Finance Research Letters 58 (2023) 104648

Contents lists available at ScienceDirect

Finance Research Letters

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

Prediction of corporate credit ratings with machine learning: Simple interpretative models

Koresh Galil [a,b,*], Ami Hauptman [c], Rosit Levy Rosenboim [d]

[a] Department of Economics, Ben-Gurion University of the Negev, Israel
[b] Ben-Gurion University of the Negev, Department of Economics, P.O. Box 653, Beer-Sheva 8410501, Israel
[c] Department of Computer Science, Sapir Academic College, Israel
[d] Department of Applied Economics, Sapir Academic College, Israel

ARTICLE INFO

JEL classification:
C45
C53
G24
G32

Keywords:
Corporate ratings
Machine learning
Classification and regression tree
Support Vector Regression
CART
SVR
Size

ABSTRACT

This study utilizes machine learning techniques, notably classification and regression trees (CART) and support vector regression (SVR), to predict corporate credit ratings. While SVR marginally outperforms in accuracy, CART offers interpretability. However, unconstrained models can produce non-monotonic relationships between credit ratings and core features, an undesired outcome. To circumvent this, we recommend restricted CART models that ensure interpretable, theory-consistent results. We underscore the importance of company size in credit rating prediction with an ideal model integrating size, interest coverage, and dividends. Although being a large-cap company is crucial, it doesn't guarantee high ratings, and small-cap companies rarely secure investment-grade ratings.

1. Introduction

Corporate credit ratings play a pivotal role in bond markets and various investment platforms (Hilscher and Wilson, 2017). These ratings assist analysts in determining a company's cost of debt and its overall weighted cost of capital. In instances where a company remains unrated, synthetic or shadow ratings act as invaluable proxies (Damodaran, 2012), supporting corporate executives during bond offerings and offering insights to risk managers and financial analysts in their credit risk assessments.

Traditional models for credit rating predictions, often referred to as shadow ratings, are effective but sometimes miss the intricate nuances of financial data. As emphasized in the special comment by Moody's rating agency (Metz and Cantor, 2006), conventional models, including linear regressions and ordered probit/logit models, possess inherent strengths. However, they may not consistently capture the dynamic interplay of credit metrics and their evolving relationships over time—a challenge that machine learning models are adept at addressing.

The literature on credit rating prediction using machine learning is extensive, with significant contributions from studies such as Addo et al. (2018); Golbayani et al. (2020); Huang et al. (2004); and Li et al. (2020). Building upon the well-documented potential of

* Corresponding author at: Economics Department of Ben-Gurion University of the Negev, Ami Hauptman is from the Computer Science Department of Sapir College, and Rosit Levy Rosenboim is from the Applied Economics Department of Sapir College, Israel.

E-mail address: galilk@bgu.ac.il (K. Galil).

https://doi.org/10.1016/j.frl.2023.104648
Received 1 July 2023; Received in revised form 21 October 2023; Accepted 22 October 2023
Available online 26 October 2023

machine learning techniques, exemplified by works like Li et al. (2020), our research offers a fresh perspective. Rather than exclusively chasing the pinnacle of accuracy, as observed in numerous preceding studies across diverse sectors and regions, we prioritize models that seamlessly merge accuracy with simplicity. Our models are distinct, leveraging a select subset of features applicable universally to all non-financial public firms, ensuring their broad applicability. Furthermore, we've crafted our models with a keen focus on theoretical consistency, ensuring they exhibit monotonicity concerning the explanatory variables. These rating prediction models are nearly as accurate as sophisticated models while still being both simple and interpretable. Importantly, the insights derived from our models are invaluable for corporate managers and financial analysts, equipping them with actionable intelligence for informed decision-making and strategic planning.

Our study proceeds as follows. The dataset and the methodology are described in Section 2. Section 3 presents the results, and Section 4 concludes.

2. Data and methodology

2.1. Data

The study's initial sample includes all firms in the COMPUSTAT database from 2005 to 2016, with an S&P issuer rating (non-default) on the financial year's last day.[1] In compliance with previous literature, we transform all ratings to numerical values, designating the value 21 for the highest rating category (AAA) and then 20 to AA+ down to 5 for CCC+ or lower. In addition, for our descriptive statistics, we also use ratings in main categories (AAA, AA, A, BBB, BB, B, CCC or lower). We exclude financial firms (SIC 6000–6999) and other firms with unique characteristics, such as agriculture, utilities, and government firms (SIC code 0 to 100, 4900 to 4999, and 9000 to 9999). Our database comprises 13,937 annual observations of 1988 firms with complete data on explanatory variables.

