

Extracting information from sequences of financial ratios with Markov for Discrimination: an application to bankruptcy prediction

Volkov, Andrey; Van den Poel, Dirk

Abstract — In this paper, we propose a method that extracts information from sequences of financial ratios and investigate the usefulness of this information for bankruptcy prediction, which constitutes an important class of financial services. We use the annual financial reports available from an external financial information services provider to extract predictors based on the Markov for Discrimination (MFD) methodology. These predictors are used as inputs in a binary classification model, which applies logistic regression to estimate the odds of bankruptcy. The results suggest that MFD-based predictors can achieve substantial predictive performance in terms of the AUC and the 5-percent predictive lift, which are two relevant performance metrics in our case.

Keywords: bankruptcy prediction; sequence analysis; Markov for Discrimination; financial services.

I. INTRODUCTION

The timely detection of reliable and accurate signs of deteriorating financial health is a highly important task that has attracted great attention from business practitioners and academics alike. The recent history of the global financial crisis and, in particular, the powerful role played by the financial services providers (such as credit rating companies) in the crisis bear witness to the fact that the further refinement of bankruptcy prediction techniques remains an extremely relevant and crucial issue.

Ever since their introduction the sixties of the 20th century [1, 2], bankruptcy prediction models have been consistently becoming more sophisticated in terms of applied algorithms and the ability to handle large amounts of data that are commonly available today [3-6]. While the majority of researchers in the area of bankruptcy prediction concentrated its efforts on the development of ever more complex and better performing algorithms, a smaller number of scholars recognized that considerable marginal gains in performance of bankruptcy prediction models could be achieved by a careful selection of independent variables with superior predictive power [6, 7]. This paper is meant to make a contribution to the latter stream of research by proposing a novel class of sequential predictors based on the Markov for Discrimination methodology that, to the best of authors' knowledge, has never been applied in the context of bankruptcy prediction before.

Compared to more traditional time-varying (such as averages, trends, and standard deviations) and non-time-varying predictors, predictors based on MFD methodology are expected to be more powerful at being able to capture information in sequences (i.e., repeated measurements over

time) of financial ratios, which can be particularly useful in discriminating between healthy and failing companies. This ability to extract information from sequences helps to overcome a common point of criticism of traditional bankruptcy prediction models, which are often assume a static, "snap-shot" view on bankruptcy [4, 8]. In this respect, our study also contributes to the body of research that views bankruptcy as a process that spans a certain time period and is characterized by certain trends taking place in the financial results of an ailing company over time [8-12].

II. METHODOLOGY

A. Markov for Discrimination

1) General model.

This paper proposes to use a statistical sequence analysis technique known as Markov for Discrimination in order to collapse a time-varying sequence of ratios to one single independent variable that is further used in bankruptcy prediction models. Markov for Discrimination [13] is a probabilistic sequence analysis technique that has originally been applied to investigate biological sequences of proteins and nucleic acids. Outside biological sciences, Markov for Discrimination has been applied by Prinzie and Van den Poel [14], who analyzed purchase sequences for consumer durable goods acquisition models. In the context of bankruptcy prediction, Markov for Discrimination score can be interpreted as the odds that a company with a given sequence of financial ratio values is likely to come from a population of failed companies, compared to healthy companies. The reader is referred to [14] for the detailed description of Markov for Discrimination method.

The Markov for Discrimination technique assumes that the observed sequences are generated by two distinct populations characterized by different Markov transition models. These models describe how an abstract agent (e.g., a company) probabilistically transitions between a set of discrete states. In the case of financial ratios, a state may correspond to a particular ratio value that a firm reports in its annual accounts.

Markov for Discrimination score is real number. In our case, positive values imply that a particular sequence is more likely to be generated by bankrupt companies than the population of non-bankrupt companies. Negative values imply that a given sequence is more likely to represent non-bankrupt companies, compared to bankrupt ones. Finally, the value of zero implies that a particular sequence has equal chances to be generated by bankrupt or non-bankrupt companies.

