A Multiclass Model for Financial Distress and Bankruptcy Prediction

Ana Sousa(✉), Ana C. Braga, and Jorge Cunha

ALGORITMI Research Centre, LASI, University of Minho, Guimarães, Portugal
amjorgesousa@gmail.com, {acb,jscunha}@dps.uminho.pt

Abstract. Corporate bankruptcy rarely occurs as an abrupt event. Instead, it is often preceded by a period of gradual financial deterioration. However, most bankruptcy prediction studies rely on binary classification models that distinguish only between healthy and failed firms, thereby overlooking intermediate states of financial instability. This study proposes a multiclass early-warning framework that represents firms' financial condition through three staged categories: healthy, financially distressed, and pre-event, where the pre-event class corresponds to the final observable accounting year before bankruptcy. The empirical analysis uses a large firm-year panel from the Orbis database, focusing on Portuguese manufacturing firms. The raw dataset comprises 37,823 firms and 151,292 firm-year observations from 1997 to 2025. After data preparation, the final analytical dataset contains 140,053 observations. Firm-level financial indicators are combined with lagged macroeconomic variables to account for both firm-specific vulnerabilities and the broader economic environment. A multinomial logistic regression model is estimated to predict firms' financial states while preserving interpretability. The model achieves good out-of-sample performance, with a balanced accuracy of around 0.75 and a macro-averaged F1 score of close to 0.70. The results indicate that the pre-event state can be clearly distinguished from both healthy and distressed firms, whereas the intermediate distress state remains more difficult to identify. Overall, the findings provide evidence consistent with a gradual deterioration process and show that interpretable econometric models can provide useful early-warning signals for firms approaching severe financial vulnerability.

Keywords: bankruptcy prediction · financial distress · multiclass classification · early warning systems · multinomial logistic regression · business failure

1 Introduction

Corporate bankruptcy generates substantial economic costs for firms, creditors, investors, employees, and public institutions. Firm failure may lead to unpaid debt, job losses, disruption of business relationships, and broader economic spillovers [7,13,15,18]. For this reason, the early identification of financially vulnerable firms remains a central objective in financial risk assessment

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16759, pp. 308–326, 2027.
https://doi.org/10.1007/978-3-032-30518-3_20

[31]. Bankruptcy prediction has a long tradition in finance and accounting. Early studies showed that financial ratios help distinguish failing from non-failing firms [1,5,24]. Classical statistical approaches, including discriminant analysis, logistic regression, and hazard models, remain relevant because they provide transparent and economically interpretable frameworks [1,4,24,26,31]. More recent research has increasingly applied machine learning techniques to bankruptcy prediction [31]. Methods such as support vector machines, decision trees, random forests, and gradient boosting can capture complex and non-linear relationships between financial indicators and failure risk [4]. However, higher predictive flexibility often comes at the cost of reduced interpretability, which remains a relevant concern in financial decision-making contexts where stakeholders need to understand the drivers of risk [31]. Despite these developments, three limitations remain important. First, most bankruptcy prediction studies rely on binary classification frameworks that distinguish only between healthy and failed firms [1,24]. However, corporate failure rarely occurs as an abrupt event. Firms often exhibit signs of financial deterioration before bankruptcy, including declining profitability, weaker cash-flow generation, and increasing leverage [1,5,23]. Binary models may therefore overlook intermediate stages of financial instability that are particularly relevant from an early-warning perspective [19,23]. Second, although machine learning models can improve predictive accuracy, their lower transparency may limit their usefulness in settings where interpretability is required. Third, reliable early-warning systems require validation designs that respect the temporal and panel structure of firm-level data. Random cross-validation may mix observations from different time periods or from the same firm across training and testing sets, producing overly optimistic estimates of predictive performance [6,26]. Motivated by these limitations, this study develops a multiclass early-warning framework that represents firms' financial condition using three staged categories: healthy, financial distress, and pre-event, with the latter corresponding to the final observable accounting year before bankruptcy. Rather than estimating dynamic transitions between financial states, the proposed framework provides a temporally ordered representation of financial deterioration. This allows the analysis to examine whether intermediate and pre-event states can be empirically distinguished while preserving the interpretability of the modelling approach. The empirical analysis is based on a large panel dataset of Portuguese manufacturing firms obtained from the Orbis database. The raw dataset contains 37,823 firms and 151,292 firm-year observations for the period 1997–2025. After data preparation, the final analytical dataset includes 140,053 observations. Firm-level financial indicators are combined with lagged macroeconomic variables to account for both firm-specific vulnerabilities and broader economic conditions. A multinomial logistic regression model is estimated as the baseline specification, allowing the probabilities of different financial states to be modelled simultaneously while maintaining economic interpretability. The study contributes to the bankruptcy prediction literature by providing empirical evidence on the usefulness of a multiclass representation of financial distress, combining descriptive, event-time, and predictive evidence, and showing that

an interpretable econometric model with a firm-disjoint validation design can provide useful early-warning signals while reducing information leakage. The remainder of the paper is organised as follows. Section 2 presents the data, modelling framework, and validation design. Section 3 reports the empirical results and discussion. Section 4 concludes.

2 Methods

2.1 Data Source and Sample Construction

Firm-level accounting data were obtained from the Orbis database maintained by Bureau van Dijk. The sample includes Portuguese non-financial firms operating in the manufacturing sector (NACE Rev. 2). Financial institutions were excluded because their balance-sheet structure, regulatory environment, and failure mechanisms differ substantially from those of non-financial firms. The analysis was further restricted to manufacturing firms, following a common approach in bankruptcy prediction studies that focus on industrial firms [1,24]. Firm information was extracted from the *Key Financials and Employees* and *Global Standard Format* (GSF) modules. The GSF format was prioritised because it provides harmonised accounting variables that improve comparability across firms and over time. Because the Orbis interface limits the number of records that can be downloaded at once, the dataset was downloaded in several batches. All raw files were preserved unchanged to ensure traceability and reproducibility of the data preparation process. Data preparation, modelling, and figure generation were implemented in `R` using a fully scripted workflow. Orbis reports financial statements using a relative time structure (*t0*–*t3*), where *t0* corresponds to the most recent accounting year available for each firm. Calendar years were reconstructed using the accounting closing date associated with each reporting period. The dataset was then reshaped into a firm–year panel, in which each observation corresponds to a firm in one fiscal year. The raw panel spans the period 1997–2025, depending on firm reporting availability, and contains 37,823 firms and 151,292 firm-year observations. After constructing financial indicators and applying the data availability filters required for modelling, the final analytical dataset used in the empirical analysis contains 140,053 firm-year observations.