We employ accounting explanatory variables, defined in alignment with the S&P rating criteria (Standard and Poor's, 2008), drawing inspiration from studies such as Blume et al. (1998); Jorion et al. (2009); Alp (2013), and Baghai et al. (2014). Initially, our model incorporated twelve explanatory variables. However, a rigorous analysis pinpointed six variables that exhibited the most significant explanatory power: Size, Interest Coverage Ratio (ICR), Total Debt Leverage, Dividend Payer, Operating Margin, and Market to Book Value of Equity.[2] Detailed definitions of these variables are provided in the Appendix. Consistent with the S&P methodology and studies like Blume et al. (1998) and Baghai et al. (2014), all financial ratios are averaged over a three-year period, and all variables except for Size and Total Debt Leverage are winsorized at the 99th percentile level, while ICR is winsorized at 1 (at the bottom) and 100 (at the ceiling).

Table 1 illustrates the distribution of annual observations across the main rating categories (AAA, AA, A, BBB, BB, B, CCC, and lower). We observe a decreasing number of AAA ratings over the years, from 9 in 2005 to just 3 in 2016. A milder pattern is observed in the AA and A rating categories. Other rating categories, however, do not display any trend over time.

Table 2 presents summary statistics of the sample. Panel (a) displays statistics for the entire sample, which aligns with the behavior of samples observed in previous literature (e.g., Alp, 2013; Baghai et al., 2014). Panel (b) shows the means of variables across the main rating categories. The table demonstrates that the means of variables display a monotonic trend across ratings. Higher ratings are linked to greater Size, Interest Coverage, Operating Margin, and Market to Book ratio. Additionally, firms with higher ratings exhibit a greater propensity to pay dividends and maintain lower leverage.

2.2. Methodology

We apply two machine-learning techniques to our data, namely classification and regression trees (CART) and support vector regression (SVR). We hereby describe both methods briefly. In the preceding sections, we describe the methods for building and evaluating our models.

2.2.1. Classification and regression trees

Classification and Regression Trees (CART) is a widely used decision tree-based algorithm for both classification and regression tasks (Breiman et al., 1984), known for its simplicity, interpretability, and robustness in various fields, such as finance (Yang et al., 2014) and medicine (Luna et al., 2019). CART predicts the value of a continuous target variable by recursively partitioning the input space, constructing a binary tree with internal nodes representing feature tests, branches corresponding to outcomes, and leaf nodes representing predicted values. The algorithm selects feature-split points to maximize the reduction in sum of squared errors (SSE) for each partition, optimizing the cost function:

$$J_m = \frac{m_{left}}{m} * MSE_{left} + \frac{m_{right}}{m} * MSE_{right}$$

[1] We chose not to use data before 2004 because of the trend in rating criteria documented in prior research (e.g. Afik and Galil, 2012; Alp, 2013). Data after 2016 is not available because of discontinuity of rating data in Compustat in early 2017.

[2] The other six financial ratios omitted from the final analysis are: R&D to Total Assets, Retained Earning to Total Assets, Capital Expenditures to Total Assets, Cash Balances to Total Assets, Tangible Assets (Property, plant, and equipment) to Total Assets, and Convertible Debt to Assets.

Table 1
Rating Distribution over Years. This table shows the breakdown of the sample across main rating categories and over the years.

Year	AAA	AA	A	BBB	BB	B	CCC or lower	Total
2005	9	50	258	385	330	172	18	1222
2006	9	57	237	380	322	193	19	1217
2007	8	60	226	348	319	205	18	1184
2008	6	54	214	345	309	214	28	1170
2009	4	48	193	348	280	234	44	1151
2010	4	43	200	356	272	243	21	1139
2011	5	37	199	366	277	212	16	1112
2012	5	33	200	377	278	208	19	1120
2013	5	35	205	392	288	200	12	1137
2014	4	41	206	390	308	207	12	1168
2015	4	41	184	403	292	211	27	1162
2016	3	37	179	402	300	189	45	1155
Total	66	536	2501	4492	3575	2488	279	13,937

Table 2
Summary statistics. This table shows the summary statistics of the sample that covers 13,937 firm-year observations over the years 2005–2016. Panel (a) shows the descriptive statistics of the explanatory variables and Panel (b) shows the variable means across main rating categories.

