Technological Forecasting & Social Change 174 (2022) 121255

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Forecasting credit ratings of decarbonized firms: Comparative assessment of machine learning models

Baojun Yu [a], Changming Li [b,*], Nawazish Mirza [c], Muhammad Umar [d]

[a] School of Management, Jilin University, Changchun, Jilin, China
[b] Engineering Technology R&D Center, Changchun Guanghua University, Changchun, Jilin, China
[c] Excelia Business School, La Rochelle, France
[d] UCP Business School, Faculty of Management Studies, University of Central Punjab, 1-Khayaban-e-Jinnah Road, Johar Town, Lahore, Pakistan

A R T I C L E I N F O

JEL Classifications:
G17
G32
O33
Q43

Keywords:
Carbon neutrality
Low carbon transitions
Machine learning
Credit ratings

A B S T R A C T

Maintaining low carbon energy transitions is a phenomenon that is critical in curtailing greenhouse emissions. However, such shifts usually warrant incremental capital expenditures, which require an uninterrupted access to financing. Credit ratings are an essential consideration of the financing process. In this paper, we assess the ability of various machine learning models, in order to forecast the credit ratings of eco-friendly firms. For this purpose, we have employed a sample of 355 Eurozone firms that are ranked on the basis of the extent of their climate change score by SDP, between the years spanning from 2010 to 2019. The study uses various machine learning methods, and the findings suggest that classification and regression trees have the most precision for the credit rating predictions. Even when the forecasting was constrained to the investment grades, speculative grades, or default categories, the accuracy remained robust. The results also suggest that a random forest ensemble can be used alongside the regression trees in order to predict default or near default ratings. Given that such firms face dynamic risk exposure towards environmental, ecological, and social factors, these results have important implications that can be taken into consideration when assessing the credit risk of pro-ecological firms.

1. Introduction

The emission of greenhouse gasses exuberates environmental as well as socioeconomic consequences. In this regard, (Wu et al., 2021) documented that carbon emission significantly impair the air quality, ultimately leading to health issues for people who come into contact with such an environment. Similarly, these emissions are also believed to have led to an increased level of sovereign risk (Chaudhry et al., 2020), and limit financial development (Caragnano et al., 2020; Umar et al., 2020b). Thus, in order to minimize the impact of greenhouse effects, a sustainable transition to low-carbon energy sources is necessary. Although the awareness of the need for a circular economy has been observed to rise sharply, a lot of work still needs to be undertaken in order to achieve sustainability goals (Tateishi et al., 2020; Umar et al., 2020a). The shift to a zero-carbon regime requires a pro-ecological regulatory environment (Jiang and Ma, 2021; Su et al., 2021b). This essentially also means that for the firms, decarbonisation will necessitate capital investment in renewable energy technologies.

The Capex investment has warranted for a low carbon business transition, and is likely to put the financial flexibility under pressure. Furthermore, the regulatory changes may warrant a complete overhaul of the traditional business models, so as to ensure a carbon-neutral footprint. Consequently, this would then impact the cost of capital, and businesses with a competitive advantage would then be in a better and more positive stand point, in order to foster their transitions. It is noteworthy that the financial flexibility is a function of credit ratings, and firms with higher credit ratings tend to have lower financing costs (Hasan et al., 2021; Su et al., 2021a). Therefore, the availability of external credit ratings can help minimize the carbon emissions effectively by facilitating the required Capex. Forecasting such ratings will help devise a financial strategy.

Previous studies have also used various machine learning models, in order to predict credit ratings. In the same reference (Golbayani et al., 2020) assessed various machine learning tools, and concluded in favor of the tree-based models. They hypothesized that such models also capture the notching, and that their predictions are analogous with the

* Corresponding author.
E-mail addresses: yubaojun@jlu.edu.cn (B. Yu), changming_li0034@163.com (C. Li), elahimn@excelia-group.com (N. Mirza), muhammad.umar@ucp.edu.pk (M. Umar).

https://doi.org/10.1016/j.techfore.2021.121255
Received 18 May 2021; Received in revised form 26 September 2021; Accepted 27 September 2021
Available online 13 October 2021

Table 1
Greenhouse gas emissions - tonnes per capita.

	2014	2015	2016	2017	2018
Austria	10.6	11.0	10.8	10.8	10.8
Belgium	11.4	11.4	11.4	11.2	10.7
Cyprus	16.1	13.9	15.0	16.0	15.3
Estonia	12.8	13.2	13.5	13.3	13.2
Finland	9.4	9.1	8.8	9.2	9.0
France	7.3	7.6	7.4	7.7	7.5
Germany	7.1	7.1	7.1	7.2	6.9
Greece	7.2	7.4	7.4	7.3	7.3
Ireland	10.6	10.7	11.4	11.6	11.3
Italy	5.8	5.8	5.9	6.0	6.3
Latvia	6.9	7.1	7.2	7.4	7.4
Lithuania	21.5	20.4	19.8	20.0	20.3
Luxembourg	7.5	5.9	5.1	5.5	5.5
Malta	11.8	12.2	12.2	12.0	11.6
Netherlands	9.2	9.3	9.4	9.6	9.2
Portugal	6.4	6.9	6.7	7.2	7.0
Slovakia	7.6	7.7	7.8	8.0	8.0
Slovenia	11.1	10.4	10.9	10.4	10.7
Spain	10.8	10.8	10.5	10.2	10.1

Source: European Environment Agency (EEA).
The table presents total national emissions per capita. The emissions include carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and the F-gasses, i.e., hydrofluorocarbons, perfluorocarbons, nitrogen trifluoride (NF_3), and sulfur hexafluoride (SF_6).

Table 2
Sample description and climate change score.

	Leadership		Management		Total
	A	A-	B	B-	
Austria	2	3	4	2	11
Belgium	2	3	7	3	15
Finland	6	10	11	3	30
France	16	22	31	3	72
Germany	15	15	29	1	60
Greece	1	0	1	1	3
Ireland	4	7	9	5	25
Italy	8	13	13	3	37
Luxembourg	0	1	1	0	2
Malta	0	0	1	0	1
Netherlands	10	7	6	2	25
Portugal	4	4	4	0	12
Spain	11	14	17	0	42
Total	79	99	134	23	335

Source: CDP Disclosures.

actual rating disclosures of Fitch, Standard & Poors, and Moody's. At another instance, (Li et al., 2020) evaluated the ability of various machine learning models to predict the issuer level ratings of commercial banks. In this regard they reported the differentiated capacity of these models and suggested that random forests are, on average better predictors of the credit ratings. However though, for higher credit risk, the study favored the classification and regression trees. At another time in the past, (Huang, 2011) analyzed the credit rating forecasting in the context of the subprime mortgage crisis and suggested that machine learning approaches tend to adequately predict the risk profile. The evidence also supports the fact that these models adapt to business cycle conditions for more precise forecasts.

However, it must be noted that these studies focus on the overall industrial firms and offer limited perspective surrounding carbon-neutral businesses. Given the significant differences, we also believe that it is not viable to standardize the conventional carbon emitters forecasting models according to carbon-neutral firms. To the best of our knowledge, no prior study has thoroughly assessed the machine learning approaches that can be taken into consideration while forecasting the credit ratings for low carbon businesses. We have therefore attempted to address this void in the literature that can aid in identifying the factors,

Table 3
Credit rating scales.

		S&P	Moody's	Fitch	Scale	Category
Max quality	Investment	AAA	Aaa	AAA	17	CR1
High quality	grade	AA+	Aa1	AA+	16	CR2
		AA	Aa2	AA	15	
		AA-	Aa3	AA-	14	
Strong		A+	A1	A+	13	CR3
repayment		A	A2	A	12	
profile		A-	A3	A-	11	
Adequate		BBB+	Baa1	BBB+	10	CR4
repayment		BBB	Baa2	BBB	9	
profile		BBB-	Baa3	BBB-	8	
Expected to	Speculative	BB+	Ba1	BB+	7	CR5
settle	grade	BB	Ba2	BB	6	
		BB-	Ba3	BB-	5	
High default		B+	B1	B+	4	CR6
risk		B	B2	B	3	
		B-	B3	B-	2	
Very High		CCC+	Caa1	CCC+	1	CR7
Default		CCC	Caa2	CCC		
Risk		CCC-	Caa3	CCC-		
Near default		CC	Ca	CC		
				C		
Default	Default	SD	C	DDD		
		D		DD		
				D		

as well as a robust model to predict creditworthiness.