Sample Selection Considerations. Restricting the sample to manufacturing firms is consistent with prior bankruptcy prediction studies focusing on industrial firms [1,24]. This restriction improves the homogeneity of the estimation sample and facilitates a more consistent interpretation of financial ratios across firms [3]. Financial ratios may have different meanings across sectors because firms differ in asset intensity, operating cycles, working-capital requirements, leverage structures, and profitability margins. Therefore, focusing on a single sector reduces cross-sector heterogeneity and strengthens internal comparability. At the same time, this design choice limits external validity, as the results may not generalise directly to other sectors or institutional contexts.

Panel Structure. The dataset is organised as an unbalanced firm–year panel. Each observation corresponds to a firm in a specific accounting year. Firms

enter and exit the database at different points, resulting in a different number of observations per firm.

2.2 Dependent Variable: Financial Condition

The dependent variable captures firms' financial condition using three mutually exclusive states that represent different stages of financial deterioration.

- **Healthy**: firms classified as *Active* in Orbis, indicating normal operating conditions.
- **Distress**: firms classified as *Active (insolvency proceedings)*, indicating firms undergoing formal insolvency or restructuring procedures.
- **Pre-event**: firms classified as *Bankruptcy* in Orbis, for which the outcome is assigned to the last accounting year observed before the bankruptcy event. Therefore, this class represents the final pre-bankruptcy observation, not the bankruptcy event itself.

This classification framework is consistent with the view that financial distress may unfold through different degrees of deterioration before failure [28]. Importantly, the proposed framework does not estimate dynamic transitions between financial states. Instead, it provides a staged and temporally ordered representation of firms' financial condition, allowing the analysis to examine whether intermediate and pre-event states can be empirically distinguished. As is typical in bankruptcy prediction datasets, the distribution of observations across classes is highly imbalanced, with healthy firms representing the vast majority of firm-year observations and financially distressed states occurring relatively infrequently [29]. Figure 1 illustrates the temporal logic used to construct the dependent variable.

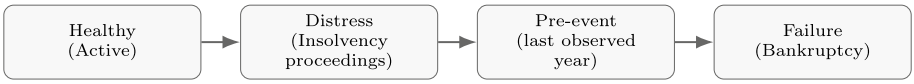

Fig. 1. Conceptual temporal logic used to construct the dependent variable.

Event Definition. The Orbis legal status variable does not always report the exact calendar year of bankruptcy. For this reason, for firms classified as *Bankruptcy*, the pre-event observation is approximated using the last available accounting year before the bankruptcy event. The bankruptcy event itself is not used as a labelled observation, reducing the risk of look-ahead bias arising from post-failure accounting adjustments or incomplete reporting.

2.3 Explanatory Variables

Predictors were selected based on a systematic review of the bankruptcy prediction literature covering 83 empirical studies [27]. The review identified the financial indicators most frequently used in prior research and ensured coverage of the main dimensions of corporate failure risk, including liquidity, profitability, leverage, cash-flow capacity, growth, and operating efficiency. The final set of variables was further refined based on analytical judgement and data availability constraints imposed by the Orbis database, ensuring consistent construction across firms and over time. Explanatory variables are grouped into three categories: (i) firm-level financial indicators, (ii) non-financial firm characteristics, and (iii) macroeconomic conditions. Firm-level variables capture internal financial vulnerabilities, while macroeconomic indicators account for the broader economic environment in which firms operate. Macroeconomic variables were merged by calendar year and included with a one-year lag, so that conditions in year $t-1$ explain firm outcomes in year t. This temporal alignment reduces the risk of look-ahead bias. Ownership-related variables were initially considered but excluded due to severe and systematic missingness for distressed and pre-event firms. In particular, including these variables would have substantially reduced the number of observations in the minority classes and could have introduced additional sample-selection bias. Financial ratios follow standard accounting definitions. Growth variables are computed as annual percentage changes, while changes in net profit and equity correspond to first differences. Table 1 summarises all explanatory variables used in the analysis.

2.4 Macroeconomic Data and Temporal Alignment

Macroeconomic indicators were collected from official public sources to capture the broader economic environment in Portugal. These variables account for aggregate conditions that may affect firms' financial stability, including economic activity, inflation, labour market conditions, financial conditions, and the business cycle [3]. All macroeconomic variables were included with a one-year lag. Since the firm-level panel covers the period 1997–2025, macroeconomic series were assembled for the period 1996–2024. This temporal alignment ensures that macroeconomic conditions observed in year $t-1$ are used to explain firm outcomes in year t, thereby reducing the risk of look-ahead bias.

2.5 Modelling Framework and Validation

Data Preprocessing. All preprocessing steps were performed using the training data only to prevent information leakage from the validation and test sets [2,8,22]. Missing values were imputed using the median calculated from the training sample [11]. Median imputation was selected as a transparent and reproducible strategy suitable for a large-scale financial dataset. However, this approach does not fully account for potentially informative missingness, which is acknowledged as a limitation of the study. Predictors were then standardised

Table 1. Explanatory variables used in the analysis.

