Explainable AI for bankruptcy prediction using XBRL filing

Maria Mora-Rodriguez, *Fujitsu Laboratories of Europe*
Alicia Rodriguez-Carrion, *Fujitsu Laboratories of Europe*
Revathy Ramanan, *XBRL International*
Manuel Rejón-Lopez, *University of Granada*
Francisco Flores-Muñoz, *The Spanish Association of Accountants* (AECA)

With the 2008 global financial crisis, there has been an increased number of corporate bankruptcies at a global level, becoming its prediction in one of the key tools in financial decision-making and the main driver for credit risk. Understanding and predicting company bankruptcy has been a widely studied topic, due to the importance for many institutions to identify key factors of a firm's repayment probability. From the first qualitative models back in 1800 from the rating agencies, there has been a progressive development of ever-evolving models.

As a result, bankruptcy prediction models research is gaining primary attention not only from the academia also by creditors, regulators and investors all over the globe. The first univariate models (Beaver, 1966) allowed to start using quantitative measurements to compare a company with itself through time, and with others of the same industry. This concept continued its sophistication towards multivariate discriminant or logit models (Altman, 1968) where 5 strictly financial variables were used to determine the probability of default, over a 66 companies database, but with such good results that this method and its variants are still a reference today. Next, structural models built upon Merton's proposals (Merton, 1974) appeared which are still in use, followed mainly by accurate and practically acceptable blended models called Z-Metrics (Altman *et .al*, 2010) mainly used today together with classical approaches.

Over these years, models based in varying algorithms have been proposed, seeking to improve accuracy, the ability to foresee bankruptcy with enough anticipation, or stability in time. The main limitation of the initial models is the use of limited number and variety of variables, which could be enhanced with macroeconomic indicators, or non-financial data such as company performance data coming from the media and social networks. However, classical models are difficult to be extended with more variables and/or more complex relationships among them. To tackle these limitations, machine learning algorithms have been extensively analysed with plenty of different datasets, based on big data and/or non-traditional metrics (invoice analysis, payable history, social media data, attention to the news in media (FRISK Scoring System 2017[i])) and novel artificial intelligence-based techniques (Zheng and Yanhui (2007), Sun and Li (2012), Iturriaga and Sanz (2015), Barboza *et al.* (2017), Mselmi *et al.* (2017), Charalambakis (2017), Zelenkov *et al*. (2017), Qu *et al.* (2019), Le *et al.* (2019), Huang and Yen (2019)) such as Support Vector Machine (SVM), genetic algorithms, neural networks among others. These approaches have shown better results in terms of accuracy as compared with traditional discriminant or logistic regression models up to 10% improvement (Barboza *et al.*, 2017). However, they have still considered 'exotic' approaches and their use still limited because of the lack of trust in 'black box' solutions (Altman, 2018). Precisely, to shed light into these 'black boxes' and, thus to make it possible to understand the rationale behind the results of those algorithms the eXplainable Artificial Intelligence (XAI) literature provides several techniques for that (Friedman (2001), Ribeiro *et al.* (2016), Apley and Zhu (2016), Lundberg and Lee (2017)). Some of these techniques have been applied to a similar case of bank distress to understand, not only the distress predictions themselves but also how the variables in the model play a role in the distress probability (Suss and Treitel (2019), Joseph (2019)).

Many studies have been focused on investigating AI models for corporate bankruptcy forecasting, however, few of them are offering explanations or trust behind those AI models or “black boxes”. Most of the studies have been based on financial data source published by third party such as data aggregators and requires much of data engineering before it can be used to train the model. Financial performance information available in structured and digital format such as XBRL can bring additional value in terms of providing good quality, comparable ready to use data. XBRL reports submitted in fulfilment of reporting obligations by the companies enhances the trust worthiness of data source.t.
Our study introduces XAI models for corporate bankruptcy forecasting using more than 10 years of open data in XBRL format, concretely financial reports collected by the Spanish and Exchange Commission (CNMV) for Spanish listed companies and solvency reports. The objective of this study is twofold: 1) to construct an Explainable AI model with high transparency and trust capacity for bankruptcy prediction 2) to demonstrate the use of public XBRL filings as an efficient data source prediction. These findings are vital for driving greater transparency and trust on financial sustainability, as investors, policymakers, auditors and stakeholders of this market are looking for more accurate information about the chances of corporate bankruptcy.