Machine learning helps in inferring statistical relationships and predictions after training the specified models that make use of extensive datasets. Here, the concept of *learning* optimizes the prediction process by limiting the reliance on exogenous assumptions. Given the robustness of these techniques, machine learning approaches have been extensively resorted to in the context of asset pricing (Gan et al., 2020), default risk (Tang et al., 2019), bankruptcy (Barboza et al., 2017), and other applications (Buehler et al., 2019), (Weigand, 2019). In our estimations, it would be reasonable to differentiate and adopt machine learning approaches for the credit rating forecast of carbon-sensitive firms. Unlike their counterparts, they are focused on the environmental, economic, as well as social sustainability factors. When looking at this in detail, we can observe that the environmental consideration will require them to sustain ecological integrity (Taghizadeh-Hesary et al., 2021; Umar et al., 2021a; Wang et al., 2021) and hence balance the consumption of natural resources (Bibi et al., 2021; Ji et al., 2021). Moreover, the economic concerns tend to warrant pursuing financial goals, without negatively impacting the environment (Umar et al., 2021b), while providing adequate yields for the stakeholders (Hasnaoui et al., 2021; Naqvi et al., 2021). These factors would make the risk exposures more dynamic as compared to those in a conventional firm and require the forecasting methods to adapt to the frequently changing vulnerabilities.

In this paper, we have resorted to the use of macro and micro-level data of 355 Eurozone firms, and have also evaluated the ability of four machine learning models in order to forecast the credit ratings. The choice of the Eurozone firms was very straightforward, given the stringent stance of decarbonisation by the member states. The European green deal has pledged to reduce the emissions by 55% by the next decade comes around. This is higher compared to an earlier target of reduction by 40% in greenhouse emissions. In addition to this, the share of renewable energy is targeted (by 2030) to be at 32%, with a similar percentage improvement in the energy efficiency as well. At present, the bulk of European emissions are contributed by power, transportation, building, industry, and agriculture. Also, the pace of decarbonisation will be driven by technology and scaling up the supply chain systems across the member states.

We have also benefitted from the CDP database, and have included firms that have been ranked at *Leadership* and *Management* levels, based on their climate change scores. The models have been trained using

Fig. 1. Variable ranking - classification and regression trees (CRT).

quarterly data from the years pertaining to 2010 to 2019. Our results indicate that classification and regression trees are the best predictors of credit ratings for low carbon firms. The findings have remained robust after limiting the predictions to investment grade, speculative grade, and default rating classes.

The rest of the paper is organized as follows. Section 2 has provided a background of the greenhouse gas emissions and CDP scores. Then, a brief introduction of the concept of credit ratings has been presented in Section 3. Moving on, Section 4 outlines the empirical strategy undertaken by this research, while Section 5 presents the machine learning models. Section 6 then introduces the rating criteria, and the results have been discussed in Section 7. Finally, Section 8 concludes the paper.

2. Greenhouse emissions and CDP climate change score

In continuation of the agendas discussed at the Paris climate conference, the European Union (EU) has aimed for a 55% reduction in greenhouse emissions by the year 2030. The final target here is to achieve zero net emissions by the year 2050. While the objective is enthusiastic, the EU can successfully accomplish it by focusing on sustainable energy sources (Akdag & Yıldırım, 2020). Although the greenhouse emissions are still high, they have come a long way from the initially observed levels in 1990. Following this context, Table 1 has presented data regarding the greenhouse gas emissions (tonnes per capita) for 19 Eurozone countries. As observable, the overall trend has been declining,

Fig. 1. (*continued*).

which is encouraging from the environmental perspective. (Paraschiv & Paraschiv, 2020) attributed this decline across the EU to the reduction in CO_2, owing to the adoption of pro-ecological fuels. It is noteworthy that Lithuania contributes towards the maximum emissions at 20.3 tonnes per capita, in the most recent reported period, followed by Cyprus (15.3) and Estonia (13.2). Luxembourg, however, has reported the minimum emissions (5.5), followed by Italy (6.3).

The corporate sector has taken up a central role in order to further reduce these emissions, and achieve the targets that have been set for 2030 (Tao et al., 2021). The individual business emissions are planned to contribute to the national tally, which will lead to the global total. Therefore, the reduction in global greenhouse gasses must be initiated through systemic changes at the firm's level. The Carbon Disclosure Project (CDP) provides a scoring-based methodology in order to rank business progress when it comes to supporting climate change and adopting to low carbon corporate models.[1] The ranks, based on annual disclosures, and related to environmental management, are categorized into five types. The A and A- firms depict the *leadership* level, while B and B- businesses qualify as *management* level. These two categories represent the significant efforts that are made to promote sustainable and pro-ecological business models. Moving on, the categories of C, C- and D, D- respectively, represent the *awareness* and *disclosure* levels, and require extensive transitions to renewable, low carbon, and less polluting fuel

[1] The details of scoring methodology is accessible at https://www.cdp.net

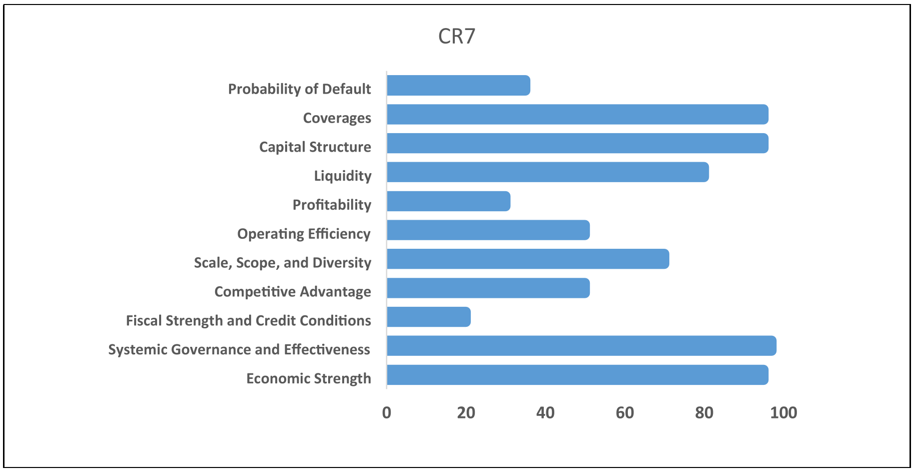

Fig. 1. (*continued*).

sources. Finally, the F category represents firms that *fail* to disclose sufficient information, and vis-à-vis to assess the climate impact. The CDP ranking is widely accepted as a benchmark to reflect and review the corporate awareness of environmental challenges, and the best practices for risk mitigation. Many academic studies have resorted to the use of the CDP data, for instance, more recently (Höck et al., 2020) and (Kouloukoui et al., 2019), assessed and analysed the firm and country-level issues related to climate change. For this research, our focus remains intact on the credit rating predictions of low greenhouse emission firms, and therefore, we have taken into consideration all the nonfinancial Eurozone firms that are ranked as A, A-, B, and B-,by CDP. The sample selection procedure has been discussed in the following section.

3. The concept of credit ratings

The credit ratings reflect an independent opinion regarding the financial flexibility and creditworthiness of an obligor. This opinion considers many factors, including the obligor's exposure to risk factors such as credit, market, liquidity, business, etc. The relevance of these risk factors is dependent on the assessed entity. For example, in the case of a banking firm, credit risk will tend to carry a higher weight, while for a manufacturing firm, business risk could be more prominent.

The ratings are usually in the form of notations which could help in differentiating the exposures from very high (investment grade) to low credit quality (speculative and default) (Pertaia et al., 2021). Moreover, the ratings can be of instruments (bonds etc.), or for the entities as a whole as well. The purpose of credit ratings is to provide a uniform and transparent system for the market participants in order to understand the repayment capacity. Consequently, they ultimately help to facilitate an efficient debt space (Szetela et al., 2019).

The rating process is very dynamic, and any material macro or micro event can trigger a rating change. Also, the credit ratings are conceptually allocated on *Point in Time* (PIT) or *Through the Cycle* (TTC) basis. The PIT approach refers to a rating allocation process in which the rating agency is likely to consider both cyclical and permanent default factors, in order to assess the current repayment capacity of an obligor. Rating agencies are expected to differentiate between transitory trends and permanent factors affecting the default risk in the TTC approach. Theoretically, this will imply that ratings will be more stable under a TTC approach.

The ratings have multiple benefits for various stakeholders. For the entities, they help in optimizing the financing cost, and provide better access to raise capital. Sometimes, they also prove to be effective when negotiating the tenure, while helping in diversifying the funding comparative credit quality, so as to price the debt and visualize any credit enhancements. The capital charge across the banking sector also benefits from the availability of a public rating. Finally, from an investor's viewpoint, the ratings provide the basis for investments, by reflecting on their attractiveness across asset classes (Alanis, 2020).

4. Data and empirical strategy

This paper aims to assess the ability of various machine learning models in order to predict the credit ratings of environment-friendly firms in the Eurozone. The choice of Eurozone stems from two points of consideration. First, the European Union has been at the forefront of the transition to sustainable business models in order to reduce greenhouse emissions. Second, as noted by (Umar et al., 2021) and (Mirza et al., 2020), the homogeneity of currency is vital for cross-country comparative assessment of credit capacity. In the EU, 19 countries have adopted the Euro, and CDP has environment-related ranks for 13 of the member states. Therefore, our sample comprises of companies from these 13 countries. The selected firms are CDP A and B category firms with publicly available credit ratings allocated by a recognized rating agency.[2] Based on this, our final sample consists of 335 nonfinancial firms. Out of this total, 79 have CDP A classification, 99 have A-, and 134 are ranked B, while 23 are categorized as B-. Among the countries, France has the most significant contribution (72), followed by Germany (60), while the minimum representation is from Malta, with only one company in the Leadership and Management category. The country-wise sample distribution also reflects on, to some extent, the pro-environment efforts, resulting in more companies represented in the CDP database. Table 2 presents the sample distribution.

