UDC 62-50

I. Zahirska

BAYESIAN NETWORK APPLICATION POSSIBILITIES FOR CORPORATE STATE MODELLING

Introduction

Accurate bankruptcy prediction tools are indispensable for successful and efficient financial management of any firm. Corporate bankruptcy does not only cause substantial losses to the business community, but also to different societal strata. In today's dynamic economic environment, the number and the magnitude of bankruptcy filings are increasing significantly. Even auditors, who have good knowledge of firms' situations, often fail to make an accurate judgment on firms' going-concern conditions. Therefore, bankruptcy prediction models have become important decision aids for auditors, creditors, and stockholders. Bankruptcy prediction models are of critical importance to various stakeholders (i.e. management, investors, employees, shareholders and other interested parties) as they may warn in time about the possible corporate failure. From a managerial perspective, financial failure forecasting tools allow to take timely strategic actions so that financial distress can be avoided. For other stakeholders, such as banks, efficient and automated credit rating tools allow to detect clients that are to default their obligations at an early stage. Hence, accurate bankruptcy prediction tools will enable various participants of the market process to increase the efficiency their core activities [1]. Bankruptcy prediction often involves incomplete information on some predictors. In this way, there has always been a strong need for a powerful tool to take into account all the relevant information available and to fill the gaps of missing information.

Techniques employed to develop bankruptcy prediction models have evolved from the simple univariate analysis (Beaver, 1966) and multiple discriminant analysis (Altman, 1968), to logit and probit models (Ohlson, 1980; Zmijewski, 1984), to neural network models (Tam and Kiang, 1992), rough set theory (McKee, 1998), discrete hazard models (Shumway, 2001), Bayesian network (BN) models (Sarkar and Sriram, 2001), and genetic programming (McKee and Lensberg, 2002).

Sarkar and Sriram (2001) developed BN models for early warning of bank failures. They found that both a naive BN model and a composite attribute BN model have comparable performance to the well-known induced decision tree classification algorithm. Some other techniques, such as rough set theory, discrete hazard models, and genetic programming, have also been introduced to the field of bankruptcy prediction. Research has shown that BNs perform well as a classification and prediction tool in different domains [2].

Problem definition

The aim of this study is to examine the possibilities of applying BN method for modeling and forecasting the corporate state, to define the advantages and drawbacks of using this method for corporate analysis, and to describe the possible ways of methodology development further on.

Problems under consideration

There exists an issue on creating a method for accurate probabilistic deduction in BN basing on the learning samples. To calculate the probabilities of the node states, instead of conditional probability tables, a matrix of empirical values of joint probability distribution of the whole network is used. The main goals are achieving the dependence of calculation speed and the size of the sample, absence of necessity of designing the network structure beforehand, and implementation simplicity [3].

There is still a lack of proper guidance in the selection of variables and the discretization of continuous variables. First, there exists a large pool of potential bankruptcy predictors, including various financial ratios, stock market information, industry level factors, etc. A method is needed to guide the selection of variables that can be used to develop a well-performing naïve BN for bankruptcy prediction. The work [2] proposes such a heuristic method based on the assumption of linear dependence as measured by correlations between variables. The proposed method aims at identifying key predictors and eliminating redundant or irrelevant ones. Secondly, BN models generally use discrete-valued variables. Through discretization, continuous variables are converted into discrete variables with several states. It is unclear whether and how the number of states into which continuous variables are discretized have an impact on BN models' performance. There remains an issue whether modelling continuous variables with continuous distributions instead of discretizing these variables can improve the model's performance.

As material supplies may be also crucial for corporate state, BNs could be thought of as useful to help engineers to fulfil the purposes of preventive maintenance. BNs provide an efficient way to represent the degradation process of an industrial system or machine. A great interest of BNs is to provide an efficient tool for modelling in a simple and readable way the most probable links between events of different nature (expert opinion, feedback experience, etc.) using conditional independence between random variables. BNs which can also be regarded as a way to introduce randomness in influence diagrams can be expected to be useful for preventive maintenance in the future. Here comes another difficulty. In order to compute the joint probability of the variable, the classical BN approach consists of asking for the marginal probabilities of the entry variables and the conditional probabilities of the other variables knowing all the possible combinations between their parents. In the maintenance modelling context, the experts are asked to provide the marginal probabilities of all the variables and not only of the parent variables as in the classical approach [4].

