CS230

Bankruptcy Prediction by Deep Learning

https://youtu.be/ZCrnqD4IqMs

by Giovanni Cialone – Winter 2020 - gcialone@stanford.edu

This Italian case study pursues the goal of developing a commercial firms insolvency prediction model. In compliance with the Basel II Accords, its major objective is an estimation of the probability of default (PD) over a certain time horizon, typically one year. The model predicts the firms that are going to fail within one year, using deep fully connected layers and Cnns.

Data

The present research utilized the company's AIDA database - Bureau Van Dijk, a Moody's Group company. After pre-processing, 14.966 Italian micro-small firms have been selected: 13,846 active and 1,120 bankrupted.

ABRUZZO	2.796	SICILIA	6.456
CALABRIA	2.148	TOSCANA	11.125
CAMPANIA	10.400	TRENTINO ALTO ADIGE, SARDEGNA	5.236
EMILIA ROMAGNA	13.753	UMBRIA	2.351
FRIULI V.G.	3.460	VAL D'AOSTA, LIGURIA MOLISE, BASILICATA	5.061
LAZIO	12.621	VENETO	15.654
LOMBARDIA	26.570	PUGLIA	6.628
MARCHE	4.885	PIEMONTE	10.502
Total not failed companies	139.646	Total failed companies	1.579

Features

The selected variables (inputs), as are 88. They are the most meaningful in terms of capacity of pointing out the critical issues related to a firm financial and economic equilibrium in the long term. According to the literature, the chosen variables are closely related to gauge liquidity, profitability, financial solidity and operating performances, namely, the liquidity ratios, EBITDA, ROE, ROI and, among others, debt ratios. Refer to Table 2

Table 2. Extracted financial variables

Current Assets EUR_Year - 2	Return on Sales Year -2
Current Assets EUR_Year - 1	Return on Sales Year -1
Date of Last Available Balance Sheet	Return on Investiment Year -2
Bank Debt to Sales Year -2	Return on Investiment Year -1
Bank Debt to Sales Year -1	Gross Sales Year -2
Debt/EBITDA ratio Year -2	Gross Sales Year -1
Debt/EBITDA ratio Year -1	Gross working capital Turnover (times) Year -2
Total Debt to Equity Year -2 ratio	Gross working capital Turnover (times) Year -1
Total Debt to Equity Year -1 ratio	Invested Capital Turnover (times) Year -2
Receivables Average Collection Period -Year -2	Invested Capital Turnover (times) Year -1
Receivables Average Collection Period -Year -1	Operation Headquarters - Region
Payables Average Settlement Period -Year -2	Total Fixed Assets Year -2
Payables Average Settlement Period -Year -1	Total Fixed Assets Year -1
Ebidta to Interest Expenses Year -2	Total Liabilities and Equity Year -2
Ebidta to Interest Expenses Year -1	Total Liabilities and Equity Year -1
Working Capital to Revenues Year -2	Total Current Liabilities Year -2
Working Capital to Revenues Year -1	Total Current Liabilities Year -1
Current Ratio Year -2	Net Income after Taxes Year -2
Current Ratio Year -1	Net Income after Taxes Year -1
Total Fixed Tangible Assets to Equity Year -2	Profits/Losses Year -2
Total Fixed Tangible Assets to Equity Year -1	Profits/Losses Year -1
(Equity+Long Term Debts)/Fixed Assets Year -2	Equity to Total Assets Year -2
(Equity+Long Term Debts)/Fixed Assets Year -1	Equity to Total Assets Year -1
Current Debts to Total Debts Year -2	(Assets – Inventories)/Debts Year -2
Current Debts to Total Debts Year -1	(Assets – Inventories)/Debts Year -1
Long Term Debts to Total Debts Year -2	Long Term debt Year- 2
Long Term Debts to Total Debts Year -1	Long Term debt Year- 1
Interest Expenses to Gross Sales Year -2	Total Assets Year- 2
Interest Expenses to Gross Sales Year -1	Total Assets Year -1
Total Equity Year -2	EBITDA Year-2
Total Equity Year -1	EBITDA Year-1
Net Financial Position Year -2	EBITDA /Gross Sales Year -2
Net Financial Position Year -1	EBITDA /Gross Sales Year -1
Company Name	Total credits Year-2
Total Assets to Equity Year -2	Total credits Year-1
Total Assets to Equity Year -1	Total debt Year-2
Return on Equity Year -2	Total debt Year-1
Return on Equity Year -1	Number of employees Year-2
Return on Assets Year -2	Number of employees Year-1
Return on Assets Year -1	