In order to have access to sufficient data for learning and forecasting, we have considered a sample period of ten years, spanning from January 2010 to December 2019. We have divided this into the *learning* and *prediction* (2017 to 2019) periods. Moreover, the quarterly data for the macro and micro variables have been downloaded from Thomson Reuters Eikon and Bank scope. In addition to this, the credit ratings have been extracted from the websites of the relevant credit rating agencies. By making use of the typical conventions of rating notations, we have then divided the ratings into three categories: investment grade, speculative grade, and default. These have been classified from CR1 (Maximum) to CR7 (Default). Finally, we have assigned the numerical values ranging from 17 (AAA) to 1 (Default), so as to include the notches in each category. The rating scales have been presented in Table 3.

The credit rating data, and the financial and macroeconomic inputs

[2] We use authorization by European Securities and Markets Authority as basis for recognition.

Fig. 2. Variable ranking - artificial neural network (ANN).

have been introduced in machine learning models using the following procedure.

a The data from 2010 to 2019 has been sorted and organized as per the rating categories (investment, speculative, and default).
b The algorithms for various machine learning models have been used to train the data.
c From the trained data, the most robust specification has been extracted. This has been based on the relevant iterations in order to ensure that a target precision is achieved.
d The best specification model will be used to forecast the final credit rating.
e The predicted credit rating is compared with the actual ratings, so as to determine the accuracy.

To measure the forecast accuracy, we have employed the F1- score, specificity, and accuracy. The predictive instances for the machine learning models have been ordered as;

True Positive (TP): Low default risk forecasted as higher ratings.
False Negative (FN): Low default risk forecasted as lower ratings.
True Negative (TN): High default risk forecasted as lower ratings.
False Positive (FP): High default risk forecasted as higher ratings.

Fig. 2. (*continued*).

The *F1 score* is an estimate of the accuracy of a forecast. This can be calculated as

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}$$

Here, the precisions refer to the ratio of positive forecasts to the total forecasts of positive class values [TP / (TP + FP)]. The recall is the proportion of positive forecasts, to true positives and false negatives [TP / (TP + FN), and this phenomenon is commonly known as sensitivity. A greater F1 score shows a higher accuracy of the model.

The *specificity* is the ratio of negative instances to the total forecasts of negative cases. We have calculated this as;

$$S = \frac{TN}{TN + FN}$$

Accuracy is the estimate of accurate forecasts. It reflects on the model's ability to produce true positives and negatives, and show them as a ratio of total forecasts. In our study, this will mean that the model can generate investment-grade ratings for good credit quality and vice versa. The accuracy is determined as follows.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

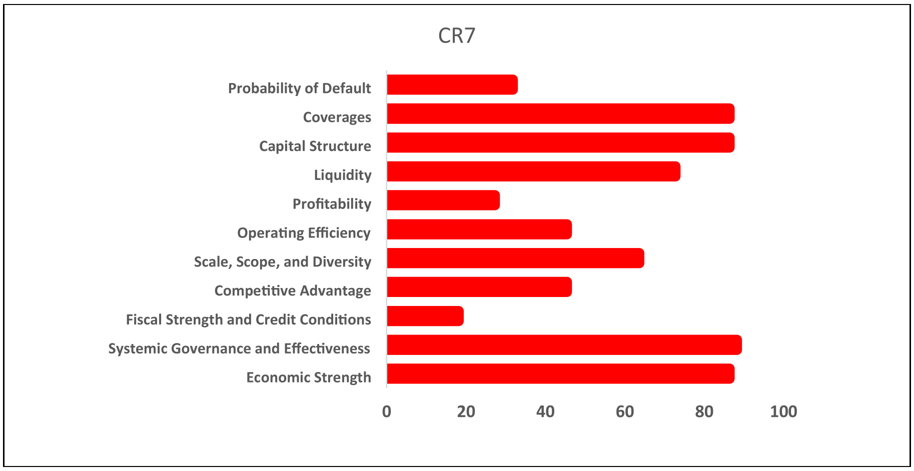

Fig. 2. (*continued*).

5. Machine learning models

The forecasting methods can be classified into two types. The first type includes conventional parametric and nonparametric models that are based on an exhaustive set of assumptions. The most commonly used methods in credit risk forecasting are the Multinomial Logistic Regression, and the Linear and Regularized Discriminant Analysis. However, studies like (Wallis et al., 2019), (Athey and Imbens, 2019), (Mullainathan and Spiess, 2017), (Varian, 2014), (Lee et al., 2014) and (Wu et al., 2014) have criticized these approaches due to their limited ability to forecast a dynamic process like credit ratings. They argued that machine learning methods have higher accuracy than the traditional parametric and nonparametric approaches.

The A higher level of accuracy stems from extensive data mining, fewer underlying assumptions, the ability to *learn* and improve, and automation factors central to the machine learning approach. These models have proven their utility in the domain of finance (Gan et al., 2020) and, specifically, for default risk, as noted by (Pang et al., 2021), (Li et al., 2020), and (Zhu et al., 2019). For this study, we have employed four models based on the machine learning approach. Our choice of these methods is based on their robust prediction performance, as documented in the extant literature by (Bas et al., 2021), (Vrontis et al., 2021), (Lee, 2021), (Li et al., 2020), and (Blazquez and Domenech, 2018).[3] These models are explained below.

5.1. Classification and regression trees (CRT)

The CRT approach helps to forecast dependent variables by employing several labelled variables (Krzywinski & Altman, 2017). This technique is especially and specifically helpful if the predicted variable is categorical or continuous in nature. As credit ratings are both categorical and continuous, therefore the CRT approach is well suited for our study. The other benefits of this approach include not assuming anything about the dataset, and no standardization requirement. This means that the model is suitable for nonlinear datasets as well, and the prediction can be based on an "if-else" algorithm.

Given that our sample is very particular, the benefits of the CRT approach make our estimates extremely robust. The CRT model estimation involves three steps. This includes constructing a maximum tree, followed by optimizing the size and ends after the data classification. The process is then iterated until the optimal objective has been achieved, which is then presented as a final rating in our study. During the iteration, multiple rounds (or nodes) may take place. These nodes are of three different types. The first type is a node without a parent, but can result in two sub-nodes. The second type is a note with a parent that results in two sub-nodes. Finally, the third node is the one with a parent, but has no sub-nodes.

5.2. Artificial neural networks

The artificial neural networks can help in computing a final variable from an array of independent variables. As noted by (Angelini et al., 2008) and (Bahrammirzaee, 2010), neural networks have demonstrated the ability to predict the credit worthiness of counterparties. The artificial neural networks comprise of an introductory layer, various hidden layers, and an output layer. The nodes interconnect these layers. In the model, the initial layer considers all the independent factors, and employing a nonlinear specification function; the output will be passed on to the following layer. This will then be repeated within the hidden layers till the final objective has been achieved. Machine learning will ensure that each layer learns from the experience and leads to the optimal output by employing appropriate weights during the process. The weight decay function of neural networks ensures that an extra layer is eliminated, as soon as its significance diminishes or new information is available (Tkáč & Verner, 2016). The functional form of the node will be as follows;

$$\tau = f(b + xw) = f\left(b + \sum_{i=1}^{N} x_i w_i\right)$$

And $x \in \vartheta_{1\times n},\ w \in \vartheta_{n\times 1},\ b \in d_{1\times 1},\ z \in d_{1\times 1}$

With b reflecting the bias of every node with n number of variables input.

In order to train the model, we have initiated a random allocation of the weights to all the inputs in the first node. These will be passed forward to estimate the output at each node, and finally the objective. The objective value is then compared with the actual in order to measure the within-sample error. The error is passed back to each node, and the relevant contribution is then calibrated and adjusted. The adjusted weights are typically used to reiterate the process till the estimation error coverages to a value of 0.0001. For our assessment, we have used three hidden layers, with a decay factor of 0.1.

[3] We have also resorted to the use of the multiple conventional approaches, but the results are not significant and their relevance is much lower than the ML models used. Therefore, in the interest of space, and to keep the discussion focused on ML approaches, we have not reported them. These additional statistics will be made available on request.