2) Estimation

Given that financial ratios are usually real valued, it was necessary to discretize them in order to obtain discrete states for the first-order Markov transition models. We opted for the simple equal-frequency binning technique to carry out the discretization. This technique divides observations into a number of bins of (roughly) equal size based on ascending ranks of the variable. Even though a great number of different discretization schemes has been proposed in literature [15], equal-frequency binning is the discretization technique that is not particularly sensitive to outliers or uneven distributions of variable values [16]. Equal-frequency binning always results in a discrete number of non-empty, balanced bins. This property is particularly useful if one wants to estimate the transition probabilities of the Markov models with a sufficient level of accuracy, since these models are quite sensitive to estimation errors [17].

To obtain a finite number of states, each observation was assigned to one of the n bins ($n = 5$ or 10) on the yearly basis. The transition probabilities were calculated based on the maximum likelihood estimator, which uses observed empirical frequencies of state-to-state transitions to calculate transition probabilities. In addition, Laplacian smoothing was applied to avoid situations in which the estimation sample did not contain any observations of a particular transition type. In case of sequences with missing values, we imputed the value of 0 for the Markov for Discrimination score, since we did not have enough information to decide which population (bankrupt or non-bankrupt companies) the sequences with missing values were likely to represent.

B. Logistic regression and predictive performance

The main purpose of the paper is to verify the relevance of MFD predictors for bankruptcy prediction modeling. We chose simple logistic regression as a modeling technique due to its high proven applicability in the area of bankruptcy prediction modeling [4, 5, 19, 20], despite the fact that bankruptcy prediction problems have been shown to be slightly non-linear [18]. In addition, we used the stepwise variable selection techniques to obtain more parsimonious models to avoid possible over-fitting issues.

The predictive performance of the models is measured by two different predictive performance measures: (1) AUC (area under the receiver operating curve) and (2) (cumulative) predictive lift. An ideal predictive model would have the AUC close to 1, while a naive (random) classification model would have the AUC value equal to 0.5. In the case of bankruptcy prediction modeling, the AUC metric can be interpreted as the chance that a randomly drawn bankrupt company will have a higher probability of going bankrupt (as assigned by the classification model) compared to the probability of going bankrupt assigned to a randomly drawn non-bankrupt company.

Lift measures the percentage of correctly positively classified examples within the top $x\%$ of observations that are most likely to be classified as positive compared to the total percentage of positive examples in the sample. Higher values of the lift metric indicate that a model is performing substantially better than a random model at identifying truly positive cases in the top $x\%$ of cases that are most likely to be positive. In the case of bankruptcy prediction the cost of misclassifying a healthy company as non-healthy are particularly high. In addition, only a small number of companies in the population indeed go bankrupt. Hence, we argue that lift is a more relevant performance measure than AUC. The former indicates how well a model is able to separate bankrupt companies from surviving ones in the top $x\%$ of companies that are assigned the highest scores by the predictive model. For example, a model that erroneously classifies top 5% of the most likely to go bankrupt companies as bankrupt will have the value of the 5%-lift equal to 0, while the AUC of such model can still be sufficiently high.

III. EMPIRICAL RESULTS

A. Data

The initial sample available for analysis consisted of bankrupt and non-bankrupt companies, selected from the Belfirst Finance dataset (Bureau Van Dijk, 2009). The initial data set included only Belgian companies that were established in the period from January 1, 2002 until December 31, 2006. We limited the selection to smaller and medium-sized companies to obtain a sample of companies of comparable size. For bankrupt companies, it was mandated that their last reported annual accounts be dated from 2007. The initial sample contained 14 752 companies, consisting of 1 090 (7.39%) bankrupt and 13 662 (92.61%) non-bankrupt companies. For each company, the sample contained 20 financial ratios, with each ratio measured over a 5-year time period, starting in the year 2003 and ending in the year 2007. The ratios selected represented different dimensions of companies' financial health, including profitability, liquidity, solvency, value added activities and growth of asset base.