Category	Variable
Liquidity	Current ratio
	Quick ratio
	Cash ratio
	Working capital to total assets
	Current assets to total assets
Efficiency	Total asset turnover
	Current asset turnover
	Fixed asset turnover
	Receivables turnover
	Inventory turnover
	Working capital to sales
Leverage and capital structure	Gearing
	Equity ratio
	Short-term leverage
	Interest coverage
	Long-term debt to equity
	Long-term debt to total assets
Profitability	ROA
	EBIT to total assets
	ROE
	Gross margin
	EBIT margin
	Net profit margin
	EBITDA margin
Cash flow	Cash flow to total assets
	Cash flow to current liabilities
	Cash to total assets
Growth	Asset growth
	Sales growth
	Change in net profit
	Change in equity
Firm characteristics	Firm size
	Firm age
	Added value per employee
Macroeconomic variables	GDP growth
	Inflation
	Unemployment rate
	Equity market return
	Interest rate
	Oil price growth

using z-score normalisation based on training means and standard deviations [2,8]. The same transformations were applied to the validation and test sets [2]. Financial ratio datasets often contain extreme values [11]. To reduce the influence of outliers in graphical analyses, the variables used in the event-time trajectories were winsorised at the 1st and 99th percentiles. The values used for model estimation remain unchanged.

Baseline Model. The objective of the study is to estimate the probability that firms belong to different financial states based on firm-level financial indicators and macroeconomic conditions. The baseline specification employs a multinomial logistic regression model. This model estimates the probability of each financial state as a function of the explanatory variables [17]. The healthy state serves as the reference category, meaning that coefficients are interpreted relative to financially stable firms.

Model Interpretability. Multinomial logistic regression was selected because the objective of this study is not only to classify firms, but also to preserve the economic interpretability of the early-warning framework. The model estimates class probabilities simultaneously and allows the financial characteristics associated with different stages of deterioration to be examined directly. This is particularly relevant in financial risk assessment, where stakeholders often need to understand the drivers of risk rather than rely only on predictive accuracy [21]. Although more flexible machine learning algorithms may achieve higher predictive performance in some settings, they often reduce transparency and may be less suitable when interpretability is a central requirement [4,21]. Therefore, multinomial logistic regression provides an appropriate balance between predictive performance, interpretability, and operational usefulness.

Alternative Specifications. Robustness was assessed using several complementary specifications. First, Elastic Net regularisation was applied within the multinomial logistic regression framework to perform data-driven variable selection and control for multicollinearity [32]. The penalty parameter was selected using cross-validation within the training sample. Second, principal component analysis (PCA) was employed as a diagnostic tool to analyse the structure of the financial variables and detect possible collinearity patterns [20]. The resulting components were not used as predictors in the final model. Third, reduced-variable models based on economically motivated financial indicators were estimated to evaluate the trade-off between model parsimony and predictive performance [4,21,32]. These specifications were used to assess the robustness of the interpretable modelling framework rather than to maximise predictive performance through black-box optimisation.

Validation Design. The dataset was partitioned into three mutually exclusive subsets at the firm level: training, validation, and testing [2]. Firm identifiers were used to ensure that observations from the same firm do not appear in more than one subset. This firm-disjoint design prevents information leakage and provides a more realistic evaluation of predictive performance [2,22]. Models were estimated using the training set. Model selection and hyperparameter tuning

were performed on the validation set. The test set was used only for the final evaluation of predictive performance.

Performance Metrics. Model performance was evaluated using overall accuracy, balanced accuracy, and the macro-averaged F1 score. Because the dataset is highly imbalanced, balanced accuracy and macro-F1 provide more informative measures than accuracy alone [16]. Model discrimination was additionally assessed using one-versus-rest ROC-AUC and precision-recall AUC metrics, with PR-based measures being particularly informative under class imbalance [25].

Temporal Instability and Target Drift. The prevalence of financial distress may vary over time as macroeconomic conditions change [10]. Such variation can alter the prediction task across different economic regimes and may lead to target drift [14]. Temporal diagnostics of class prevalence were therefore used to interpret differences between validation and test performance and to assess the stability of early-warning signals over time [6].

3 Results and Discussion

This section presents the empirical results of the multiclass bankruptcy prediction framework. It begins by describing the main characteristics of the sample and the differences across financial states. It then examines the macroeconomic context and event-time patterns in financial indicators prior to bankruptcy. Finally, it reports the data partition, model selection procedure, predictive performance, robustness checks, discrimination ability, and early-warning evidence based on model predictions.

3.1 Descriptive Evidence on Financial Conditions

Table 2 reports descriptive statistics by financial condition. This comparison provides an initial assessment of whether the three proposed financial states differ in terms of liquidity, profitability, leverage, cash-flow capacity, efficiency, growth, and firm characteristics.

Clear differences emerge across the three financial states. Healthy firms generally exhibit stronger liquidity, higher profitability, more stable capital structures, and greater cash-flow capacity than distressed and pre-event firms, supporting the staged classification framework proposed in this study. Liquidity differences are particularly visible in the quick and cash ratios, while profitability separates the groups through ROA, ROE, and margin indicators. Capital structure and debt-sustainability measures also point to greater vulnerability among distressed and pre-event firms, reflected in lower equity ratios, higher short-term leverage, and weaker interest coverage. Cash-flow indicators follow a similar pattern. Some variables, especially working-capital and long-term debt indicators, do not follow a strictly monotonic pattern, suggesting heterogeneity in firms' adjustment strategies and financing structures. Overall, the descriptive evidence is consistent with prior bankruptcy prediction research, which shows that firms approaching

Table 2. Descriptive statistics by financial condition.