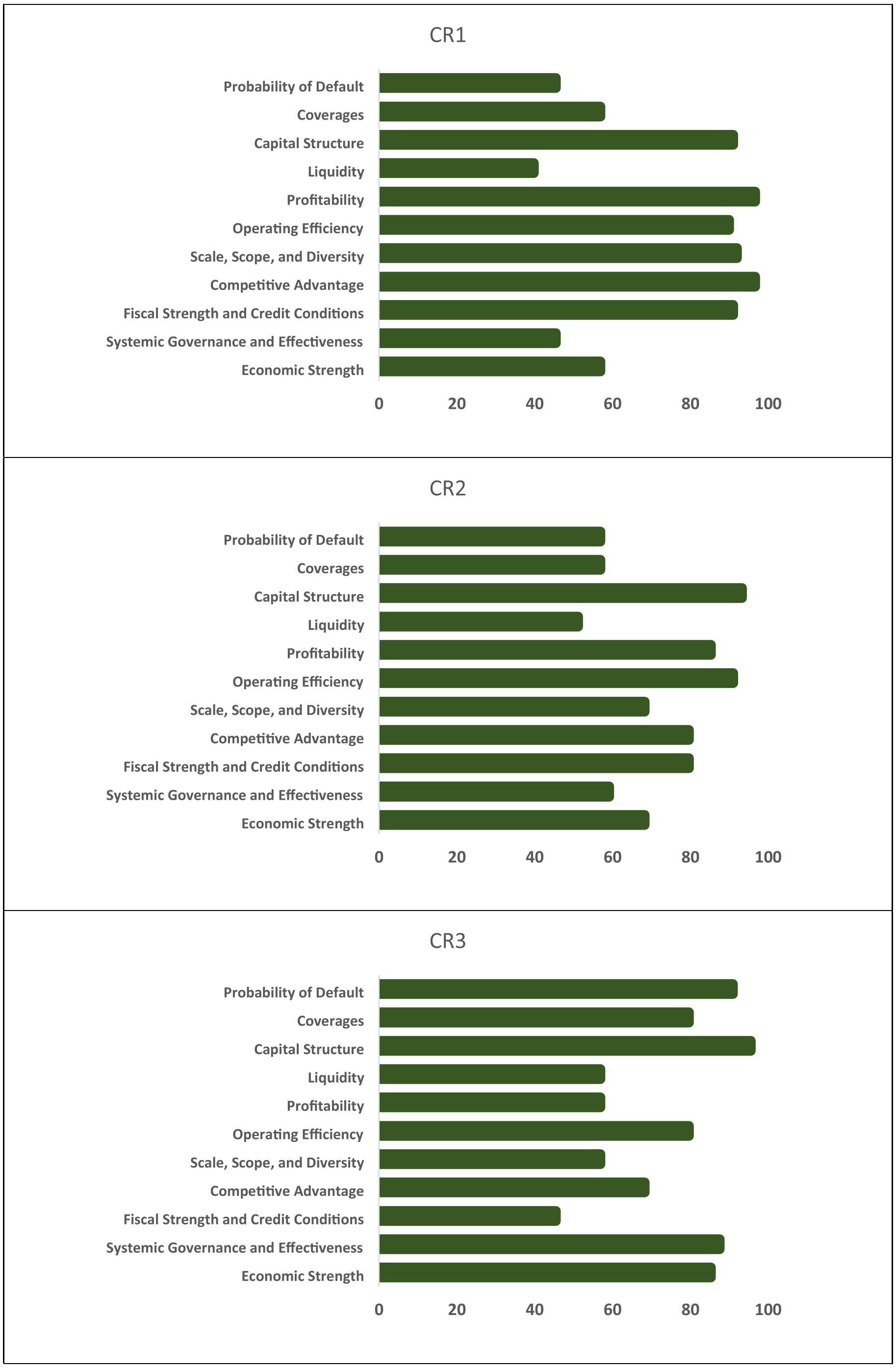

Fig. 3. Variable ranking – random forest ensemble (RFE).

5.3. *Random forests ensemble*

Many studies such as (Tang et al., 2019) and (Malekipirbazari & Aksakalli, 2015) have documented the precision of random forests, in order to duly predict the obligor's strength. The random forest includes various decision trees, which collectively form an ensemble. Each tree in the forest will produce a prediction, and the most precise one of these will qualify to be the model output. The rationale behind this measure is that the models with low correlations will lead to collective predictions. This would also help in eliminating the individual noise. Therefore, even if a few decision trees are subject to errors, the precision of the others will help in getting a more robust output. In order to achieve a low correlation, two specifications of the random forest are beneficial. Primarily, each tree is allowed to be created by generating it with a replacement random sample. The decision trees have been trained from this sample. The second specification relates to feature randomness. This allows us to consider all the variables, and while the nodes are split, the input that will produce maximum separation will be selected. Hence, the training is randomized by allowing the selection of variables from a unique set.

5.4. *Support vector machine*

The support vector machine is a predictive model that employs

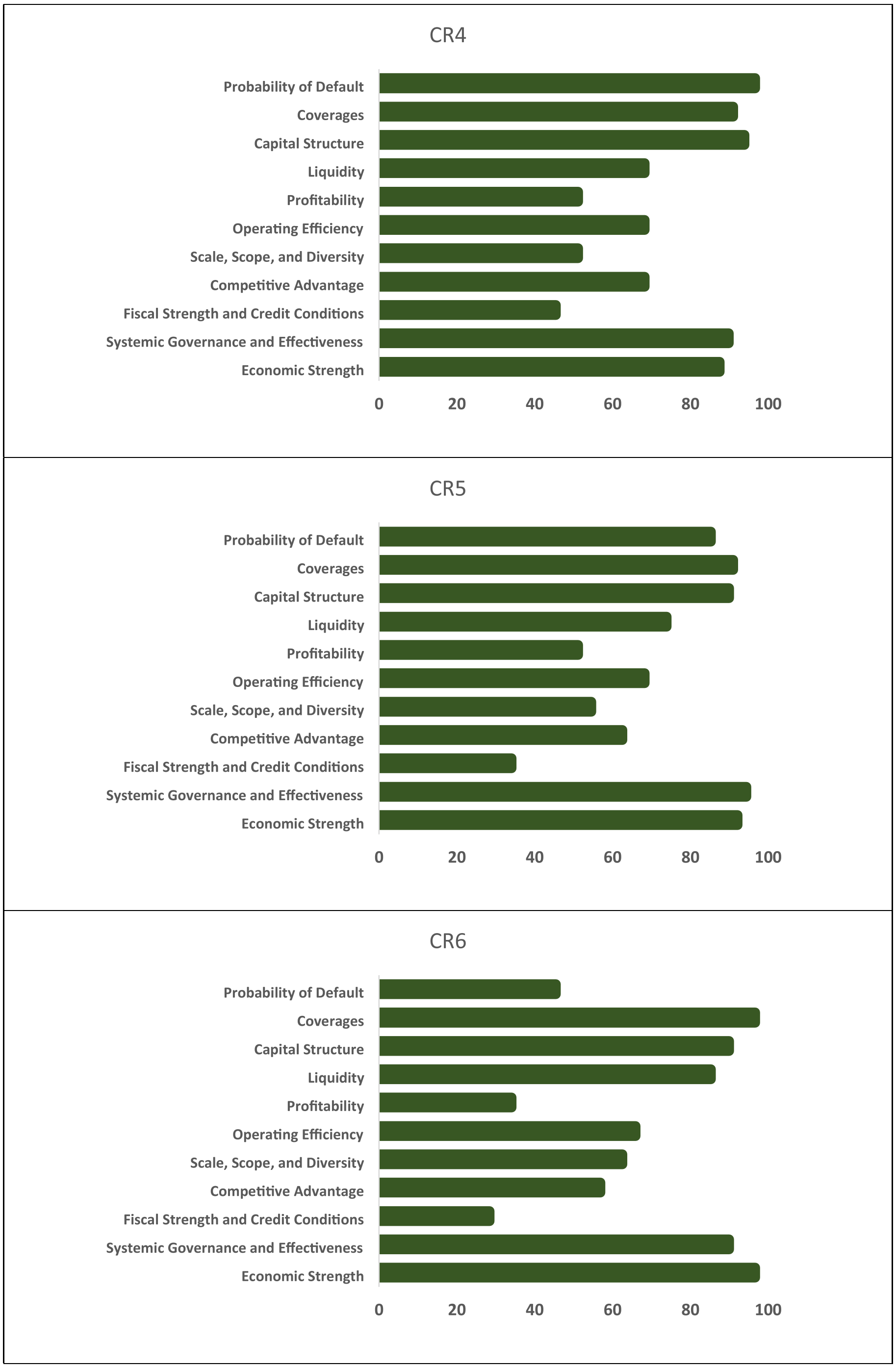

Fig. 3. (*continued*).

classifications (Guenther & Schonlau, 2016). While training the sample, a limit is imposed to distribute similar points across different classes. Thereafter, the sample points are evaluated if they lie within the boundary. On establishing the boundary, the core points are extracted, while the training data becomes redundant. The core points are referred to as the support vector. The prime benefit of SVM is its employability for separable and inseparable linear data (Tiwari, 2017). In the case of separable data, the established boundary will segregate the data with a maximum margin. However, for more complex inseparable data x, SVM transform input into an extreme dimension H, in a way that $x \in M^I \rightarrow \phi(x) \in M^H$along with a kernel function ϕ(x), in order to create the limits. Given the nature of credit rating notations, we have followed a multiclass approach. In this, we have separated each class, and the model has been trained to distinguish between the types in a clear manner. This is important because, in credit ratings, the transition across notations reflects significant repayment capacity differences. This is also true for the notching. The test phase of class A, with some pattern B is thus determined as follows;

$A = \begin{cases} n, & \text{if} \quad d_n(B) + t_l > 0 \\ 0, & \phantom{\text{if}} \quad d_n(B) + t_l \leq 0 \end{cases}$ with $d_n(B) = \max\{d_i(B)\}_{i=1}^{N_l}$, $d_i(B)$ is the space between B, and SVM limit of class *i*, and t_l represents the origin of the classification.

