities in a corporate bank portfolio: A logistic model approach. *European Journal of Operational Research*, 135(2), 338–349. [https://doi.org/10.1016/S0377-2217\(01\)00045-5](https://doi.org/10.1016/S0377-2217(01)00045-5)
- Yang, Z. R., Platt, M. B., & Platt, H. D. (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research*, 44(2), 67–74. [https://doi.org/10.1016/S0148-2963\(97\)00242-7](https://doi.org/10.1016/S0148-2963(97)00242-7)
- Yeh, C. C., Chi, D. J., & Lin, Y. R. (2014). Going-concern prediction using hybrid Random Forests and rough set approach. *Information Sciences*, 254(1), 98–110. <https://doi.org/10.1016/j.ins.2013.07.011>
- Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, 40(2), 185–211. [https://doi.org/10.1016/0304-405X\(95\)00844-5](https://doi.org/10.1016/0304-405X(95)00844-5)
- Yu, L., Yue, W., Wang, S., & Lai, K. (2010). Support vector machine based multiagent ensemble learning for credit risk evaluation. *Expert Systems with Applications*, 37(2), 1351–1360. <https://doi.org/10.1016/j.eswa.2009.06.083>
- Zhao, Z., Xu, S., Kang, B. H., Kabir, M. M. J., Liu, Y., & Wasinger, R. (2014). Investigation and improvement of multi-layer perception neural networks for credit scoring. *Expert Systems with Applications*, in press
- Zhou, L., Lai, K. K., & Yen, J. (2014). Bankruptcy prediction using SVM models with a new approach to combine features selection and parameter optimisation. *International Journal of Systems Science*, 45(3), 241–253. <https://doi.org/10.1080/00207721.2012.720293>
- Zingales, L. (2000). In search of new foundations. *Journal of Finance*, 55(4), 1623–1653

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.