Variable	Healthy	Distress	Pre-event
Liquidity			
Current ratio	1.78 (1.04–3.44)	1.08 (0.65–1.72)	0.97 (0.70–1.25)
Quick ratio	1.37 (0.70–2.78)	0.77 (0.39–1.31)	0.58 (0.38–0.82)
Cash ratio	0.31 (0.07–1.04)	0.04 (0.01–0.19)	0.03 (0.01–0.09)
WC/Total assets	0.21 (0.05–0.39)	0.16 (-0.01–0.36)	0.28 (0.14–0.42)
Current assets/Total assets	0.77 (0.55–0.93)	0.73 (0.48–0.91)	0.68 (0.49–0.84)
Profitability			
ROA (%)	2.45 (-0.12–8.31)	-0.54 (-11.54–1.12)	0.07 (-8.09–0.97)
ROE (%)	8.22 (1.19–21.60)	1.00 (-19.04–9.75)	0.98 (-9.10–7.28)
Net profit margin	0.02 (0–0.07)	-0.01 (-0.20–0.01)	0 (-0.08–0.01)
EBIT margin	3.76 (0.56–9.79)	0.46 (-13.24–3.83)	2.73 (-5.53–5.61)
EBIT/Assets	0.04 (0–0.10)	0 (-0.12–0.03)	0.02 (-0.06–0.06)
EBITDA margin	7.47 (2.29–14.93)	2.45 (-9.25–7.67)	6.51 (-1.23–10.74)
Gross profit margin	0.64 (0.48–0.84)	0.68 (0.50–0.94)	0.56 (0.41–0.70)
Leverage/Capital structure			
Equity ratio	0.36 (0.12–0.62)	0.12 (-0.14–0.29)	0.15 (-0.04–0.27)
Short-term leverage	0.36 (0.19–0.64)	0.60 (0.37–0.90)	0.72 (0.50–0.93)
Long-term debt/Assets	0.14 (0.01–0.31)	0.25 (0.10–0.44)	0.07 (0–0.28)
Long-term debt/Equity	0.18 (0–0.77)	0.44 (-0.05–1.70)	0 (0–0.77)
Gearing	34.40 (2.79–112.87)	132.03 (42.99–285.54)	77.83 (2.04–161.68)
Interest coverage	5.77 (1.48–24.96)	0.48 (-7.54–2.42)	1.00 (0.20–1.58)
Cash-flow			
Cash-flow/Total assets	0.06 (0.01–0.13)	0 (-0.10–0.05)	0.03 (-0.04–0.06)
Cash-flow/Current liabilities	0.17 (0.02–0.50)	0.01 (-0.14–0.09)	0.05 (-0.05–0.11)
Cash/Total assets	0.12 (0.03–0.30)	0.03 (0.01–0.11)	0.02 (0.01–0.06)
Efficiency			
Total assets turnover	1.11 (0.67–1.75)	0.86 (0.49–1.52)	1.05 (0.71–1.51)
Current assets turnover	1.65 (1.00–2.63)	1.37 (0.79–2.41)	1.73 (1.20–2.44)
Fixed assets turnover	5.38 (2.19–15.05)	4.35 (1.47–13.81)	4.03 (1.82–9.27)
Inventory turnover	8.71 (3.67–25.62)	6.23 (2.57–20.27)	5.51 (3.12–10.65)
Receivables turnover	4.98 (2.93–9.40)	3.89 (1.99–8.52)	3.43 (2.20–5.60)
WC/Sales	0.18 (0.05–0.38)	0.16 (0–0.43)	0.26 (0.13–0.48)
Growth and variation			
Total assets growth	0.04 (-0.06–0.19)	-0.01 (-0.12–0.15)	0.05 (-0.07–0.21)
Sales growth	0.06 (-0.08–0.25)	-0.04 (-0.23–0.16)	-0.01 (-0.16–0.14)
Δ Net profit	0.32 (-12.73–16.02)	-2.90 (-63.44–19.18)	-2.06 (-82.19–21.12)
Δ Equity	6.16 (-2.66–40.24)	-8.60 (-121.17–7.88)	1.71 (-89.02–25.76)
Firm characteristics			
Firm age	15 (6–27)	17 (7–31)	13 (6–23)
Employees	6 (2–15)	17 (7–38)	40 (18.75–91.25)
Added value per employee	20.63 (13.71–31.81)	13.55 (8.83–19.67)	16.25 (10.95–19.22)

Notes: Median values are reported, with 25th–75th percentiles in parentheses. Growth variables exhibit high dispersion due to firm heterogeneity.

failure tend to exhibit weaker liquidity, lower profitability, higher leverage, and reduced cash-flow capacity [1,5,24,31].

Ownership Structure. The BvD Independence Indicator was initially considered as a proxy for ownership structure. However, it is available for 98.1% of healthy firm-year observations but is almost entirely missing for distressed firms (99.6%) and completely absent for pre-event observations. This structural imbalance makes comparisons across financial states unreliable and may introduce bias; therefore, the variable was excluded from the descriptive analysis and predictive modelling framework.

3.2 Macroeconomic Context

To contextualise firm-level behaviour, macroeconomic indicators representing economic activity, labour market conditions, inflation, financial conditions, and

market performance are analysed for the period 1996–2024. Average GDP growth was 1.51%, with considerable volatility (SD = 3.07), reflecting different phases of the business cycle. The unemployment rate averaged 8.23% and reached a maximum of 16.5% during the European sovereign debt crisis. Inflation averaged 2.2%, also showing variation across the period. Public debt increased substantially over time, reaching 134% of GDP. Equity market returns and oil price growth also displayed high volatility. Figure 2 presents the evolution of real economic conditions in Portugal, measured through GDP growth and the unemployment rate. Two major shocks are particularly evident. The European sovereign debt crisis (2010–2013) was associated with a contraction in economic activity and a sharp increase in unemployment, while the COVID-19 shock in 2020 produced a sudden decline in GDP followed by a rapid recovery.

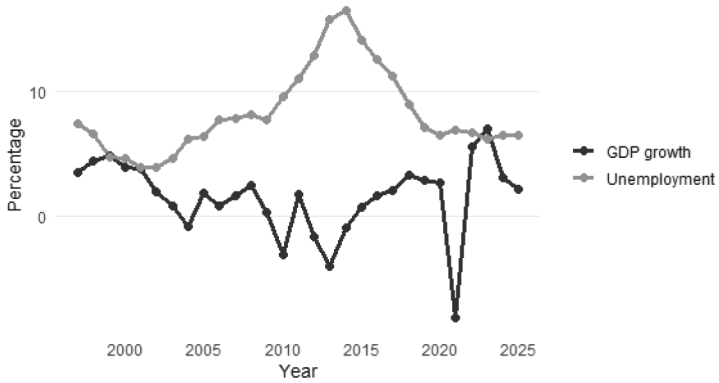

Fig. 2. Real economy in Portugal (1996–2024): GDP growth and unemployment.