6. Credit rating criteria

The credit ratings are a function of various macroeconomic and firm-

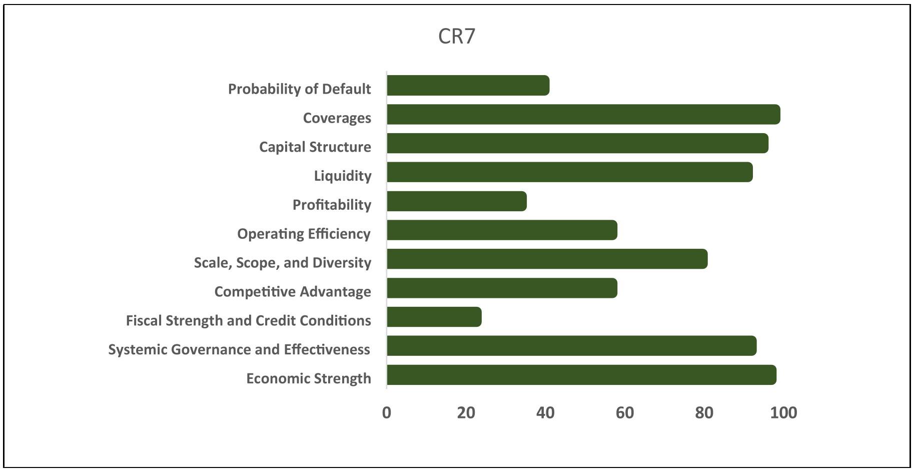

Fig. 3. (*continued*).

specific factors that evolve over the life cycle (Blomkvist et al., 2021). This top-down approach is the principle that forms the basis of the corporate rating criteria of most credit rating agencies. We have thus followed a similar style for this study, and have explained the country, business, and financial risks below, and their measurements for credit rating predictions.

6.1. Country risk profile

We have considered multiple macroeconomic factors that are central to corporate performance and repayment ability. These include:

6.1.1. Economic strength

Economic strength is critical for the robust performance of the corporate sector. This strength stems from the growth and scale of the economy (Öğüt et al., 2012). In order to quantify economic growth, we have taken into consideration the GDP growth rate, volatility of the GDP growth, and the level of diversification.

6.1.2. Systemic governance and effectiveness

The level of systemic governance and effectiveness would impact the country's risk, increasing the domestic institutions' default risk. Systemic governance and efficiency emanate from the transparency, effectiveness of legal structure, continuity of policies, and capacity of the regulators to intervene (Kanagaretnam et al., 2014). We have measured this factor by using three widely used indices. These include the World Bank Government Effectiveness Index, the World Bank Rule of Law Index, and the World Bank Control of Corruption Index.

6.1.3. Fiscal strength and credit conditions

The level of public finances and state debt influence a country's susceptibility to risk. This will also impact the operating environment, and lead to varying credit conditions. To quantify fiscal strength, we have taken into consideration the government debt to GDP, as our key metric. Moreover, for credit conditions, we have employed growth in the private sector credit relative to GDP.

6.2. Business risk

There is a significant impact of business cycles on the default risk (Fève et al., 2021), and therefore the business risk is an essential criterion for corporate ratings. Thus, we have estimated the business risk from the following four factors:

6.2.1. Competitive advantage

Competitive advantage refers to the strategic positioning that results in customer attractiveness towards products and services (Lassala et al., 2021). This is an assessment of the fragility and sustainability of the business model that may stem from product differentiation, positioning, brand reputation, technological advancements, etc. We have used a proxy of competitive advantage through market share and revenue growth.

6.2.2. Scale, scope and diversity

The scale, scope and diversity are a measure of concentration or diversity of the products. The factors that contribute include domestic and geographic diversity, volumes, and maturity of the products and services. Therefore, we have assessed the scale, scope and diversity through the volatility of sales volume and the Hirschman Herfindahl index (HHI) (Afzal and Mirza, 2012).

6.2.3. Operating efficiency

Operational efficiency relates to the quality and flexibility of a firm's capital investments, working capital, cost structures, and expense management. Businesses with optimal cost structures and controls can benefit from high capacity utilization. Similarly, the optimality in the working capital can reduce the cash cycles, leading to a higher repayment capacity (Zeidan and Shapir, 2017). To quantify the operating efficiency of a sample firm, we have used the operating leverage and cash conversion cycle.

6.2.4. Profitability

Profitability is vital to evaluate a firm's long-term sustainability. The assessment has two sub-components, including the level and volatility of profitability. The level of profitability is estimated by taking into account the return on capital employed (ROCE), and the EBITDA margin. The volatility of profitability is measured as the standard deviation of the EBITDA margin across the years (Mirza et al., 2020).

6.3. Financial risk

The financial risk profile directly impacts the repayment capacity of an obligor. It includes assessing sources and uses of financing, the mix of debt and equity, and the cash flow coverage. We have estimated the financial risk profile from the following factors:

6.3.1. Liquidity

Liquidity refers to the ability of a firm to satisfy short-term obligations. The focus of liquidity is on the uses and sources of cash, and stress factors that can impair the short-term cushion. We have measured liquidity as a function of the current ratio, and the operating cash flows to current liabilities (Achim et al., 2021).

Fig. 4. Variable ranking - support vector machine (SVM).

6.3.2. *Capital structure*

The assessment of capital structure can help understand a firm's reliance on external financing (Attaoui et al., 2021). Additionally, it also helps evaluate the impact of credit on leverage, and this consequently reflects on the financial flexibility. We have estimated the capital structure through debt to equity ratio, and the proportion of sponsors' debt in total debt.

6.3.3. *Coverages*

The coverages indicate a firm's cash flow sufficiency to settle the financial liabilities. The cash flow sufficiency is also vital to withstand economic cycles and support growth (Chen et al., 2019). Therefore, we have used OCF to interest, EBITDA to interest, and debt payback to measure the coverage.

6.3.4. *Probability of default*

The probability of default is an ex-ante estimation of distance from bankruptcy. It is based on the market value of assets, a default point, and the standard deviation of assets. In order to estimate the probability of default, we have employed the iterative approach of (Umar et al., 2021c) and (Nawazish et al., 2013).