Panel (a) – Descriptive statistics

	Mean	Median	Standard Deviation	Minimum	Maximum
Size	8.348	8.325	1.790	−4.851	13.229
Interest Coverage	17.157	6.674	26.985	1.000	100.000
Dividend payer	0.662	1.000	0.473	0.000	1.000
Total Debt Leverage	0.318	0.283	0.204	0.012	1.129
Operating Margin	0.207	0.167	0.156	−0.159	0.783
Market to Book	1.560	1.316	0.829	0.386	12.554

Panel (b) – Variable means across main rating categories

	Size	Interest Coverage	Dividend payer	Total Debt Leverage	Operating Margin	Market to Book
AAA	11.556	41.247	0.985	0.181	0.296	1.887
AA	10.975	49.232	0.978	0.190	0.290	1.783
A	9.939	31.005	0.943	0.212	0.254	1.772
BBB	8.808	19.238	0.852	0.252	0.212	1.601
BB	7.745	8.972	0.514	0.357	0.197	1.490
B	6.500	5.296	0.236	0.485	0.161	1.333
CCC or lower	5.062	2.876	0.097	0.612	0.075	1.415
Total	8.348	17.157	0.662	0.318	0.207	1.560

ensuring an accurate fit to the training data. CART provides easily interpretable and visualized if-then-else rules, facilitating decision-making processes.

2.2.2. *Support vector machine for regression (SVR)*

Support Vector Regression (SVR) is an extension of the Support Vector Machine (SVM) algorithm for regression tasks, which has been widely used in various fields, such as finance (Tas and Atli., 2022), and biology (Batta et al., 2022), due to its ability to handle high-dimensional data and its robustness to noise. SVR employs nonlinear projections to map the input data into a higher-dimensional feature space, where the regression function can be effectively estimated. By minimizing the norm of the weight vector while satisfying the constraints, SVR seeks to identify the optimal hyperplane that separates the projected data points, maximizing the margin between predicted and actual values in the higher-dimensional space. This allows SVR to capture complex relationships between variables and improve regression performance.

2.2.3. *Experiments*

We trained and tested our prediction models using the six explanatory variables described in Section 2.1. The overall data set for learning consisted of 12,559 rows, each containing values for all six variables, while the credit ratings were used as tags.

We generated all non-empty subsets of the explanatory variables. For each subset size $N \in [1, \ldots, 6]$, we conducted a group of experiments, iterating over all possible combinations of N variables. For each combination, we trained two types of learning models: an SVR model, and a regression-tree model.

The SVR model was trained with the default parameter set used in the scikit-learn Python library.[3] Regarding regression-tree models, for each variable combination, we tried all depths in the range $[2, \ldots, 6]$. We repeated each experiment configuration

[3] The default parameter set produced the best results. See scikitlearn.org for a comprehensive description of SVR parameters.

around 1000 times to address the random nature of model training.

In regression tree construction, variables are selected at each step based on optimality constraints, without the requirement of selecting all variables in the current combination, resulting in cases where specific trees did not contain all available variables.

Finally, all experiments with CART models were conducted two times: First, with applying monotonicity constraints, and second without such constraints. We consider a model monotonic if improving on one or more variables cannot lead to a lower predicted rating.

2.2.4. Evaluation

For each experiment, we performed a random train-test split, where 0.8 of the data was used for training and 0.2 for testing. All models were evaluated using R-squared, RMSE and notch-distance accuracy (ACC) measures. Notch-distance was derived by notch-distance as described in Golbayani et al. (2020): let d denote the absolute prediction error. If $d \leq 0.5$, distance is calculated as 0. If $0.5 < d \leq 1.5$, the distance is 1, and if $d > 1.5$, the distance is considered 2 or above and counted as binary prediction error. For each experiment configuration we kept the best trees for reference, using R-squared as the determining factor.