Further, we eliminated from the initial sample companies that did not provide enough data to estimate Markov for Discrimination variables. Such companies contained sequences of insufficient length or sequences with gaps, possibly caused by irregularities and omissions in the reported annual accounts. A sequence was considered to have insufficient length if the number of non-missing observations in this sequence was less than or equal to one. Such sequences do not contain actual transitions and, as such, did not contribute enough information for the estimation of Markov for Discrimination scores. In this study, it was opted to drop 4189 (7.78% bankrupt and 92.22% non-bankrupt) companies that had sequences of

insufficient length for more than 2 ratios, resulting in the data set of 10563 (7.23% bankrupt and 92.77% non-bankrupt) companies. Remaining sequences with insufficient length were assigned the value of 0 for the Markov for Discrimination score. The initial dataset also contained sequences that contained gaps, i.e. missing values between non-missing ones. While such sequences potentially contain useful transitions, it is not possible for them to estimate the Markov for Discrimination score with enough precision due to the absence of observations for one or more periods. If missing values in such sequences represent transitions that are particularly discriminative for the bankrupt versus non-bankrupt companies, the Markov for Discrimination score calculated only on the basis of non-missing transition might be erroneous. In total, there were 1 497 observations containing at least one sequence with a gap, representing 120 (8.02%) bankrupt and 1 377 (91.98%) of non-bankrupt companies. It was decided to eliminate these observations from the analysis, resulting in the final dataset that contained 9 066 (7.10% bankrupt and 92.90% non-bankrupt) companies.

In order to avoid issues with over-fitting and to validate the results, the final dataset was randomly divided into training and validation and test sets, comprising 50%30% and 20% of all observations, respectively. The training set was used to calibrate logistic regression models, while the validation set was used to validate the classification performance. The test sample was used to assess the predictive performance of the best model identified in the study. Unless stated otherwise, the performance metrics mentioned in the study are calculated on the basis of the validation set.

B. Markov for Discrimination predictors

Table I presents several performance measures (based on the validation set) of logistic regression models with only Markov for Discrimination scores as predictor variables.

Two classes of models were compared: models obtained by binning each ratio on per annum basis into 5 (MFD5) and 10 (MFD10) bins.

TABLE I.

Model	Variable selection	AUC	Lift	Cum. lift
MFD5	None	0.7500	4.698	0.235
	Stepwise	0.7441	4.398	0.220
MFD10	None	0.6827*	2.999	0.150
	Stepwise	0.6827*	2.999	0.150

Note: Column maximum in bold; * indicates 1% significance level.

In terms of AUC, the MFD5 model without variable selection achieves the highest performance (AUC = 0.75), while the MFD5 model with stepwise selection is not significantly different. All MFD10 models perform significantly worse. This difference in performance can be explained by the fact that the transition probabilities are likely to be estimated less precisely in the latter case, since the number of states increases exponentially with the number of bins. Hence, one also needs a considerably higher number of observations to estimate the transition probabilities with enough precision.

In terms of the top-5% lift, MFD5 models again perform notably better than MFD10 models. Cumulatively, the two MFD5 models are able to correctly identify 23.5% and 22.0% of all failing companies in the top 5% of observations in the validation dataset, while MFD10 models are able to identify only 15.0% of truly bankrupt companies.

IV. CONCLUSION

The objective of this study is two-fold. Firstly, we wish to contribute to the underdeveloped body of literature that tries to increase the performance of bankruptcy prediction models by concentrating on variables with more predictive power. In contrast, most of the studies dealing with bankruptcy prediction concentrate on benchmarking of different kinds of classification algorithms. Secondly, we also contribute to the body of research that views bankruptcy as a process that spans a certain time period, during which companies go through evolutions reflected in the value of their financial ratios.

We propose to use predictors based on the Markov for Discrimination technique. These predictors are able to extract the sequential information from the series of financial ratios and, thus, implicitly incorporate the view of bankruptcy as a process. To the best of the authors' knowledge, this type of predictors has never been applied in the area of bankruptcy prediction before.

The results obtained demonstrate that the predictors based on MFD scores are able to perform well in bankruptcy prediction modeling tasks. However, the precision, with which the transition probabilities of the Markov transition models are estimated, plays an important role in the final model performance. While more fine-grained discretization schemes are probably better suited to capture the complexity of the problem, they also need progressively more observations to estimate the transition probabilities with sufficient accuracy. The future research needs to investigate this trade-off between a larger number of states and the potential deterioration of the predictive performance due to insufficient number of observations. A possible way to research this problem would be to consider the choice of the parameters of a discretization scheme (e.g., the number of bins) as an optimization problem.