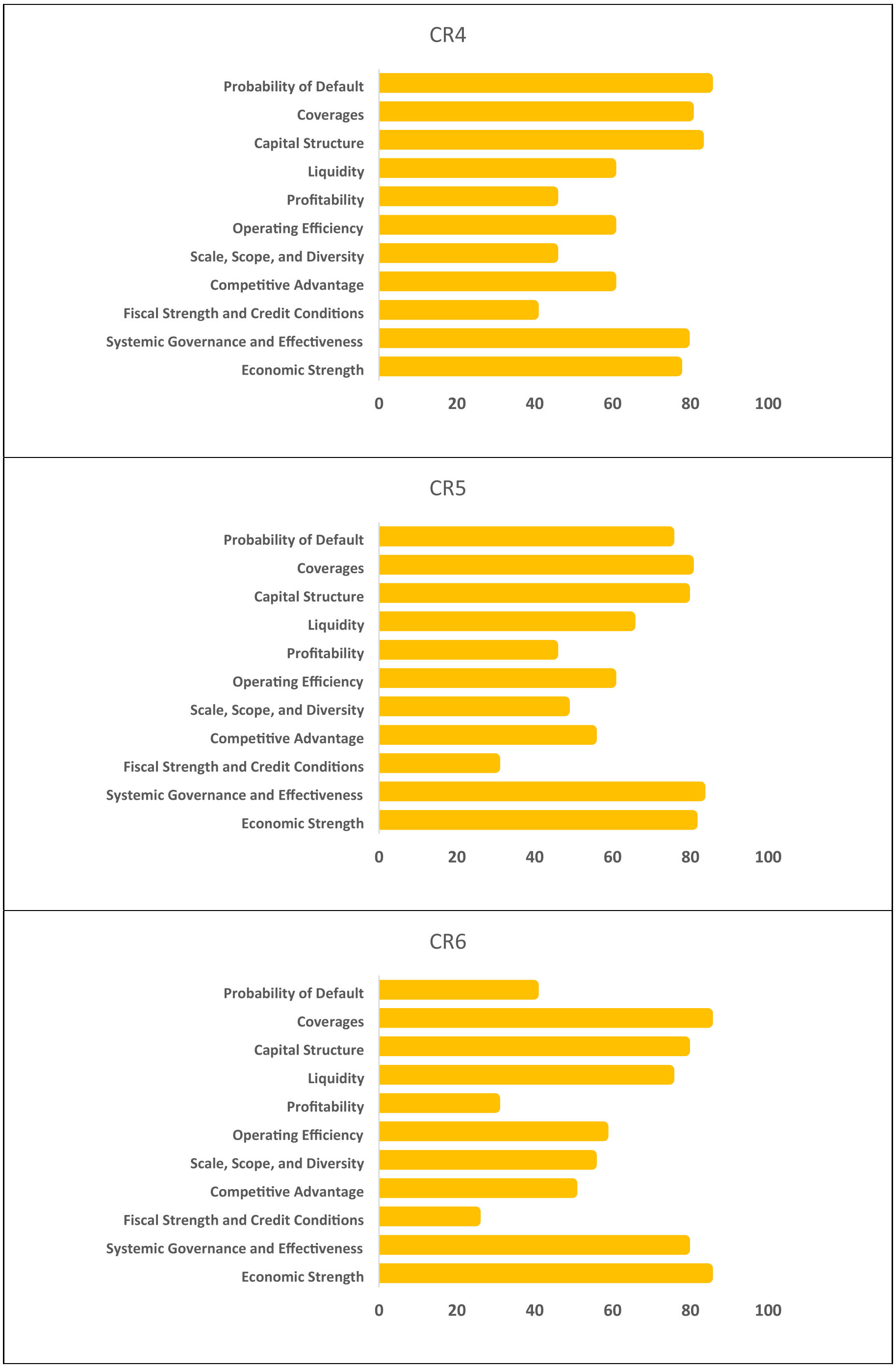

Fig. 4. (*continued*).

7. Results and discussion

We begin by presenting the variable ranking under various machine learning models. These ranks are an outcome of the data training for the four specifications employed in this study. For each of the four models, we have segregated the variable importance according to the different rating classes. The segregation is pertinent as the relevance of the variables would not be the same for investment and speculative grades (Driss et al., 2021). The variable ranking for classification and regression trees is presented in Fig. 1. For CR1 and CR2 ratings, we have observed a greater relevance of country risk profile and business risk than financial risk, except for capital structure. As we transit to slightly lower ratings (CR3 to CR5), the importance of financial risk increases continuously. Finally, the financial risk is hugely significant for high and very high default ratings (CR6 and CR7). While the capital structure is always vital, we have also observed an increase in the relevance of coverages. This is understandable because, as a firm's cash flows deteriorate, the repayment capacity will be impaired.

The variable ranking for artificial neural networks is presented in Fig. 2. Similar to CRT, the capital structure tended to remain highly significant for all ratings. We also observed a higher relevance of economic strength and systematic governance for all the ratings from CR2 to CR. The importance of capital structure (even for the best credit quality), economic strength, and systemic governance is in line with the findings of (Gopalakrishnan & Mohapatra, 2020), who suggested a strong link between debt structure, economic uncertainty, and default

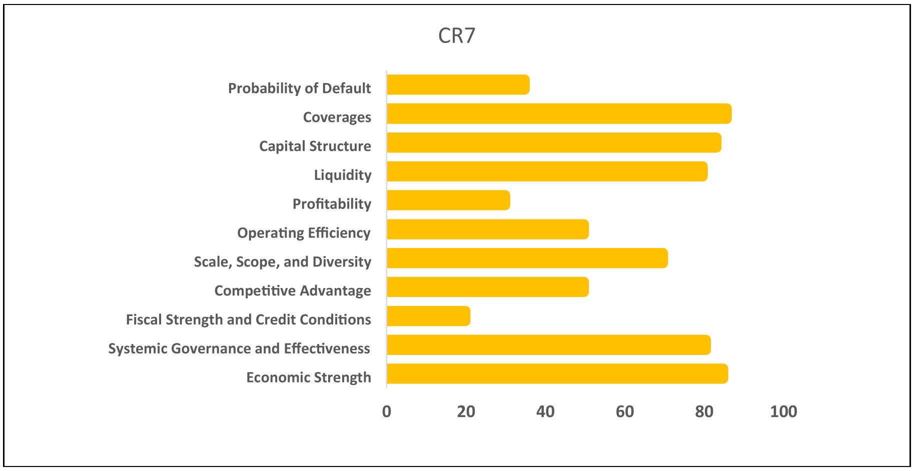

Fig. 4. (*continued*).

Table 4
Comparative model accuracy - for all ratings (CR1 to CR7).

Model	Rank	F1 Score	Specificity	Accuracy
CRT	1	0.89	0.93	0.95
ANN	3	0.65	0.71	0.72
RFE	2	0.81	0.85	0.89
SVM	4	0.41	0.52	0.54

Table 5
Comparative model accuracy - across rating classes.

Panel A	Investment grade (CR1 to CR4)			
Model	Rank	F1 Score	Specificity	Accuracy
CRT	1	0.91	0.94	0.96
ANN	4	0.58	0.60	0.69
RFE	3	0.71	0.73	0.76
SVM	2	0.85	0.88	0.91
Panel B	Speculative grade (CR5 and CR6)			
Model	Rank	F1 Score	Specificity	Accuracy
CRT	1	0.93	0.93	0.97
ANN	3	0.72	0.74	0.77
RFE	4	0.64	0.66	0.67
SVM	2	0.84	0.89	0.91
Panel C	Default (CR7)			
Model	Rank	F1 Score	Specificity	Accuracy
CRT	1	0.91	0.92	0.94
ANN	4	0.68	0.41	0.55
RFE	2	0.81	0.83	0.83
SVM	3	0.75	0.77	0.79

risk. The coverages continue to remain critical for firms with lower credit quality.

Fig. 3 highlights the importance of our selected variables vis-à-vis the random forests ensemble (RFE). While the overall rankings are similar to CRT and ANN, an interesting factor is the firm's liquidity that became increasingly important in the lower rating class. Although liquidity is essentially a short-term phenomenon, consistent liquidity issues in firms with lower credit quality can trigger a default (Zhang et al., 2020). We observe the probability of default to be the most significant for CR3 and CR4 firms, primarily because they are borderline between investment and speculative grades.

The variable rankings for our last specification that support vector machine (SVM) are shown in Fig. 4. As observed before, the capital structure remained significant for all the rating classes. For higher rating classes (CR1 and CR2), the impact of liquidity on an eventual rating remained lower. The profitability and operating efficiency become relevant for higher ratings, but the effect of profitability diminishes as we transit to lower ratings. This is plausible because as a firm moves closer to default and the year-on-year profitability is insufficient to support the credit profile. The recourse at that stage stems from the economic system and the cash injections from the sponsors. Therefore, it is not surprising to see a greater relevance of country risk profile and coverages for lower credit quality firms.

Once the models are trained, and ranking is established, we have proceeded with the predictions, and have reported the comparative precision of the four models. Table 4 presents the statistics related to forecasting for the consolidated output of credit ratings. Across the four specifications, we have observed CRT to be the best model, with the highest F1 score, specificity, and accuracy. The second-best model appears to be the RFE, followed by ANN and SVM. We believe that the ability of CRT to predict the ratings of firms with lower emissions stems from the fact that we do not need the data to be linear or parametric. Furthermore, as the output (credit ratings) is a class-based continuous variable, CRT appears to be an appropriate fit. As for RFE, the ability to predict credit ratings emanates from the randomness and bagging of the algorithm.

As a measure for robustness, we have also presented the ranking of these models across various rating classes in Table 5. The CRT models remain at the top for investment grade (CR1 to CR4), speculative-grade (CR5 and CR6), and the default category (CR7). Therefore, we can conclude that CRT is the optimal machine learning approach for environmentally friendly firms to forecast credit quality. However, unlike CRT, RFE does not appear to be a robust predictor across all rating classes. It remains a good predictor for the default ratings but gives in against SVM and ANN for speculative grades. Similarly, it has a lower forecasting ability than SVM for investment-grade entities. Therefore, for a more restricted output (ratings for a given class), our results suggest that besides the CRT approach, the SVM and ANN take preference over RFE. As mentioned earlier, this is possible because the RFE models work best with randomness, and with constrained forecasts, the predictability deteriorates.

8. Conclusion and recommendations

The transition to low carbon business models by investing in renewable energy technologies is imperative for ecological well-being. The adoption of sustainable business prototypes will require substantial capital investments. Furthermore, the evolution of the regulatory environment for a zero carbon footprint may need an overhaul of the business drivers. Consequently, there will be an impact on the financial flexibility, and typically the firms with a better credit standing will have a competitive advantage in achieving a carbon-neutral status. The credit ratings reflect the creditworthiness, and this study compares the ability of various machine learning models to predict the ratings of low carbon firms.