3. Results

In constructing our prediction models, we generated sets of 1 to 6 explanatory variables. Table 3 presents the best models based on R-squared, RMSE, and ACC measures, focusing on interpretable CART models and including SVR models for comparison. SVR models outperform CART models in some combinations of three variables and in all combinations of four to six variables. Additional predictors generally improve performance measures, but not all additions enhance accuracy. The value of added predictors diminishes, with the highest R-squared increasing from 0.6968 for a single predictor model to 0.7889 for a six-predictor model, reflecting their high correlation. Zmijewski (1984) found that three predictors suffice for a robust bankruptcy prediction model.[4]

Size emerges as the best single predictor for ratings, surpassing other predictors in both CART and SVR models. This finding is quite astonishing, given that Size is not directly associated with the ability of a firm to serve its debts.[5] Previous studies favored ratios like ICR (Damodaran, 2012) or cash flow to debt ratios (Beaver, 1966) for bankruptcy prediction because this ratio combines both the firm's ability to generate profits/cash and the amount of debt it has to serve. Notably, Beaver (1966) also conducted univariate analysis and did not consider Size a bankruptcy predictor.

Comparing our models' performance measures with previous literature has limitations. Our panel database predicts ratings across different time periods, while previous studies mainly focused on single-year predictions. Rating agencies consider the business cycle and avoid rapid rating changes (Löffler, 2004, 2005). Moreover, rating criteria change over time (e.g., Blume et al., 1998, Afik and Galil, 2022). Our panel dataset-based model is more robust and adaptable than single-year sample models. Previous studies often focused on specific industries, while our model is designed for a wide range of industries. Lastly, we prioritize simplicity and interpretability with a limited number of variables, whereas previous studies used numerous variables and uninterpretable models.

In comparison to Wallis et al. (2019), who utilized 27 variables and a random-forest model to predict S&P 500 corporate ratings for 308 firms in 2016–2017 with an ACC of 0.646, our best CART model with three variables achieved an ACC of 0.676. Our model, based on a sample of 13,937 firm-year observations spanning various sectors and 12 years, demonstrates strong performance despite its simplicity and use of only three explanatory variables.

Table 4 presents the CART model estimation using Size as the predictor, which achieves the highest accuracy measures among single-variable models. For convenience, Size is transformed into the market value of equity in 2022 prices. The model exhibits non-monotonic behavior, where slight increases in market value above a threshold lead to rating drops. This non-monotonic pattern persists throughout the table and is observed in other estimated models with one or more predictors. While this model demonstrates high accuracy, its results are difficult to comprehend, making it challenging for corporate executives to rely on its predictions for desired rating grades. It should be acknowledged that uninterpretable models like SVR or neural-network models may also exhibit non-monotonic behavior.

To ensure practicality for analysts and executives, we sought monotonic models in our estimation. Monotonic models are defined as those where improvements in variables do not result in lower predicted ratings. We carefully examined all estimated models and selected only the monotonic ones. Table 5 presents a comparison of performance measures for the best monotonic CART models. The accuracy measures of these models only show a slight decrease compared to the unrestricted models in Table 3. The top-performing model, incorporating ICR, Size, and DIVP as the three variables, outperforms the best models with four or five predictors. None of the models using all six predictors were found to be monotonic in our search.

Table 6 presents our best single-variable monotonic model estimate. The model utilizes Size to classify firms into 14 rating grades, excluding AAA, AA+, and AA- which are difficult to distinguish from AA. Notably, Damodaran's (2012) model also predicts up to 13

[4] Zmijewski (1984) used only three variables, each representing a different group of financial performance. Return on assets (net income to total assets) represented profitability, total debt to total financial leverage and current assets to current liability represented liquidity.

[5] We also explore two alternative proxies for size: sales and total assets. Models employing the market value of equity as a proxy significantly outperform those employing sales or total assets. Nevertheless, the models utilizing the alternative proxies for size still exhibit significantly superior performance compared to models disregarding size. Our findings align with those of Dang et al. (2018), who discovered that across 20 distinct areas of corporate finance, the three proxies exhibit consistent coefficient signs and similar statistical significance. However, in certain domains, the R-squared value exhibits significant fluctuations in models employing the alternative proxies.

Table 3

Accuracy measures for unrestricted models. This table shows the accuracy measures for various rating unrestricted prediction models. The accuracy measures are R-squared (R^2), RMSE, and ACC as defined in the text. SIZE is the log of market value of equity in 1985 prices, ICR is interest coverage, DIVP is a dummy variable that gets the value 1 if the firm pays dividends, and zero otherwise, OM is operating margin. TDL is total debt leverage, and MB is the market-to-book ratio. Variables definitions appear in the appendix.