Three Models

1) **The Baseline**: A simple sequential fully connected model has been used as benchmark. The shape of the input layer is (88 x 19,383).

2) **The Deep Sequential Model:** This model has the input shape as the Baseline. The architecture is much more complex : 01 input layer with shape (88, 19383), 17 inner layers, 512 neurons each and 1 output layer. This model totals 262,656 parameters of each of the layers from 2 to 17 and 45,568 parameters for the inner layer 1. Adding 513 parameters of the output layer, the final number of trainable parameters is 4,248,577.

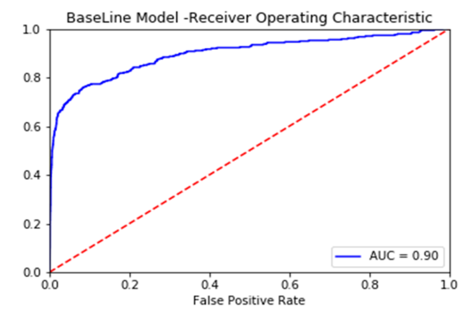
3) **The CNN Model:** To apply the CNN it is necessary to modify the data structure to obtain a 3D matrix. In fact, these expect that each unit corresponds to a data matrix. Actually, the data are obtained observing two years, so we have in total 88 variables, 44 for time -1 and 44 variables for time -2. So, each row of the original training set (88, 19385) has been reshaped in an array of dimensions ((7, 7) 2) for 19383 samples.

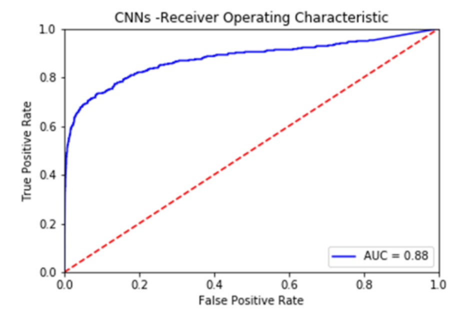
Results

True Positives	242	275	243
False Positives	93	60	92
Sensibilty	0.7224	0.8205	0.7254
Specificity	0.9116	0.8148	0.9162
Accuracy	0.8975	0.8148	0.9018
Avg.Loss	5.5021	0.3063	1.9490
AUC	0.85	0.90	0.88
Max Accuracy	0.9496	0.9424	0.9510
Avg Accuracy	0.8708	0.8794	0.9010
Table 3. Comparison Matrix			

Table 3. Comparison Matrix

Sensitity is the major objective of the models, the metrics refers to the TPs (True Positives); AUC is the Arera Under the CURVE (ROC CURVE).





Discussion

In this paper, the large amount of data for small and medium-sized Italian companies collected from financial and income statements have been processed , applying two different Neural Networks architectures: (i) **a deep sequential model** and **(ii) a Convolutional architecture**, using a simple a very simple sequential one as a baseline. The results obtained show that all models, including the baseline, achieve good results, probably due to the good quality of the data. The model with the best performances was the Sequential Architecture which reached the highest AUC value, 0.90 and the highest senibility 0.8205. The CNN Architecture showed the best specificity (numbers of True Positives caputered). The Sequntial model captures 270 out 335 True positives.