The results indicate interesting differences in the relevance of the rating factors across various rating classes. The investment-grade ratings associate a higher weight to the country's risk profile. In comparison, the firms in speculative-grade categories are more reliant on making their decisions based on the business risks involved. Finally, the default or near default ratings are more sensitive to the financial risk profile. In the machine learning specifications, classification and regression trees demonstrate the maximum precision for predicting the credit ratings. This was consistent for unconstrained forecasting, and across the rating classes.

These findings are innovative regarding carbon-neutral firms, as the earlier studies on credit rating predictions have favoured random forest algorithms. The results can help various stakeholders to use appropriate methods to forecast the creditworthiness of these firms. At the managerial level, firms with better credit ratings need to be more conscious of the macroeconomic risk factors, as an abrupt change could trigger a downgrade as well. For mid-rated firms, capital structure choices should be closely monitored, as a lapse can lead to a deterioration in the credit quality. Finally, the managers of lower-rated firms should be cognizant of the financial risks involved. The implications can extend beyond the cost of capital for low carbon firms, as creditworthiness is essential for valuation, corporate governance, and investment liquidity.

The machine learning models benefit from a sizable data source without exhaustive assumptions providing more precise forecast estimates. The automation in the process is cost-effective, and can enable rating agencies to offer their services more competitively. The dynamism of carbon-neutral firms is also adequately captured and can expedite achieving sustainability goals. Future research that is based on machine learning models and credit ratings can focus on the soft information employed in the process. Machine learning tools are famous for their processing criteria of hard information that includes numeric data. However, credit rating opinions are sometimes supported by qualitative evidence. Therefore, it will be interesting to assess how these models perform, especially when underlying information is both qualitative and quantitative.

CRediT authorship contribution statement

Baojun Yu: Conceptualization, Investigation, Writing – review & editing. **Changming Li:** Supervision, Writing – review & editing. **Nawazish Mirza:** Data curation, Software, Validation, Project administration, Writing – original draft. **Muhammad Umar:** Conceptualization, Methodology, Writing – review & editing.

References

Achim, M.V., Safta, I.L., Văidean, V.L., Mureşan, G.M., Borlea, N.S., 2021. The impact of covid-19 on financial management: evidence from Romania. Econ. Res. Ekon. Istraz. 10.1080/1331677X.2021.1922090.

Afzal, A., Mirza, N., 2012. Interest rate spreads in an emerging economy: the case of Pakistan's commercial banking sector. Econ. Res. Ekon. Istraživanja 25, 987–1004. https://doi.org/10.1080/1331677X.2012.11517543.

Alanis, E., 2020. Is there valuable private information in credit ratings? North Am. J. Econ. Finance 54, 101293. https://doi.org/10.1016/J.NAJEF.2020.101293.

Athey, S., Imbens, G.W., 2019. Machine learning methods that economists should know about. Annu. Rev. Econ. 11, 685–725. https://doi.org/10.1146/annurev-economics-080217-053433.

Attaoui, S., Cao, W., Duan, X., Liu, H., 2021. Optimal capital structure, ambiguity aversion, and leverage puzzles. J. Econ. Dyn. Control 129, 104176. https://doi.org/10.1016/J.JEDC.2021.104176.

Bahrammirzaee, A., 2010. A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. Neural Computing and Applications 19 (8), 1165–1195.

Barboza, F., Kimura, H., Altman, E., 2017. Machine learning models and bankruptcy prediction. Expert Syst. Appl. 10.1016/j.eswa.2017.04.006.

Bas, J., Cirillo, C., Cherchi, E., 2021. Classification of potential electric vehicle purchasers: a machine learning approach. Technol. Forecast. Soc. Change 168, 120759. https://doi.org/10.1016/J.TECHFORE.2021.120759.

Bibi, A., Zhang, X., Umar, M., 2021. The imperativeness of biomass energy consumption to the environmental sustainability of the United States revisited. Environ. Ecol. Stat. 10.1007/s10651-021-00500-9.

Blazquez, D., Domenech, J., 2018. Big data sources and methods for social and economic analyses. Technol. Forecast. Soc. Change 130, 99–113. https://doi.org/10.1016/J.TECHFORE.2017.07.027.

Blomkvist, M., Löflund, A., Vyas, H., 2021. Credit ratings and firm life-cycle. Finance Research Letters 39, 101598.

Buehler, H., Gonon, L., Teichmann, J., Wood, B., 2019. Deep hedging. Quant. Finance. 10.1080/14697688.2019.1571683.

Caragnano, A., Mariani, M., Pizzutilo, F., Zito, M., 2020. Is it worth reducing GHG emissions? Exploring the effect on the cost of debt financing. J. Environ. Manage. 270, 110860 https://doi.org/10.1016/j.jenvman.2020.110860.

Chaudhry, S.M., Ahmed, R., Shafiullah, M., Duc Huynh, T.L., 2020. The impact of carbon emissions on country risk: evidence from the G7 economies. J. Environ. Manage. 265, 110533 https://doi.org/10.1016/j.jenvman.2020.110533.

Chen, N.Y., Chen, K.C., Liu, C.C., 2019. Debt-financed repurchases and credit ratings with the respect of free cash flow and repurchase purpose. North Am. J. Econ. Finance 47, 23–36. https://doi.org/10.1016/J.NAJEF.2018.11.013.

Driss, H., Drobetz, W., El Ghoul, S., Guedhami, O., 2021. Institutional investment horizons, corporate governance, and credit ratings: International evidence. Journal of Corporate Finance 67, 101874.

Fève, P., Sanchez, P.G., Moura, A., Pierrard, O., 2021. Costly default and skewed business cycles. European Economic Review 132, 103630.

Gan, L., Wang, H., Yang, Z., 2020. Machine learning solutions to challenges in finance: an application to the pricing of financial products. Technol. Forecast. Soc. Change. 10.1016/j.techfore.2020.119928.

Golbayani, P., Florescu, I., Chatterjee, R., 2020. A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. North Am. J. Econ. Finance 54, 101251. https://doi.org/10.1016/j.najef.2020.101251.

Guenther, N., Schonlau, M., 2016. Support vector machines. The Stata Journal 16 (4), 917–937.

Gopalakrishnan, B., Mohapatra, S., 2020. Insolvency regimes and firms' default risk under economic uncertainty and shocks. Economic Modelling 91, 180–197.

Hasan, I., Kim, S.J., Politsidis, P.N., Wu, E., 2021. Loan syndication under Basel II: how do firm credit ratings affect the cost of credit? J. Int. Financ. Mark. Inst. Money 72, 101331. https://doi.org/10.1016/j.intfin.2021.101331.

Hasnaoui, J.A., Rizvi, S.K.A., Reddy, K., Mirza, N., Naqvi, B., 2021. Human capital efficiency, performance, market, and volatility timing of asian equity funds during COVID-19 outbreak. J. Asset Manag. 2021, 1–16. https://doi.org/10.1057/S41260-021-00228-Y.

Höck, A., Klein, C., Landau, A., Zwergel, B., 2020. The effect of environmental sustainability on credit risk. Journal of Asset Management 21 (2), 85–93.

Huang, S.C., 2011. Using Gaussian process based kernel classifiers for credit rating forecasting. Expert Syst. Appl. 38, 8607–8611. https://doi.org/10.1016/j.eswa.2011.01.064.

Ji, X., Zhang, Y., Mirza, N., Umar, M., Rizvi, S.K.A., 2021. The impact of carbon neutrality on the investment performance: evidence from the equity mutual funds in BRICS. J. Environ. Manage. 297, 113228 https://doi.org/10.1016/J.JENVMAN.2021.113228.

Jiang, Q., Ma, X., 2021. Spillovers of environmental regulation on carbon emissions network. Technol. Forecast. Soc. Change 169, 120825. https://doi.org/10.1016/j.techfore.2021.120825.

Kanagaretnam, K., Lim, C.Y., Lobo, G.J., 2014. Influence of national culture on accounting conservatism and risk-taking in the banking industry. The Accounting Review 89 (3), 1115–1149.

Kouloukoui, D., de Oliveira Marinho, M.M., da Silva Gomes, S.M., Kiperstok, A., Torres, E.A., 2019. Corporate climate risk management and the implementation of climate projects by the world's largest emitters. Journal of Cleaner Production 238, 117935.

Lever, J., Krzywinski, M., Altman, N., 2017. Points of significance: Principal component analysis. Nature methods 14 (7), 641–643.

Lassala, C., Orero-Blat, M., Ribeiro-Navarrete, S., 2021. The financial performance of listed companies in pursuit of the Sustainable Development Goals (SDG). Econ. Res. Ekon. Istraz. 34, 427–449. https://doi.org/10.1080/1331677X.2021.1877167.

Lee, C., 2021. A review of data analytics in technological forecasting. Technol. Forecast. Soc. Change 166, 120646. https://doi.org/10.1016/J.TECHFORE.2021.120646.

Li, J.P., Mirza, N., Rahat, B., Xiong, D., 2020. Machine learning and credit ratings prediction in the age of fourth industrial revolution. Technol. Forecast. Soc. Change 161, 120309. https://doi.org/10.1016/j.techfore.2020.120309.