BN principles

BNs represent a type of probabilistic networks where getting new knowledge about probabilities on the nodes is implemented using Bayes formula and its generalizations. BNs are graphical models that combine elements of both graph and probability theory. Broadly stated, BNs are directed acyclic graphs (DAGs) with a set of probability tables. A BN encodes the probability distribution of a set of random variables by specifying a set of conditional independence assumptions together with a set of relationships among these variables and their related joint probabilities. DAGs model the set of relationships among the variables; each of a DAG nodes represents a variable and the arcs are the causal or influential links between the variables. A set of conditional probability functions associated with each node model the uncertain relationship between the variable and its parents. BN conditional independence assumptions yield more compact models than those based on full joint probability distributions, relaxing the issue of computational complexity when considering a large number of variables. This visual layout provides powerful knowledge representation formalism.

BNs are locally structured, meaning that each node interacts with just its parent nodes. For discrete random variables, the conditional probability distributions are a set of tables, expressing a multinomial distribution.

There are two possible scenarios to consider when learning a BN from a sample of N cases: (1) the structure of the BN is known and parameters must be estimated; and (2) the structure of the BN is not known and, as a result, it must be found and the parameters estimated [5].

Formally BN is a pair, $< G, B >$, where the first component G is a DAG that corresponds to the random variables and that is recorded as a set of independence conditions: each variable is independent from its parents in G. The second component of the pair, B, is a set of parameters defining the network.

BNs are featured by the existence of probable cause-effect relations between the nodes. In the graph nodes we get the probabilities that are used to calculate new probabilities in the casual nodes, as can be defined from the observed result nodes. Basis of the calculation algorithm is formed by the Bayes formula and its derivatives, depending on the network topology.

BNs make up graphical models of events and processes basing on the combination of some probability theory and graph theory results. In the literature BNs may be sometimes met under such names as Bayesian belief networks, causal networks, and probabilistic networks.

BN is a convenient tool for describing quite complicated processes and events having uncertainty. They have shown up special usability for developing machine learning algorithms. The main idea of graphical model creation is unit conception, meaning the possibility of dividing a complex system into the simple elements. To unite the separate elements into a system probability theory results are used. They provide the model with a practical usability in general, as well as with an opportunity of connecting the graphical models to databases. Such graph-theoretical approach gives the researcher an opportunity to model the processes from the range of strongly interacting variables, as well as to create data structures for the further development of efficient algorithms for their processing and making decisions.

BNs are most relevantly used for diagnostics problems in the wide sense: defining the probable illness by observed symptoms, finding the possible alarm signal source in object security systems, unexpected device behaviour definition in an adjusted system, corporate state estimation and forecasting, etc. Analytic system simplicity is achieved by the simple network topologies, small number of nodes and using the only function – diagnostics function.

Quite obviously, one of the first applications of BNs was of a medical purpose. For example, Quick Medical Reference (QMR) system was developed in 1980 to provide a therapist with a database of statistic and empirical data to facilitate diagnosing. There also other systems like PathFinder, Child, Munin etc. [3]. The work [6] proposes a method of investigating the influence of computer working environment on a human being. There also exist military systems based on BNs (like “Operation Dardanelles” used to model the tactic task of maintaining coastline security from the seaside), space systems (like Vista developed for NASA Mission Control Centre), IT (e.g. help and support wizards, anti-spam filters), image and video processing, system security systems for industries [4].

BN application peculiarities

Among different techniques, BN models have many attractive features. One of the main benefits of using BNs is the possibility of taking into account both discrete and continuous variables, uncertainties, and virtual absence of variable number restrictions [3]. Moreover, BNs are easy to interpret. Users report that BN representations are quite intuitive and easy to understand. They perform well as a classification tool, have no restriction on variables’ underlying distributions, and have no requirement of complete information. Contemporary studies focus on one type of BN models: naive Bayes, which are simple to implement and have been shown to perform well in bankruptcy prediction.

Unlike most regression techniques, BNs do not have any requirements on the underlying distributions of variables. BNs can easily model complex relationships among variables including partial mediators and ‘‘interaction effects’’. BNs do not require complete information for observations. Observations that have some missing variables can still be used to train or test BN models. This feature is valuable for bankruptcy studies because bankruptcy samples are usually small and bankrupt firms tend to have missing information. BNs are dynamic and interactive. They can easily be updated with new information as it is learned. Subjective human knowledge can easily be incorporated into models. Compared to other machine learning techniques, such as neural networks, BN models are more transparent and intuitive because relationships among variables are explicitly represented by the direct acyclic graph [2].