Figure 3 summarises financial conditions over the same period, including public debt, equity market returns, and interest rates.

Public debt rose sharply from the early 2000s to the mid-2010s and remained high thereafter. Equity market returns were highly volatile, reflecting changing market expectations and broader financial uncertainty. Interest rates followed a long downward trend until around 2021, followed by an increase associated with post-pandemic monetary tightening. These macroeconomic patterns are relevant because firms' financial conditions are observed under heterogeneous economic regimes. Adverse macroeconomic conditions may increase corporate vulnerability by reducing demand, tightening financing conditions, and weakening firms' ability to service debt [12,30]. However, macroeconomic shocks affect firms unevenly. Their impact depends on firm-specific financial resilience, including liquidity, profitability, leverage, and cash-flow capacity. Therefore, macroeconomic variables are included as lagged controls that complement firm-level indicators rather than replace them as the main source of early-warning information. Having established the macroeconomic context, the next subsection turns

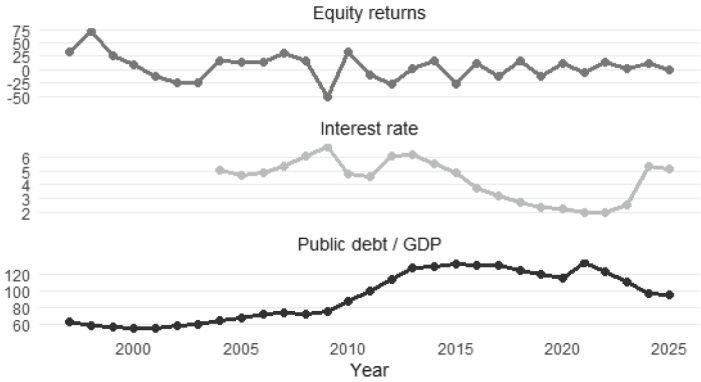

Fig. 3. Financial conditions in Portugal (1996–2024): public debt, equity market returns, and interest rates.

to firm-level trajectories before bankruptcy to examine whether financial deterioration is observable before the event.

3.3 Event-Time Financial Deterioration Prior to Bankruptcy

To examine patterns of financial deterioration before bankruptcy, an event-time analysis was conducted around the bankruptcy event. For bankrupt firms, the reference year (t_0) corresponds to the pre-event observation, defined as the last observed accounting year before bankruptcy. Observations from $t-1$ to $t-3$ represent the three preceding fiscal years. This alignment ensures that all information used in the analysis refers strictly to periods before bankruptcy and therefore avoids look-ahead bias. The notation t_0 is used only within this event-time analysis and should not be confused with the *t0* notation used by Orbis to denote the most recent accounting year available for each firm. Four representative indicators were selected to capture key dimensions of financial distress: liquidity (current ratio), profitability (ROA), capital structure (equity ratio), and operating cash-flow generation (cash flow to total assets). These indicators reflect core dimensions of firms' financial condition that have consistently been identified in the bankruptcy prediction literature as important determinants of corporate failure [1,5,24]. Because financial ratios often contain extreme values, the variables were winsorised at the 1st and 99th percentiles. Median values and interquartile ranges (IQRs) are used to summarise the trajectories. For comparison, a benchmark group of healthy firms was constructed using balanced four-year windows of observations. Each healthy firm contributes a comparable time window aligned to a pseudo-event year, allowing trajectories to be expressed relative to the same event-time structure ($t-3$ to t_0). The number of healthy firms was randomly sampled to approximate the size of the bankrupt sample. Figure 4 shows the median trajectories and interquartile ranges of the selected indicators.



Fig. 4. Event-time trajectories of selected financial indicators for bankrupt and healthy firms. Median values and interquartile ranges are shown. The dashed line marks t_0.

Bankrupt firms exhibit consistently weaker liquidity, with current ratios remaining below those of healthy firms throughout the event-time period. Clear differences are also observed in profitability. Bankrupt firms display persistently lower ROA values, approaching zero before failure, whereas healthy firms maintain positive profitability. Capital structure indicators show a similar pattern: firms approaching bankruptcy have markedly lower equity ratios, indicating weaker capital buffers. In addition, bankrupt firms exhibit lower operating cash flow, while healthy firms maintain higher and more stable cash-flow levels relative to total assets. Overall, the event-time evidence indicates that firms approaching bankruptcy exhibit weaker financial conditions several years before the event. This pattern is consistent with the view that bankruptcy is generally preceded by gradual financial deterioration, reflected in lower liquidity, declining profitability, reduced equity buffers, and weaker operating cash flows, in line with prior findings in the bankruptcy prediction literature [1,5,24].

3.4 Data Partition and Model Selection

The dataset was partitioned into three mutually exclusive subsets at the firm level: training, validation, and testing. Firm identifiers were used to ensure that observations from the same firm were not included in more than one subset. This firm-disjoint design reduces information leakage and provides a more realistic assessment of predictive performance. The training set contains 84,020 firm-year observations from 22,693 firms. The validation set contains 27,989 observations from 7,565 firms, and the test set contains 28,044 observations from 7,565 firms.

Class proportions remain highly stable across the three subsets, as shown in Table 3. In the training sample, 96.66% of observations correspond to healthy firms, 3.09% to distressed firms, and 0.25% to the pre-event condition. The validation and test sets exhibit nearly identical distributions, indicating that the partition preserves the overall class structure of the dataset. This imbalance is

typical in bankruptcy prediction contexts, where failure events are relatively rare and class distributions are highly skewed [29].