References

Altman, Edward I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, *23*(4), pp.589-609.

Altman, E.I., Rijken, H., Watt, M., Balan, D., Forero, J. and Mina, J., 2010. The Z-Metrics™ methodology for estimating company credit ratings and default risk probabilities. *RiskMetrics Group*, pp.18-21.

Altman, Edward I., 2018. A Fifty-Year Retrospective on Credit Risk Models, the Altman Z -Score Family of Models and their Applications to Financial Markets and Managerial Strategies (December 11, 2018). Journal of Credit Risk, Vol. 14, No. 4

Apley, D.W. and Zhu, J., 2016. Visualizing the effects of predictor variables in black box supervised learning models. arXiv preprint arXiv:1612.08468.

Barboza, F., Kimura, H. and Altman, E., 2017. Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, pp.405-417.

Beaver, William H. 1966. “Financial Ratios as Predictors of Failures,” Journal of Accounting Research, 4, 71-111.

Charalambakis, Evangelos C., 2014. On corporate financial distress prediction: What can we learn from private firms in a small open economy?. *Working Paper No. 188, Bank of Greece, Athens, Greece.*

Friedman, Jerome H., 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics, pp.1189-1232.

Huang, Y.P. and Yen, M.F., 2019. A new perspective of performance comparison among machine learning algorithms for financial distress prediction. Applied Soft Computing, 83, p.105663.

Iturriaga, F.J.L. and Sanz, I.P., 2015. Bankruptcy visualization and prediction using neural networks: A study of US commercial banks. Expert Systems with applications, 42(6), pp.2857-2869.

Joseph, Andreas, 2019. Shapley regressions: A framework for statistical inference on machine learning models. *arXiv preprint arXiv:1903.04209*.

Le, T., Vo, B., Fujita, H., Nguyen, N.T. and Baik, S.W., 2019. A fast and accurate approach for bankruptcy forecasting using squared logistics loss with GPU-based extreme gradient boosting. Information Sciences, 494, pp.294-310.

Lundberg, S.M. and Lee, S.I., 2017. A unified approach to interpreting model predictions. In Advances in neural information processing systems, pp. 4765-4774.

Merton, Robert C., 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, *29*, 2, pp.449-470.

Mselmi, N., Lahiani, A. and Hamza, T., 2017. Financial distress prediction: The case of French small and medium-sized firms. *International Review of Financial Analysis*, *50*, pp.67-80.

Qu, Y., Quan, P., Lei, M. and Shi, Y., 2019. Review of bankruptcy prediction using machine learning and deep learning techniques. Procedia Computer Science, 162, pp.895-899.

Ribeiro, M.T., Singh, S. and Guestrin, C., 2016, August. " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 1135-1144.

Sun J. and Li Hui 2012. Financial distress prediction using support vector machines: Ensemble vs. individual. Appl. Soft Comput. 12, 8, pp.2254–2265.

Suss, J. and Treitel, H., 2019. Predicting bank distress in the UK with machine learning. Staff Working Paper No. 831, Bank of England.

Zelenkov, Y., Fedorova, E. and Chekrizov, D., 2017. Two-step classification method based on genetic algorithm for bankruptcy forecasting. Expert Systems with Applications, 88, pp.393-401.

Zheng, Q. and Yanhui, J., 2007, August. Financial distress prediction based on decision tree models. In 2007 IEEE International Conference on Service Operations and Logistics, and Informatics, pp. 1-6.

[i] FRISK Scoring System 2017 http://www.creditriskmonitor.com