No. of Variables used	Variables used	Regression Tree				SVR		
		Depth	R^2	RMSE	ACC	R^2	RMSE	ACC
1	SIZE	4	0.6968	1.9793	0.5653	0.6668	2.0001	0.5727
1	ICR	4	0.5371	2.4402	0.4772	0.4898	2.5237	0.4697
1	DIVP	3	0.3542	2.8029	0.4426	0.3475	2.7663	0.4165
2	ICR, SIZE	6	0.7469	1.7029	0.6548	0.7452	1.7774	0.6470
2	ICR, OM	3	0.5408	2.4275	0.4655	0.4929	2.4246	0.5180
2	ICR, DIVP	6	0.6026	2.1922	0.5456	0.5837	2.2307	0.5276
2	TDL, SIZE	6	0.6962	1.8885	0.6057	0.6711	1.9428	0.5861
2	SIZE, DIVP	5	0.7081	1.8559	0.6057	0.6393	1.9936	0.5776
2	SIZE, OM	6	0.6762	1.9056	0.6011	0.6550	1.9866	0.5782
3	ICR, OM, SIZE	6	0.7422	1.7282	0.6732	0.7810	1.5751	0.6994
3	ICR, TDL, SIZE	6	0.7351	1.7120	0.6758	0.7593	1.6331	0.6823
3	ICR, SIZE, DIVP	6	0.7804	1.6319	0.6746	0.7890	1.6386	0.6826
3	ICR, SIZE, MB	6	0.7610	1.7101	0.6736	0.7453	1.7881	0.6513
3	TDL, SIZE, DIVP	6	0.7342	1.8262	0.6274	0.7471	1.7693	0.6332
4	ICR, OM, SIZE, DIVP	6	0.7640	1.6562	0.6745	0.8210	1.4584	0.7190
4	ICR, SIZE, DIVP, MB	6	0.7815	1.6445	0.6757	0.7903	1.6338	0.6810
4	OM, TDL, SIZE, DIVP	6	0.7383	1.8258	0.6231	0.7680	1.7249	0.6714
5	ICR, OM, TDL, DIVP, SIZE	6	0.7723	1.5804	0.6928	0.8193	1.4472	0.7504
6	ICR, OM, TDL, SIZE, DIVP, MB	7	0.7889	1.5967	0.7066	0.8339	1.3818	0.7472

Table 4

An unrestricted rating model using Size. This table shows the estimated unrestricted tree regression model with the highest R-squared. SIZE is the log of the market value of equity in 1985 prices. For convenience, we transform Size to the market value of equity in December 2022 prices. The model has an R-squared of 0.698, RMSE of 1.9793, and ACC measure of 0.5653**.**

Rating	Size	Market Value of Equity (December 2022 prices)		
AA+	11.761-	359,850	–	∞
BBB-	11.744–11.761	353,785	–	359,850
AA-	11.377–11.744	245,105	–	353,785
A+	10.333–11.377	86,288	–	245,105
A-	10.206–10.333	75,997	–	86,288
A	10.072–10.206	66,466	–	75,997
A-	9.709–10.072	46,233	–	66,466
AAA	9.405–9.709	34,113	–	46,233
BBB+	9.051–9.405	23,943	–	34,113
BBB	8.665–9.051	16,276	–	23,943
A-	8.129–8.665	9523	–	16,276
BB	8.126–8.129	9494	–	9523
BBB-	8.11–8.126	9344	–	9494
AA+	8.109–8.11	9334	–	9344
BBB-	7.571–8.109	5450	–	9334
B	7.569–7.571	5440	–	5450
BB+	7.268–7.569	4026	–	5440
BBB	6.996–7.268	3067	–	4026
BB	6.673–6.996	2220	–	3067
BB-	6.626–6.673	2118	–	2220
BB	6.538–6.626	1940	–	2118
BB-	6.237–6.538	1436	–	1940
B+	4.98–6.237	408	–	1436
B	3.67–4.98	110	–	408
B-	1.495–3.67	13	–	110
CCC	0–1.495	0	–	13

classes using ICR. In our model, firms require a market value exceeding 12,971 million USD to be classified as investment-grade (BBB- or higher), and a market value exceeding 496,552 million USD for an AA rating or higher. Small-cap companies with a market value up to 2 billion USD are limited to a maximum rating of BB-, while mid-caps with a market value up to 10 billion USD cannot attain an investment-grade rating. These predictions may contradict actual observations as the model disregards other vital features.