Before model estimation, a diagnostic principal component analysis (PCA) was conducted on the training data to examine the structure of the financial variables and identify possible collinearity patterns. The first principal component explains 14.2% of the variance, while the first five components together explain 42.6%. This indicates that the information contained in the predictors is distributed across several dimensions rather than being dominated by a single latent factor. PCA was used only as a diagnostic tool and not for dimensionality reduction in the final predictive model.

Variable selection was performed using Elastic Net regularisation within a multinomial framework. The model was estimated using five-fold cross-validation at the firm level on the training set, evaluating several values of the mixing parameter α. Model selection was based on predictive performance on the validation set, using the macro-F1 score as the main criterion. This metric was prioritised because it gives equal weight to all classes and is therefore more appropriate than overall accuracy in the presence of strong class imbalance.

Table 3. Sample split and class distribution.

Split	Observations	Firms	Healthy (%)	Distress (%)	Pre-event (%)
Train	84,020	22,693	96.66	3.09	0.25
Validation	27,989	7,565	96.70	3.02	0.29
Test	28,044	7,565	96.59	3.15	0.27

As reported in Table 4, the specification with $\alpha = 0.1$ achieved the highest validation macro-F1 and was therefore selected as the final model. The selected model retains 32 predictors capturing liquidity, profitability, leverage, cash flow, efficiency, firm growth, firm characteristics, and macroeconomic conditions. Test metrics are reported for transparency only and were not used for model selection, ensuring an unbiased final evaluation of predictive performance. This result is consistent with the descriptive evidence presented earlier, suggesting that financial deterioration is multidimensional and reflects both firm-level vulnerabilities and broader economic conditions.

3.5 Predictive Performance and Robustness

The test set, which was not used to estimate or select the model, was used to assess predictive performance on unseen data, following standard practice for evaluating generalisation ability [6]. Under the default classification rule, based on the highest predicted probability, the model achieves an overall accuracy of 97.17%, a balanced accuracy of 0.747, and a macro-F1 score of 0.701. These results are reported in Table 5. Because the dataset is highly imbalanced, balanced accuracy and macro-F1 provide a more informative evaluation than overall

Table 4. Elastic Net model selection results.

Specification	α	Pred.	F1 (Val.)	Bal. Acc. (Val.)	F1 (Test)	Bal. Acc. (Test)
alpha_0.1	0.1	32	0.7015	0.7306	0.7009	0.7471
alpha_0.5	0.5	29	0.6994	0.7359	0.7071	0.7482
alpha_1.0	1.0	22	0.6983	0.7338	0.7114	0.7481
alpha_0.9	0.9	23	0.6950	0.7274	0.7095	0.7435

Note: Pred. = predictors retained; Val. = validation; Bal. Acc. = balanced accuracy. Test metrics are reported for completeness only and were not used for model selection.

accuracy alone [9,22,29]. Table 5 also reports class-specific F1 scores, showing substantial differences across financial states.

Table 5. Predictive performance on the test set across alternative specifications.

Specification	Acc.	Bal. Acc.	Macro-F1	Healthy F1	Distress F1	Pre-event F1
Default model	0.9717	0.7471	0.7009	0.9855	0.3859	0.7312
Weighted model	0.8732	0.8826	0.6342	0.9303	0.2780	0.6944
Dual-threshold	0.9715	0.7620	0.7268	0.9855	0.4524	0.7425

Note: Acc. = accuracy; Bal. Acc. = balanced accuracy; F1 = F1-score. Class-specific F1 values are reported for each class.

Healthy firms are identified with a very high F1 score of 0.986, reflecting the model's strong ability to recognise the majority class. The pre-event state also achieves a relatively strong F1 score of 0.731, indicating that the model identifies firms approaching bankruptcy with useful reliability. From an early-warning perspective, this result is particularly important because the pre-event state corresponds to the final observable stage before bankruptcy. In contrast, the distress class is more difficult to detect, with an F1 score of 0.386. This weaker performance should not be interpreted only as model failure. Rather, it suggests that intermediate financial distress is empirically heterogeneous and partially overlaps with both healthy firms and firms approaching the pre-event state [3]. Distressed firms may therefore display financial profiles that are not sharply separated from neighbouring states, making this class the most challenging to identify in a staged early-warning framework. To assess robustness, two additional specifications were examined. First, a weighted multinomial model was estimated using class weights to account for class imbalance. This specification improves balanced accuracy, which increases from 0.747 to 0.883, but this improvement comes at the cost of lower overall accuracy and weaker macro-F1 performance. Specifically, overall accuracy decreases from 0.972 to 0.873, while macro-F1 decreases from 0.701 to 0.634 (Table 5). These results suggest that the weighted model over-corrects for class imbalance by favouring minority-class predictions at the expense of more false positive classifications. Second, a

dual-threshold classification rule was applied to the probabilities produced by the main model. The thresholds were calibrated on the validation set and then applied unchanged to the test set. Under this rule, overall accuracy remains almost unchanged at 0.972, while balanced accuracy increases from 0.747 to 0.762 and macro-F1 rises from 0.701 to 0.727. In particular, the F1 score for distress increases from 0.386 to 0.452, while the F1 score for the pre-event class increases slightly from 0.731 to 0.743. These results indicate that the estimated class probabilities contain useful information about the distress state, but that the default maximum-probability classification rule does not fully exploit this signal.

