Bankruptcy Forecasting Using Case-Based Reasoning: an Italian Case Study

Fabio Sartori[1] and Alice Mazzucchelli[2]

[1] Department of Informatics, Systems and Communication,
University of Milano-Bicocca,
Viale Sarca 336/14, 20126, Milano, Italy
sartori@disco.unimib.it

[2] Department of Business Administration, Finance, Management and Law,
University of Milano-Bicocca,
Via Bicocca degli Arcimboldi 8, 20126, Milano, Italy
alice.mazzucchelli@unimib.it

1 Abstract

Over the last decade, the globalization phenomenon has determined the growing integration of Markets, as well as a concrete structural modification of enterprises' capability to compete. The increasing level of uncertainty and the complexity of the competitive scenario imply the need for continuous investments to improve both process and product innovation: many firms are not able to revise them, with the consequence to close from one year to another. The recent global market crisis, as demonstrated by recent cases of enterprise bankruptcy, has amplified this problem: in 2013, according to CERVED data[3], 111000 business closure have been registered in Italy, one of the countries where the financial crisis is more significantly affecting both the labour market and the enterprise competitiveness.

In this scenario, how to forecast enterprise bankruptcy has become an important and multidisciplinary research trend. According to [1], prediction models can be divided into three main categories: *Statistical models*, which include the well known *Multiple Discriminant Analysis* and *Logit* approaches, *Artificial Intelligence Expert System models* (AIES) and *Theoretical models*. Among AIES methods, Case Based Reasoning [2] is cited as a possible approach.

In this presentation, we want contribute to the debate presenting the initial results of a research project conducted on the Italian Small and Medium Enterprises (SMEs) that aims at identifying potentially bankruptcy firms before their situation becomes critical; given a database of inactive firms (i.e. the case base) and a database of active enterprises, the goal of the system is comparing the second dataset with the first one, in order to intersect them and discover firms potentially moving from the active state to the inactive one. To this scope, a general-purpose CBR platform, namely CREPERIE (see [3]), has been adopted

[3] https://www.cervedgroup.com/c/document_library/get_file?uuid=01c76c98-b328-4d79-bda0-184f4b4cbedf&groupId=20536

in the analysis phase. With respect to statistical and traditional methods based on the Z"-Score [4], CBR results to be more objective, since it works on uniformly weighted balance indexes.

Our sample was constructed from 2012 income statement and balance sheet information stored in the AIDA database, a Bureau Van Dijk (BvD) database containing demographic information, geographical location, industry, financial statement, balance sheet and assessment (at least five years) of more than one million of Italian enterprises, mainly limited companies. As for the 14,218 Monza and Brianza enterprises, we have extracted demographic information, legal status, balance sheet, income statement and any information about bankruptcy procedures activated in the considered year. This original dataset was divided based on the legal status of the enterprises into two distinct subsets: *bankrupted enterprises* and *not–bankrupted enterprises*: in our research, the first have triggered a bankruptcy procedure prescribed by Italian law in 2012; on the other hand, the latter have not activated any procedure and therefore they are considered healthy.

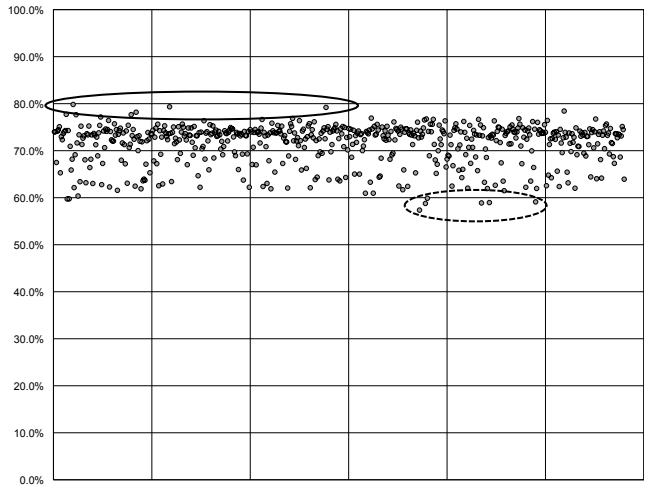

Fig. 1. Overall similarity (on the vertical axis) calculated by CREPERIE: active enterprises (on the horizontal axis) are clustered into three groups

The CREPERIE platform has divided the chosen 580 active enterprises into three subsets; the firms with higher similarity (highlighted by a solid oval in the Figure) should be close to bankruptcy, the ones with lower similarity (highlighted by a pointed oval in the Figure) should be completely active. In order to demonstrate this, the enterprises have been compared with the risk rating

assigned them by DFKA[4]; this rating is calculated according to the analysis of enterprise balance and belong to three main categories: *low risk* (ratings AAA, AA+, AA, A+, A), *medium risk* (ratings BBB, BB+, BB, B+, B), *high risk* (ratings CCC, CC+, CC, C), *default* (rating D).

Table 1. Comparison between enterprises' clusters according to CBR Overall similarity and their ratings

Enterprise ID	*Overall similarity*	*Rating*	*Enterprise ID*	*Overall similarity*	*Rating*
Enterprise A	57.4%	AAA	Enterprise I	78.4%	D
Enterprise B	58.8%	AA+	Enterprise J	78.2 %	D
Enterprise C	58.9%	AA+	Enterprise K	68.1%	BB+
Enterprise D	59.0 %	AAA	Enterprise L	68.2%	BB
Enterprise E	59.1 %	AAA	Enterprise M	68.4%	BB
Enterprise F	79.8%	D	Enterprise N	68.6%	B+
Enterprise G	79.3%	D	Enterprise O	68.7 %	CCC
Enterprise H	79.2%	CC	Enterprise P	68.9 %	CCC

Table 1 shows the correspondence between CBR results and DFKA ratings: the lowest similarity enterprises (from A to E, i.e. the enterprises in pointed oval in Figure 1) are completely sane, and their default risk is substantially null; the five highest similarity firms (from F to J, i.e. the enterprises in solid oval in Figure 1) are very close to default. The remaining six enterprises (from K to P) have been selected from the cloud of enterprises whose situation is not clear: their similarity value put them close to the half of the cloud, that is the region between 59.1% (i.e. the similarity value of E) and 78.2% (i.e. the similarity value of J). Also these six enterprises confirm the correspondence between CBR results and the rating; moreover, the rating decreases according to the similarity increasing, from BB+ (Enterprise K) to CCC (Enterprise P). Then, currently active firms can be further investigated to identify proper actions to avoid bankruptcy in the next future, by comparing their balance indexes with the ones of the most similar failed enterprises.

References

1. Aziz, M.A., Dar, H.A.: Corporate governance. Journal of Contemporary Accounting and Economics **6** (2006) 18 – 33
2. Kolodner, J.: Case–Base Reasoning. Morgan Kaufmann Publisher, Inc., 2929 Campus Drive, Suite 260, San Mateo CA 94403 (1993)
3. Manenti, L., Sartori, F.: Metadata support to retrieve and revise solutions in case-based reasoning. IJMSO **6** (2011) 185–194
4. Altman, E.I., Danovi, A., Falini, A.: Z-score modelsÕ application to italian companies subject to extraordinary administration. Journal of Applied Finance **23** (2013)

[4] http://www.dfka.it/area-crediti/rating/