Table 7 displays the best monotonic model found in our search, utilizing the market value of equity, ICR, and dividend payment as predictors. The model classifies ten grades. According to this model, a firm needs a market value greater than 481,877 to be classified as AA, regardless of other features. A mid-cap firm with a market value of 10 billion USD will receive an investment-grade rating of

Table 5

Accuracy measures for monotonic models. This table shows the accuracy measures for monotonic rating prediction models with the highest R-squared measure among all estimated unrestricted models. The accuracy measures R-squared (R^2), RMSE, and ACC as defined in the text. SIZE is the log of market value of equity in 1985 prices, ICR is interest coverage, DIVP is a dummy variable that gets the value 1 if the firm pays dividends, and zero otherwise, OM is operating margin. TDL is total debt leverage, and MB is market to book ratio. Variables definitions appear in the appendix.

No. of Variables used	Variables used	Regression Tree			
		Depth	R^2	RMSE	ACC
1	SIZE	5	0.683	2.0423	0.5666
1	ICR	4	0.503	2.4911	0.466
1	DIVP	5	0.3277	2.9882	0.405
2	ICR, SIZE	4	0.756	1.7753	0.637
2	ICR, OM	3	0.5494	2.4335	0.4952
2	ICR, DIVP	3	0.5355	2.3688	0.4936
2	TDL, SIZE	3	0.6138	2.1903	0.5191
2	SIZE, DIVP	3	0.6841	2.0273	0.5488
2	SIZE, OM	3	0.6744	2.0825	0.5435
3	ICR, OM, SIZE	3	0.7107	1.8941	0.6024
3	ICR, TDL, SIZE	3	0.7064	1.9413	0.6067
3	ICR, SIZE, DIVP	4	0.763	1.7504	0.641
3	ICR, SIZE, MB	3	0.7119	1.9087	0.586
3	TDL, SIZE, DIVP	3	0.676	2.0231	0.5812
4	ICR, OM, SIZE, DIVP	3	0.7133	1.9573	0.587
4	ICR, SIZE, DIVP, MB	3	0.719	1.9041	0.5886
4	OM, TDL, SIZE, DIVP	3	0.6839	2.0128	0.5796
5	ICR, OM, TDL, DIVP, SIZE	3	0.7128	1.9159	0.5966
6	ICR, OM, TDL, SIZE, DIVP, MB	Not found	–	–	–

Table 6

A monotonic rating model using Size. This table shows the estimated monotonic tree regression model with the highest R-squared. SIZE is the log of the market value of equity in 1985 prices. For convenience, we transform Size to the market value of equity in December 2022 prices. The model has an R-squared of 0.683, RMSE of 2.0423 and ACC measure of 0.5666.

Rating	Size	Market Value of Equity (December 2022 prices)		
AA	12.083-	496,552	–	∞
A+	11.586–12.083	302,079	–	496,552
A	10.704–11.586	125,047	–	302,079
A-	10.135–10.704	70,788	–	125,047
BBB+	9.774–10.135	49,338	–	70,788
BBB	8.986–9.774	22,436	–	49,338
BBB-	8.438–8.986	12,971	–	22,436
BB+	7.705–8.438	6232	–	12,971
BB	7.209–7.705	3795	–	6232
BB-	6.344–7.209	1598	–	3795
B+	5.775–6.344	905	–	1598
B	5.155–5.775	487	–	905
B-	3.752–5.155	120	–	487
CCC	−3.752	0	–	120

BBB- if its ICR is above 3.884 and it pays dividends. However, if the firm's ICR drops below 6.022 and it stops paying dividends, it will be downgraded to BB. A small-cap firm with a market value of 2 billion USD and dividend payments will receive a BB- rating if its ICR is below 3.884. Improving the ICR above 3.884 may result in an upgrade to BB. Size remains a dominant factor in improving a firm's rating over time. The model suggests that mid-cap firms can achieve an investment-grade rating of BBB+ with a high ICR and consistent dividend payments. However, to attain a rating of A or higher, a firm must be a large cap with a market value exceeding 72 billion USD.

It is evident from our models and real data (not shown for brevity) that ratings of AA-AAA are mostly limited to high-cap companies, while small-cap companies struggle to obtain BBB ratings. It should be noted that based on this table, a larger Size is necessary but not sufficient for higher ratings. Large-cap companies may still receive speculative-grade ratings, while small-cap companies cannot reach AA-AAA ratings.