Future

It's very likely that these architectures will provide, in a future wider investigation, more interesting results. It is worth noting that the results obtained in this paper show a predictive capacity of the applied methods higher than that of similar works in the literature, that generally use only listed companies. On the contrary, this approach is completely independent of market values and can be applied to small and medium-sized enterprises. Ultimately, the models can find wider application, not only to the italian case but also to other countries where accounting standards are similar and the inputs variables have same metrics.

REFERENCES

[1] FitzPatrick, P. J. (1932). A comparison of ratios of successful industrial enterprises with those of failed companies. The Certified Public Accountant, 598-605 (Oct), 656662 (Nov) , 727-731 (Dec).
[2] Beaver W. (1966). Financial ratios as predictors of failure. Journal of Accounting 5(5), 71-111.
[3] Altman E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate. The Journal of Finance 23(4), 589-609.
[4] Altman E. I. (2000). Predicting Financial Distress of Companies: Revisiting The Z-score and ZETA models. Stern School of Business, New York university, pp. 1-54.
[5] Jodi L., Gissel D.G. (2007). A Review of Bankruptcy Prediction Studies: 1930 to present. Journal of Financial Education, Vol. 33 (Winter 2007), 1-42.
[6] Daniel T. (1968). Discriminant: analysis for the prediction of business failures. PH.D. dissertation, University of Alabama.
[7] Richard H.G., Jackson, A.W. (2013). The performance of insolvency prediction and credit risk. The British Accounting Review, 45, 183-202.
[8] Ohlson J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research 18(1), 109-131.
[9] Zavgren C.V. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic Analysis. Journal of Business Finance & Accounting 12(1), 19-45.
[10]Bell T.B., Ribar G.S., Verchio J.R. (1990), Neural nets vs. logistic regression: A comparison of each model's ability to predict commercial bank failures, in Auditing Symposium X: Proceedings of the 1990 Deloitte & Touche/University of Kansas Symposium on Auditing Problems, pp. 29-53.
[11]Black F., Scholes M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81.
[12]Merton R. C. (1974). On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance, 29(2), 449–470.
[13]Angelini E., di Tollo G., Roli A. (2007). A neural network approach for credit risk evaluation. The Quarterly Review of Economics and Finance. Elsevier, v. 48(4), pp. 733-755.
[14]Jackson, R.H.G., Wood, A. (2013) The performance of insolvency prediction and credit risk models in the UK: A comparative study. The British Accounting Review 45, 183-202 (2013).
[15]Zhang Y., Wang S., Ji G. (2013). A Rule-Based Model for Bankruptcy Prediction Based on an Improved Genetic Ant Colony Algorithm, Vol. 2013, A. ID 753251. Hindawi Publishing Corp. - Mathematical Problems in Engineering.
[16]Barboza F., Kimura H., Altman E. (2017) Machine learning models and bankruptcy prediction. Expert System with Applications, 83, 405-417.
[17]Le H.H., Viviani J.L. (2018). An improvement by implementing a machine-learning approach to classical financial ratios. Research in International Business and Finance, 44 (2018) pp. 16-25
[18]Hosaka T. (2018). Bankruptcy prediction using image financial ratios and convolutional neural network. Expert Systems with Applications, Volume 117, 1 March 2019, pp. 287-299 .
[19]Breiman L. (2001). Random Forests, Machine Learning. 45 (1): 5–32.
[20]Friedman J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, Vol. 29, No. 5, pp. 1189-1232
[21]Goodfellow I., Bengio Y., Courville A. (2016). Deep Learning, series Adaptive Computation and Machine Learning, MIT press, ISBN-13: 978-0262035613
[22]Mai F., Tian S., Lee C., Ma L. (2018). Deep learning models for bankruptcy prediction using textual disclosures. European Journal of Operational Research (2018)
[23]Alaka H. (2017). Big data analytics for construction firms insolvency prediction models. Bristol: University of West of England.
[24]Falkenstein E.G., Boral A., Lea V.C. (2000). RISKCALC for private companies: MOODY's default model. Moody's.
[25] Haibo He et al. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning
[26] Chawla n. et al. (2002). SMOTE: Synthetic Minority Over-sampling Technique