Malekipirbazari, M., Aksakalli, V., 2015. Risk assessment in social lending via random forests. Expert Systems with Applications 42 (10), 4621–4631.

Mirza, N., Rahat, B., Naqvi, B., Rizvi, S.K.A., 2020. Impact of Covid-19 on corporate solvency and possible policy responses in the EU. Q. Rev. Econ. Finance. 10.1016/j.qref.2020.09.002.

Mullainathan, S., Spiess, J., 2017. Machine learning: an applied econometric approach. J. Econ. Perspect. 31, 87–106. https://doi.org/10.1257/jep.31.2.87.

Naqvi, B., Mirza, N., Rizvi, S.K.A., Porada-Rochoń, M., Itani, R., 2021. Is there a green fund premium? Evidence from twenty seven emerging markets. Glob. Finance J. 100656 https://doi.org/10.1016/J.GFJ.2021.100656.

Nawazish, M., Saeed, M.S., Abbas Rizvi, K., 2013. The pricing of size, book to market and financial leverage in Euro stocks. Ekon. Istraz. 26, 177–190. https://doi.org/10.1080/1331677x.2013.11517615.

Öğüt, H., Doğanay, M.M., Ceylan, N.B., Aktaş, R., 2012. Prediction of bank financial strength ratings: The case of Turkey. Economic Modelling 29 (3), 632–640.

Pang, S., Hou, X., Xia, L., 2021. Borrowers' credit quality scoring model and applications, with default discriminant analysis based on the extreme learning machine. Technological Forecasting and Social Change 165, 120462.

Pertaia, G., Prokhorov, A., Uryasev, S., 2021. A new approach to credit ratings. J. Bank. Finance 106097. https://doi.org/10.1016/J.JBANKFIN.2021.106097.

Su, C.-W., Cai, X.-Y., Qin, M., Tao, R., Umar, M., 2021a. Can bank credit withstand falling house price in China? Int. Rev. Econ. Finance 71, 257–267. https://doi.org/10.1016/j.iref.2020.09.013.

Su, C.-W., Yuan, X., Tao, R., Umar, M., 2021b. Can new energy vehicles help to achieve carbon neutrality targets? J. Environ. Manage. 297, 113348 https://doi.org/10.1016/j.jenvman.2021.113348.

Szetela, B., Mentel, G., Brożyna, J., 2019. Modelling European sovereign default probabilities with copulas. Econ. Res. Ekon. Istraz. 32, 1716–1726. https://doi.org/10.1080/1331677X.2019.1629325.

Taghizadeh-Hesary, F., Rasoulinezhad, E., Yoshino, N., Sarker, T., Mirza, N., 2021. Determinants of the Russia and Asia–Pacific energy trade. Energy Strategy Rev. 38, 100681 https://doi.org/10.1016/J.ESR.2021.100681.

Tang, L., Cai, F., Ouyang, Y., 2019. Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China. Technol. Forecast. Soc. Change. 10.1016/j.techfore.2018.03.007.

Tao, R., Umar, M., Naseer, A., Razi, U., 2021. The dynamic effect of eco-innovation and environmental taxes on carbon neutrality target in emerging seven (E7) economies. Journal of Environmental Management 299, 113525. https://doi.org/10.1016/j.jenvman.2021.113525.

Tateishi, H.R., Bragagnolo, C., de Faria, R.N., 2020. Economic and environmental efficiencies of greenhouse gases' emissions under institutional influence. Technol. Forecast. Soc. Change 161, 120321. https://doi.org/10.1016/j.techfore.2020.120321.

Umar, M., Ji, X., Kirikkaleli, D., Alola, A.A., 2021a. The imperativeness of environmental quality in the United States transportation sector amidst biomass-fossil energy consumption and growth. J. Clean. Prod. 285, 124863 https://doi.org/10.1016/j.jclepro.2020.124863.

Umar, M., Ji, X., Kirikkaleli, D., Shahbaz, M., Zhou, X., 2020a. Environmental cost of natural resources utilization and economic growth: can China shift some burden through globalization for sustainable development? Sustain. Dev. 28, 1678–1688. https://doi.org/10.1002/sd.2116.

Umar, M., Ji, X., Kirikkaleli, D., Xu, Q., 2020b. COP21 Roadmap: do innovation, financial development, and transportation infrastructure matter for environmental sustainability in China? J. Environ. Manage. 271, 111026 https://doi.org/10.1016/j.jenvman.2020.111026.

Umar, M., Ji, X., Mirza, N., Naqvi, B., 2021b. Carbon neutrality, bank lending, and credit risk: evidence from the Eurozone. J. Environ. Manage. 296, 113156 https://doi.org/10.1016/j.jenvman.2021.113156.

Umar, M., Ji, X., Mirza, N., Rahat, B., 2021c. The impact of resource curse on banking efficiency: evidence from twelve oil producing countries. Resour. Policy 72, 102080. https://doi.org/10.1016/j.resourpol.2021.102080.

Varian, H.R., 2014. Big data: new tricks for econometrics. J. Econ. Perspect. 28, 3–28. https://doi.org/10.1257/jep.28.2.3.

Vrontis, D., Morea, D., Basile, G., Bonacci, I., Mazzitelli, A., 2021. Consequences of technology and social innovation on traditional business model. Technol. Forecast. Soc. Change 170, 120877. https://doi.org/10.1016/J.TECHFORE.2021.120877.

Wallis, M., Kumar, K., Gepp, A., 2019. Credit rating forecasting using machine learning techniques. In: Sun, Z. (Ed.), Managerial Perspectives on Intelligent Big Data Analytics. IGI Global, Hershey, PA, USA, pp. 180–198. https://doi.org/10.4018/978-1-5225-7277-0.ch010.

Wang, Q.-S., Su, C.-W., Hua, Y.-F., Umar, M., 2021. How can new energy vehicles affect air quality in China?— From the perspective of crude oil price. Energy Environ. 0958305X211044388 https://doi.org/10.1177/0958305X211044388.

Weigand, A., 2019. Machine learning in empirical asset pricing. Financ. Mark. Portf. Manag. 10.1007/s11408-019-00326-3.

Wu, H.C., Hu, Y.H., Huang, Y.H., 2014. Two-stage credit rating prediction using machine learning techniques. Kybernetes 43, 1098–1113. https://doi.org/10.1108/K-10-2013-0218.

Wu, P., Guo, F., Cai, B., Wang, C., Lv, C., Liu, H., Huang, J., Huang, Y., Cao, L., Pang, L., Gao, J., 2021. Co-benefits of peaking carbon dioxide emissions on air quality and health, a case of Guangzhou, China. J. Environ. Manage. 282, 111796 https://doi.org/10.1016/j.jenvman.2020.111796.

Zhang, X., Ouyang, R., Liu, D., Xu, L., 2020. Determinants of corporate default risk in China: The role of financial constraints. Economic Modelling 92, 87–98.

Zeidan, R., Shapir, O.M., 2017. Cash conversion cycle and value-enhancing operations: theory and evidence for a free lunch. J. Corp. Finance 45, 203–219. https://doi.org/10.1016/J.JCORPFIN.2017.04.014.

Zhu, Y., Zhou, L., Xie, C., Wang, G.J., Nguyen, T.V., 2019. Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. International Journal of Production Economics 211, 22–33.

Baojun Yu is a Professor at School of Management, Jilin University, China. His research interests include information systems, digital economy, IT governance, etc. He has extensive professional and consulting experience in business process improvement, IT investment, and organizational change.

Changming Li is a lecturer at School of Engineering, Changchun Guanghua University, China. His research interests include IoT, artificial intelligence, digital economy. He has vast experience in software system optimization, digital economic analysis, artificial intelligence algorithm, etc.

Nawazish Mirza is a Full Professor of Finance at Excelia Business School, France. He graduated with a Ph.D. from the University of Paris Dauphine and worked in investment banking and credit ratings before his academic stint. His research interests include technological innovations, credit risk, fund management, financial intermediation, and valuations. He has extensive consulting experience in credit ratings, investment banking, and the valuation of new technologies. His recent research has been published in Technological Forecasting and Social Change, International Review of Economics and Finance, Economic Modelling, Pacific-Basin Finance, Finance Research Letters, Journal of Asset Management, Resources Policy, and Journal of Environmental Management, among others.

Muhammad Umar is an Assistant Professor in UCP Business School, Faculty of Management Studies, University of Central Punjab, Lahore, Pakistan. He has done his Ph.D. in Management Science and Engineering from the School of Business, Qingdao University, China. He has 13 years of work experience in different Academic and Research Institutions. His research interests focus on technological-based Finance, financial markets, risk management, energy finance, resources and environmental management, and environmental economics. Umar's work is featured in various well-reputed journals, including Technological Forecasting & Social Change, Journal of Cleaner Production, Energy, Energy Policy, Pacific-Basin Finance Journal, Journal of Environmental Management, Resources Policy, Economic Research-Ekonomska Istraživanja, and Science of the Total Environment.