Another extension of the current methodology could consider different estimation methods that can be applied to obtain accurate and reliable estimates of transition

probabilities in Markov transition models (e.g., bootstrapping). Additionally, it might be interesting to look at the performance of predictors based on higher-order Markov transition models. Finally, it might be fruitful to benchmark the performance of the MFD-based predictors against more traditional sequential and non-sequential predictors.

REFERENCES

- [1] E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *The Journal of Finance*, vol. 23, pp. 589-609, 1968.
- [2] W. Beaver, "Financial ratios predictors of failure," *Journal of Accounting Research*, vol. 4 (Suppl.), pp. 71-111, 1967.
- [3] M. A. Aziz and H. A. Dar, "Predicting corporate bankruptcy: where we stand?," *Corporate governance*, vol. 6, pp. 18-33, 2006.
- [4] S. Balcaen and H. Ooghe, "35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems," *The British Accounting Review*, vol. 38, pp. 63-93, 2006.
- [5] A. I. Dimitras, S. H. Zanakis, and C. Zopounidis, "A survey of business failures with an emphasis on prediction methods and industrial applications," *European Journal of Operational Research*, vol. 90, pp. 487-513, 1996.
- [6] A. Verikas, Z. Kalsyte, M. Bacauskiene, and A. Gelzinis, "Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: A survey," *Soft Computing*, vol. 14, pp. 995-1010, 2010.
- [7] M. Niemann, J. H. Schmidt, and M. Neukirchen, "Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity," *Journal of Banking and Finance*, vol. 32, pp. 434-446, 2008.
- [8] N. Chancharat, P. Davy, M. McCrae, and G. Tian, "Firms in financial distress: a survival model analysis," in *20th Australasian Finance & Banking Conference*, 2007.
- [9] J. Argenti, *Corporate Collapse: The Causes and Symptoms*. McGraw-Hill, London, 1976.
- [10] M. Luoma and E. K. Laitinen, "Survival analysis as a tool for company failure prediction " *OMEGA International Journal of Management Science*, vol. 19, pp. 673-678, 1991.
- [11] E. K. Laitinen, "Financial predictors for different phases of the failure process," *OMEGA International Journal of Management Science*, vol. 21, pp. 215-228, 1993.
- [12] H. Ooghe and S. De Prijcker, "Failure processes and causes of company bankruptcy: a typology," *Management Decision*, vol. 46, pp. 223-242, 2008.
- [13] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison, *Biological Sequence Analysis. Probabilistic models of Proteins and Nucleic Acids*. UK: Cambridge University Press, 1998.
- [14] A. Prinzie and D. Van den Poel, "Predicting home-appliance acquisition sequences: Markov/Markov for Discrimination and survival analysis for modelling sequential information in NPTB models," *Decision Support Systems*, vol. 44, pp. 28-45, 2007.
- [15] S. Kotsiantis and D. Kanellopoulos, "Discretization techniques: a recent survey," *GESTS International Transactions on Computer Science and Engineering*, vol. 32, pp. 47-58, 2006.
- [16] E. Cantú-Paz, "Supervised and unsupervised discretization methods for evolutionary algorithms," in *GECCO-2001 Genetic and Evolutionary Computation Conference*, San Francisco, California, USA, 2001.
- [17] D. I. Simester, P. Sun, and J. N. Tsitsiklis, "Dynamic catalog mailing policies," *Management Science*, vol. 52, pp. 683-696, 2006.
- [18] T. Van Gestel, B. Baesens, and D. Martens, "From linear to non-linear kernel-based classifiers for bankruptcy prediction " *Neurocomputing*, vol. 73, pp. 2955-2970, 2010.
- [19] J. Ohlson, "Financial ratios and the probabilistic prediction of bankruptcy.," *Journal of Accounting Research*, vol. 18, pp. 109-131, 1980.
- [20] C. V. Zavgren, "Assessing the vulnerability of financial characteristics of American industrial firms: a logistic analysis," *Journal of Banking and Finance*, vol. 12, pp. 19-45, 1985.