During our robustness tests, we assessed the models under the challenging environment of the 2008–2009 financial crisis. Indicators such as RMSE and ACC revealed a marginal decline in performance across all unrestricted models during this period. Remarkably, our recommended three-variables (ICR, SIZE, DIVP) monotonic model still surpasses other monotonic CART models, even though it fares slightly worse than in non-crisis periods. Taking cues from Li et al. (2020) and Edirisinghe et al. (2022), we integrated macroeconomic variables into our analysis. These included: the risk-free rate (represented by one-year Treasury bills yield), the GDP nominal annual growth rate, the GDP real annual growth rate, and the inflation rate (represented by the annual consumer price index growth rate). Some newly estimated models exhibited marginally superior accuracy compared to our suggested model. However, none

Table 7
A monotonic rating model using Size, Interest Coverage (ICR) and Dividend-Payer (DIVP). This table shows the estimated monotonic tree regression model with the highest R-squared among those using only three explanatory variables. The explanatory variables are Market value of equity (in 2002 prices), ICR (interest coverage) and an indicator DIVP on whether the firm pays dividends or not. The model has R-squared of 0.7630, RMSE of 1.7504 and ACC measure of 0.6410.

Rating	Market Value of Equity (December 2022 prices)			Interest Coverage			Dividend
AA	481,877	–	∞		–		
A+	320,118	–	481,877		–		
A	72,146	–	320,118	12.020	–		
BBB+	72,146	–	320,118	− ∞	–	12.020	
BBB+	12,790	–	72,146	6.022	–	∞	pay
BBB-	12,790	–	72,146	6.022	–	∞	no pay
BBB-	12,790	–	72,146	− ∞	–	6.022	pay
BBB-	4066	–	12,790	3.884	–	∞	pay
BB	4066	–	12,790	3.884	–	∞	no pay
BB	1528	–	4066	3.884	–	∞	
BB	12,790	–	72,146	− ∞	–	6.022	no pay
BB-	865	–	12,790	− ∞	–	3.884	pay
B+	0	–	865	3.884	–	∞	
B+	865	–	12,790	− ∞	–	3.884	no pay
B	0	–	865	2.198	–	3.884	
CCC+	0	–	865	− ∞	–	2.198	

demonstrated monotonicity. Given the dual imperatives of simplicity and alignment with financial theory, no model conclusively surpassed our recommended model. Due to space constraints, detailed results aren't showcased but can be provided upon request.

4. Conclusions

In our study, we utilized machine-learning techniques to predict corporate credit ratings, offering key insights for corporate managers and financial analysts. Our findings indicate that while SVR models have a marginal advantage in accuracy, their interpretability is limited. Conversely, restricted CART models provide accuracy, transparency and consistency with financial theory. A central discovery is the significant influence of company size on credit ratings. Notably, large-cap companies don't consistently secure top-tier ratings, despite their size. In contrast, small-cap firms frequently face challenges in obtaining investment-grade ratings. Small-cap entities aiming to improve their ratings might adopt growth strategies, whereas large-cap companies might emphasize other financial indicators. For mid-cap companies, a focus on size, dividends, and a robust Interest Coverage Ratio proves advantageous. Additionally, our model serves as a valuable resource for financial analysts, particularly when evaluating shadow ratings for unrated entities.

Authors statement

During the preparation of this work the author(s) did not use AI technologies.

Declaration of Competing Interest

None.

Data availability

The authors do not have permission to share data.

Appendix –. Variables Definition

Size (SIZE) follows the definition by Blume et al. (1998). It is the equity market value of the firm (PRCC_f * CSHO), in million dollars, adjusted by the U.S. consumer price index (CPI) of January 1985 and then converted to its natural logarithm value.

Interest Coverage Ratio (ICR) follows the definitions by Blume et al. (1998) and Alp (2013). It is the ratio of operating income after depreciation (OIADP) plus interest expense (XINT) to interest expenses (XINT).

Total Debt Leverage (TDL) follows the definition by Alp (2013). It is the ratio of debt (DLTT+DLC) to total assets (AT).

Dividend payer (DIVP) follows the definition by Alp (2013). It is a dummy variable that equals 1 if the dividend per share (DVPSX_F) is positive, and equals zero otherwise.

Market to Book (MB) follows the definition by Alp (2013). It is the sum of total assets (AT) and market value of equity minus the book value of equity, all divided by total assets (AT). Market value of equity is the fiscal-year closing price (PRCC_F) times the shares outstanding (CSHO). The book value of equity is stockholder's equity (SEQ) minus preferred stock plus balance-sheet deferred taxes and investment tax credit (TXDITC). If data item TXDITC is missing, it is set to zero. If data item SEQ is unavailable, it is replaced by either common equity (CEQ) plus preferred stock par value (PSTK), or total assets (AT) minus total liabilities (LT). Preferred stock is preferred stock liquidating value (PSTKL), if missing then, preferred stock redemption value (PSTKRV), or preferred stock par value (PSTK).