3.6 Discrimination Performance and Early-Warning Evidence

The model's discrimination ability was further evaluated using one-versus-rest ROC-AUC and precision-recall AUC metrics. In the test set, the ROC-AUC values were 0.895 for healthy firms, 0.886 for distressed firms, and 0.999 for the pre-event class. These results indicate strong discrimination across financial states, particularly for firms approaching the pre-event condition. Because the dataset is highly imbalanced, precision-recall performance provides additional information on minority-class detection and complements ROC-based evaluation [16,25]. The PR-AUC values are 0.994 for healthy firms, 0.432 for distressed firms, and 0.718 for the pre-event state. Together with the class-specific results reported in Table 5, these metrics confirm that the model clearly separates healthy firms from those approaching the pre-event state, while the intermediate distress class remains more difficult to identify, which is a common challenge in imbalanced classification settings [16,29]. The early-warning ability of the model was analysed by aligning observations relative to the year in which the pre-event condition occurs (t_0), corresponding to the final observed accounting year before bankruptcy. This event-time perspective makes it possible to examine whether warning signals become more frequent as firms approach the pre-event year. Table 6 reports the proportion of observations classified as either distress or pre-event in each event-time window, together with the share directly classified as pre-event.

Table 6. Event-time early-warning evidence.

Event time	N	Flagged (distress/pre-event) %	Predicted pre-event %
$t-3$	5	40.0	20.0
$t-2$	13	69.2	15.4
$t-1$	38	76.3	50.0
t_0	75	93.3	82.7

Warning signals become more frequent as firms approach the pre-event year. The share of observations classified as distress or pre-event rises from 40.0% at

$t-3$ to 93.3% at t_0, while the share directly predicted as pre-event increases from 20.0% to 82.7%. Figure 5 illustrates this progression.

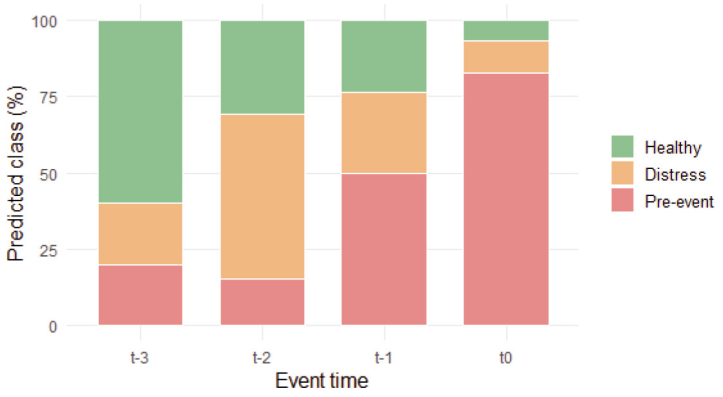

Fig. 5. Distribution of predicted states across event time relative to the pre-event year.

The earliest event-time windows contain relatively few observations and should therefore be interpreted with caution. Even so, as the pre-event year approaches, the share of healthy predictions declines, while distress and especially pre-event classifications become increasingly frequent. This pattern is consistent with the view that financial deterioration tends to emerge gradually rather than appearing only as an abrupt event [3,26]. Overall, the evidence suggests that the framework provides indicative early-warning information by distinguishing firms in more advanced stages of financial deterioration from financially healthy firms.

4 Conclusion

This study develops and evaluates a multiclass early-warning framework that distinguishes between healthy firms, financially distressed firms, and firms in the final observable pre-event state before bankruptcy. Using a large firm-year panel of Portuguese manufacturing firms, the analysis combines descriptive evidence, event-time patterns, and predictive modelling to examine whether different stages of financial deterioration can be empirically distinguished. The findings are consistent with the view that bankruptcy is often preceded by gradual financial deterioration rather than appearing only as an abrupt event. Firms approaching bankruptcy exhibit weaker liquidity and profitability, reduced operating cash flows, and smaller equity buffers. These patterns suggest that financial fragility may become observable before bankruptcy and contain relevant forecasting information. The multinomial logistic model achieves robust predictive performance, with balanced accuracy around 0.75 and a macro-averaged

F1 score close to 0.70 on the test set. The pre-event state is clearly distinguishable from both healthy and distressed firms, whereas the intermediate distress state remains more difficult to classify. This reflects the empirical heterogeneity of distressed firms and their partial overlap with neighbouring financial states. The results also suggest that firm-level financial indicators provide direct early-warning signals, while macroeconomic variables capture broader economic conditions and provide relevant contextual information. This interpretation is consistent with the idea that financial distress is closely associated with firm-specific weaknesses, although the broader economic environment remains important for understanding the conditions under which firms operate. From a practical perspective, the multiclass framework offers advantages over traditional binary bankruptcy models by distinguishing between different degrees of financial vulnerability. This distinction may help stakeholders identify firms at more advanced stages of deterioration and support earlier intervention before bankruptcy occurs. Nevertheless, the results should be interpreted in light of several limitations. Bankruptcy events are rare, leading to strong class imbalance and a limited number of observations in the early pre-event windows. Median imputation provides a transparent and reproducible missing-data strategy, but it does not fully account for potentially informative missingness. Finally, the focus on Portuguese manufacturing firms improves sectoral comparability and supports a more consistent interpretation of financial ratios, but limits external validity, as the results may not generalise directly to other sectors, countries, legal environments, or accounting systems. Future research could compare the proposed multiclass approach with modern machine learning benchmarks and apply it to other sectors and countries. Explicitly dynamic approaches, such as time-to-event models, multi-horizon prediction objectives, transition models, or dynamic model updating, could also be used to model financial deterioration more formally over time. Overall, the findings suggest that combining economically interpretable indicators with a multiclass perspective and a rigorous validation design can provide a realistic and operational early-warning framework for identifying firms at different stages of financial vulnerability.