BN is a powerful tool to model associations between relevant variables of a problem. This kind of modelling requires the intervention of experts. Most simple rules are normally used to develop the network structure, to evaluate the probabilities that are needed for inference, and to keep the most reliable probabilities, required to compute the joint probability of the network. Marginal and conditional probabilities are determined first from operating experience and secondly from expertise. There is a problem to define a strategy to avoid a too heavy and too unstable acquisition of expert information [4].

Based on a 10-fold analysis, the naive BN consisting of eight selected variables have an average prediction accuracy of 81.12% for the bankruptcy sample and 81.85% for the non-bankruptcy sample. This prediction accuracy is appealing given the difficult nature of bankruptcy prediction. Bankruptcy prediction often involves incomplete information on some predictors. There is an issue how to select second-order variables that can compensate for missing information on selected predictors. Empirical evidence does not show a significant improvement upon models’ performance by incorporating second-order variables. Similar results are observed even after we restrict sample firms to those with at least two first-order variables missing. Nevertheless, the results obtained do not deny the possible superiority of the cascaded model over the naive model in situations where missing information on first-order variables are even more substantial. Secondly, we investigate the impact on a naïve Bayesian model’s performance of the number of states into which continuous variables are discretized. The naive Bayesian model consists of eight variables, six of which are continuous. Using an average training sample size of 801 bankruptcies and 6239 non-bankruptcies, the author finds that the model performance is the best with the six continuous variables being discretized into two or three states. When the number of states is increased to four or more, the model’s performance deteriorates, probably due to over-fitting. Finally, the performance of the naive Bayes model with continuous variables being discretized and the performance of the model with continuous variables being modeled with normal distributions are compared. Results show that replacing discretization with probability density functions does not increase the model's performance. On the contrary, modelling continuous variables with normal distributions leads to a significant decrease in predicting non-bankruptcy sample. We also experiment to identify and use the best-fit distributions for continuous variables. The results are substantially similar to those using the normal distribution. One potential explanation is that normal distributions do not represent variables' underlying distributions very well. More importantly, the above reported results could also be applicable to contexts other than bankruptcy prediction. However, the sample proportion of bankruptcies used in this study is larger than the realistic population proportion of bankruptcies, which leads to the ignorance of the prior during our study process. There are other important bankruptcy predictors which are not examined by the study. Finally, this study focuses on only one type of BN models: naïve Bayes. Future research is also needed to explore how to better apply other types of BN models to bankruptcy prediction [2].

Conclusions

Financial activities of firms or industries are accompanied by gathering business information that needs to be processed, and there is a necessity of making decision based on this data. BN models are implemented in some software complexes that can, for instance, analyze risks and forecast portfolio profitability, as well as investment project liability.

One of the main benefits of using BNs is the possibility of taking into account both discrete and continuous variables, uncertainties, and virtual absence of variable number restrictions. BNs work with quantity and quality market indicators, dynamic input of new information, using obvious dependence between existent factors that influence financial indices. BNs are also especially easy to interpret. Still there are issues about choosing the first-order variables and filling the gaps of missing information, as well as finding the strategy of getting reliable expert information.

In general, BNs perform well as a tool for describing complicated processes, thus, it can be successfully used for bankruptcy prediction.

References

1. Tony Van Gestel, Bart Baesens, Johan A.K. Suykens, Dirk Van den Poel, Dirk-Emma Baestaens, Marleen Willekens. Bayesian kernel based classification for financial distress detection // European Journal of Operational Research 172 (2006) p.979–1003.

2. Sun L., Shenoy P.P., 2006. Using Bayesian networks for bankruptcy prediction: Some methodological issues // European Journal of Operational research 180 (2007) p.738-753.
3. Терентьев А.Н., Бидюк П.И., Коршевнюк Л.А. Алгоритм вероятностного вывода в байесовских сетях // Материалы IX международной научно-технической конференции “Системный анализ и информационный технологии” (САІТ-2007), г. Киев, 15–19 мая 2007. – К.: Екмо, 2007. – С. 76.
4. G. Celeux, F. Corset, A. Lannoy, B. Ricard. Designing a Bayesian network for preventive maintenance from expert opinions in a rapid and reliable way // Reliability Engineering and System Safety 91 (2006) p.849–856.
5. Eitel J.M. Laurıa, Peter J. Duchessi. A methodology for developing Bayesian networks: An application to information technology (IT) implementation // European Journal of Operational Research 179 (2007) p.234–252.
6. Міщенко Н.М., Феліжанко О.Д., Щоголева Н.М. Програмні засоби захисту користувачів від шкідливого впливу комп’ютерного середовища // Математичні машини і системи. – 2003. – № 1. – С. 95-1.