Operating Margin (OM) follows the definition by Blume et al. (1998) and Alp (2013). It is operating income before depreciation (OIBDP) divided by sales (SALE).

References

Addo, P.M., Guegan, D., Hassani, B., 2018. Credit risk analysis using machine and deep learning models. Risks 6 (2), 38.

Afik, Z., & Galil, K. (2022). Have ratings become more accurate?. Available at SSRN 4219053.

Alp, A., 2013. Structural shifts in credit rating standards. J. Finance 68 (6), 2435–2470.

Baghai, R.P., Servaes, H., Tamayo, A., 2014. Have rating agencies become more conservative? Implications for capital structure and debt pricing. J. Finance 69 (5), 1961–2005.

Batta, I., Abrol, A., Calhoun, V.D., Alzheimer's Disease Neuroimaging Initiative, 2022. SVR-based multimodal active subspace analysis for the brain using neuroimaging data. bioRxiv 2022. -07.

Beaver, W.H., 1966. Financial ratios as predictors of failure. J. Account. Res. 71–111.

Blume, M.E., Lim, F., MacKinlay, A.C., 1998. The declining credit quality of US corporate debt: myth or reality? J. Finance 53 (4), 1389–1413.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Cart. Classification and Regression Trees. Wadsworth and Brooks/Cole, Monterey, CA, USA, 1984.

Damodaran, A., 2012. Investment valuation: Tools and Techniques For Determining the Value of Any Asset. John Wiley & Sons.

Edirisinghe, C., Sawicki, J., Zhao, Y., Zhou, J., 2022. Predicting credit rating changes conditional on economic strength. Finance Res. Lett. 47, 102770.

Golbayani, P., Florescu, I., Chatterjee, R., 2020. A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. N. Am. J. Econ. Finance 54, 101251.

Hilscher, J., Wilson, M., 2017. Credit ratings and credit risk: is one measure enough? Manage. Sci. 63 (10), 3414–3437.

Huang, Z., Chen, H., Hsu, C.J., Chen, W.H., Wu, S., 2004. Credit rating analysis with support vector machines and neural networks: a market comparative study. Decis. Support Syst. 37 (4), 543–558.

Jorion, P., Shi, C., Zhang, S., 2009. Tightening credit standards: the role of accounting quality. Rev. Account. Stud. 14 (1), 123–160.

Li, J.P., Mirza, N., Rahat, B., Xiong, D., 2020. Machine learning and credit ratings prediction in the age of fourth industrial revolution. Technol. Forecas. Soc. Change 161, 120309.

Luna, J.M., Chao, H.H., Diffenderfer, E.S., Valdes, G., Chinniah, C., Ma, G., Simone II, C.B., 2019. Predicting radiation pneumonitis in locally advanced stage II–III non-small cell lung cancer using machine learning. Radiother. Oncol. 133, 106–112.

Löffler, G., 2004. An anatomy of rating through the cycle. J. Bank. Finance 28 (3), 695–720.

Löffler, G., 2005. Avoiding the rating bounce: why rating agencies are slow to react to new information. J. Econ. Behav. Organ. 56 (3), 365–381.

Metz, A., Cantor, R., 2006. Moody's Credit Rating Prediction Model. Special Comment. Moody's Investor Service.

Standard and Poor's, 2008. Corporate Ratings Criteria (New York, NY).

Tas, E., Atli, A.H., 2022. A comparison of SVR and NARX in financial time series forecasting. Int. J. Comput. Econ. Econom. 12 (3), 303–320.

Wallis, M., Kumar, K., Gepp, A., 2019. Credit rating forecasting using machine learning techniques. In: *Managerial perspectives on intelligent big data analytics*. IGI Global, pp. 180–198.

Yang, R., Zuo, A., 2014. Identifying the credit level of a company with bagging-CART integrated algorithm. In: 2014 Seventh International Symposium on Computational Intelligence and Design, 2. IEEE, pp. 371–374.

Zmijewski, M.E., 1984. Methodological issues related to the estimation of financial distress prediction models. J. Account. Res. 59–82.