Acknowledgments. This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UID/00319/2023.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Altman, E.I.: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Financ. **23**(4), 589–609 (1968). https://doi.org/10.2307/2978933
2. Apicella, A., Donnarumma, F., Isgrò, F., Prevete, R.: Don't push the button: exploring data leakage risks in machine learning pipelines. Artif. Intell. Rev. **58**, 1–58 (2025). https://doi.org/10.1007/s10462-025-11326-3

3. Balcaen, S., Ooghe, H.: 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. Br. Account. Rev. **38**(1), 63–93 (2006). https://doi.org/10.1016/j.bar.2005.09.001
4. Barboza, F., Kimura, H., Altman, E.I.: Machine learning models and bankruptcy prediction. Expert Syst. Appl. **83**, 405–417 (2017). https://doi.org/10.1016/j.eswa.2017.04.006
5. Beaver, W.H.: Financial ratios as predictors of failure. J. Account. Res. **4**, 71–111 (1966). https://doi.org/10.2307/2490171
6. Bergmeir, C., Benítez, J.M.: On the use of cross-validation for time series predictor evaluation. Inf. Sci. **191**, 192–213 (2012). https://doi.org/10.1016/j.ins.2011.12.028
7. Bernstein, S., Colonnelli, E., Giroud, X., Iverson, B.: Bankruptcy spillovers. J. Financ. Econ. **133**(3), 608–633 (2019). https://doi.org/10.1016/j.jfineco.2018.09.010
8. Bouke, M.A., Abdullah, A.: An empirical study of pattern leakage impact during data preprocessing on machine learning-based intrusion detection models reliability. Expert Syst. Appl. **230**, 120715 (2023). https://doi.org/10.1016/j.eswa.2023.120715
9. Dasilas, A., Rigani, A.: Machine learning techniques in bankruptcy prediction: a systematic literature review. Expert Syst. Appl. **255**, 124761 (2024). https://doi.org/10.1016/j.eswa.2024.124761
10. Duffie, D., Saita, L., Wang, K.: Multi-period corporate default prediction with stochastic covariates. J. Financ. Econ. **83**(3), 635–665 (2007). https://doi.org/10.1016/j.jfineco.2005.10.011
11. El Madou, K., Marso, S., El Kharrim, M., El Merouani, M.: Evolutions in machine learning technology for financial distress prediction: a comprehensive review and comparative analysis. Expert. Syst. **41**(2), e13485 (2024). https://doi.org/10.1111/exsy.13485
12. European Central Bank: Loss-given-default and macroeconomic conditions. ECB Working Paper Series (2954) (2024). https://doi.org/10.2866/10039
13. Franks, J., Lóránth, G.: A study of bankruptcy costs and the allocation of control. Rev. Finance **18**(3), 961–997 (2014). https://doi.org/10.1093/rof/rft020
14. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., Bouchachia, A.: A survey on concept drift adaptation. ACM Comput. Surv. **46**(4), 1–37 (2014). https://doi.org/10.1145/2523813
15. Graham, J.R., Kim, H., Li, S., Qiu, J.: Employee costs of corporate bankruptcy. J. Financ. **78**(4), 2087–2137 (2023). https://doi.org/10.1111/jofi.13251
16. He, H., Garcia, E.A.: Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. **21**(9), 1263–1284 (2009). https://doi.org/10.1109/TKDE.2008.239
17. Hosmer, D.W., Lemeshow, S., Sturdivant, R.X.: Applied Logistic Regression, 3rd edn. Wiley, Hoboken (2013). https://doi.org/10.1002/9781118548387
18. Jia, H., Li, C., Li, D.: Bankruptcy social cost, government intervention and stock price crash risk. Financ. Res. Lett. **64**, 105449 (2024). https://doi.org/10.1016/j.frl.2024.105449
19. Johnsen, T., Melicher, R.W.: Predicting corporate bankruptcy and financial distress: information value added by multinomial logit models. J. Econ. Bus. **46**(3), 269–286 (1994). https://doi.org/10.1016/0148-6195(94)90038-8
20. Jolliffe, I.T., Cadima, J.: Principal component analysis: a review and recent developments. Phil. Trans. R. Soc. A **374**(2065), 20150202 (2016). https://doi.org/10.1098/rsta.2015.0202

21. Jones, S., Johnstone, D., Wilson, R.: Predicting corporate bankruptcy: an evaluation of alternative statistical frameworks. J. Bus. Finance Account. **44**(1–2), 3–34 (2017). https://doi.org/10.1111/jbfa.12218
22. Kapoor, S., Narayanan, A.: Leakage and the reproducibility crisis in machine-learning-based science. Patterns **4**(9), 100804 (2023). https://doi.org/10.1016/j.patter.2023.100804
23. Lau, A.H.L.: A five-state financial distress prediction model. J. Account. Res. **25**(1), 127–138 (1987). https://doi.org/10.2307/2491262
24. Ohlson, J.A.: Financial ratios and the probabilistic prediction of bankruptcy. J. Account. Res. **18**(1), 109–131 (1980). https://doi.org/10.2307/2490395
25. Saito, T., Rehmsmeier, M.: The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE **10**(3), e0118432 (2015). https://doi.org/10.1371/journal.pone.0118432
26. Shumway, T.: Forecasting bankruptcy more accurately: a simple hazard model. J. Bus. **74**(1), 101–124 (2001). https://doi.org/10.1086/209665
27. Sousa, A., Braga, A.C., Cunha, J.: Mapping the evolution of bankruptcy prediction models: a systematic literature review (2026). Manuscript submitted to International Review of Financial Analysis
28. Sun, J., Li, H., Huang, Q., He, K.: Predicting financial distress and corporate failure: a review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. Knowl.-Based Syst. **57**, 41–56 (2014). https://doi.org/10.1016/j.knosys.2013.12.006
29. Veganzones, D., Séverin, E.: An investigation of bankruptcy prediction in imbalanced datasets. Decis. Support Syst. **112**, 111–124 (2018). https://doi.org/10.1016/j.dss.2018.06.011
30. Xing, K., Luo, D., Liu, L.: Macroeconomic conditions, corporate default, and default clustering. Econ. Model. **118**, 106079 (2023). https://doi.org/10.1016/j.econmod.2022.106079
31. Zhao, J., Ouenniche, J., De Smedt, J.: Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. Mach. Learn. Appl. **15**, 100527 (2024). https://doi.org/10.1016/j.mlwa.2024.100527
32. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. Roy. Stat. Soc. B **67**(2), 301–320 (2005). https://doi.org/10.1111/j.1467-9868.2005.00503.x