A Data Mining Model for Detecting Financial and Operational Risk Indicators of SMEs

Ali Serhan Koyuncugil, and Nermin Ozgulbas

Abstract—In this paper, a data mining model to SMEs for detecting financial and operational risk indicators by data mining is presented. The identification of the risk factors by clarifying the relationship between the variables defines the discovery of knowledge from the financial and operational variables. Automatic and estimation oriented information discovery process coincides the definition of data mining. During the formation of model; an easy to understand, easy to interpret and easy to apply utilitarian model that is far from the requirement of theoretical background is targeted by the discovery of the implicit relationships between the data and the identification of effect level of every factor. In addition, this paper is based on a project which was funded by The Scientific and Technological Research Council of Turkey (TUBITAK).

Keywords—Risk Management, Financial Risk, Operational Risk, Financial Early Warning System, Data Mining, CHAID Decision Tree Algorithm, SMEs.

I. INTRODUCTION

Small and medium enterprises (SMEs) are playing a major and increasing role in the transition economies of Central and Eastern Europe. When overall world economy is considered, it is noted that SMEs, which constitute 95% of total enterprises in the world, account for 66% of total employment and 55% of total production [1]. According to the 6th Annual Report of the European Small Business Observatory (2005) there are 19.3 million of the enterprises in the European Union, 99.8% of them are defined as SMEs and employ approximately 75 million people. There are only 35,000 enterprises, with more than 250 employees, but 18 million enterprises employ fewer than 10 people [2]. SMEs constitute 99.89% of total enterprises and play a major role in Turkish economy. According to data from 2002 General Industrial Enterprise Census on SMEs in Turkey conducted by Turkish Statistical Institute (TSI), SMEs employ approximately 3 million people in total. SMEs account for 77% of total employment in Turkey and 38% of total value added generated in Turkey [3].

The majority of the transition economies have acknowledged that SMEs are an engine of economic growth and a source of sustainable development, crucial for industrial restructuring and for the generation of new jobs, creating income for the population; although unemployment is still an increasingly problem among these countries. In Turkey, economic stagnation, high inflation, instability, frequent change in economic measures, inability to compete with firms, which have high competitive power as well as factors, such globalization, alignment with the EU acquis, and the Basel-II result in failure and performance degradation of SMEs. The past years of isolation from international business, the lack of entrepreneurship culture, managerial shortcomings caused by lack of professional managers and skilled employees, deficiencies in management information systems, difficult access to credits, and the inhospitable policies and bureaucratic regulations make this process even more difficult and risky to start up and beginner entrepreneurs. SMEs represent the backbone of the Turkish economy, constituting an important contribution to GDP and to the sustainability of employment levels. But SMEs expose too much different kind of risks like listed above and financial and operational performance of SMEs are affected by these risks.

The effective management and identification of risk is a complicated task, as well as a fundamental element of business success or failure. Nowadays, Basel II Capital Accord, which will become effective in 2009, has been the center of attention of both credit companies and SMEs in Turkey. Basel II is the second of the ‘Basel Accords’, which are recommendations on banking laws and regulations issued by the ‘Basel Committee on Banking Supervision’. The purpose of Basel II, which was initially published in June 2004, is to create an international standard that banking regulators can use when creating regulations about how much capital banks need to put aside to guard against the types of risks banks face. Basel-II, which introduces risk-based capital management and risk-based credit pricing, would negatively/positively affect amount/price of loans to be utilized by SMEs. Under Basel-II banks will from now on not only consider financial risk of SMEs but also operational risk thereof before they grant loans to SMEs. With minimum financial and operational risk SMEs would get higher ratings from banks and independent auditing institutions thereby increasing their chances to receive loan facilities with more favorable conditions and minimum costs [4].

It was argued that the new capital requirements for lending to SMEs could be too high since, possibly, the value of the asset correlation for that type of firm was substantially lower than that for larger firms. An excessive level of capital might have discouraging effects on the willingness of banks to grant
loans to SMEs. In particular, a substantial increase in the
wedge between regulatory capital needs for SMEs versus
those for larger firms might result in a shortage of finance for
the former. SMEs are expected to experience problems in
receiving an appropriate credit rating and in obtaining low-
cost credits from the banks, unless they take the necessary
steps to improve their corporate structure and use effective
risk management techniques. Benefits of risk management can
summarize as early warning to avoid distress, road maps for
good credit rating, better business decision making, and
greater likelihood of achieving business plan and objectives.

Most of studies and reports on SMEs in Turkey emphasize
that insufficiencies of capital and lack of financial
management are the most important problems like Europe and
other countries. Turkish SMEs have to take risks to grow but
manage risks to protect the business. These findings and
trends in sector indicated that SMEs need risk management to
deal with not only Basel-II, but also increasing globalization,
negotiations towards full accession to the EU, financial and
economic regulations introduced in the markets where SMEs
are operating.

Therefore, this study was designed as a project to bring out
the financial and operational risk factors into open as early
warning signals for SMEs in Turkey. Our SME project is the
biggest study as covered amount and also first study that
designed a data mining model for financial and operation risk
in Turkey. The Scientific and Technological Research Council
of Turkey has been funded this project.

This paper presents a data mining model to SMEs for
detecting financial and operational risk indicators by data
mining. The identification of the risk factors by clarifying the
relationship between the variables defines the discovery of
knowledge from the financial and operational variables. Automatic and estimation oriented information discovery
process coincides the definition of data mining. During the
formation of model, an easy to understand, easy to interpret
and easy to apply utilitarian model that is far from the
requirement of theoretical background is targeted by the
discovery of the implicit relationships between the data and
the identification of effect level of every factor.

II. METHODOLOGY AND MODEL

The risk model has five phases. Data for model collected in
the first and second phases of study and these phases
completed. We have studied on the other three phases.
Completed and other phases of the study and methodologies
are presented below.

The first phase of study consists of the process of
determining the financial variables of SMEs which will be
used in data mining model. Such data obtained by means of
financial analyses of balance sheets and income statements of
companies available through Turkish Central Bank. Available
data of all firms collected between the years of 1994-2006.

In this phase firstly, firms classified in accordance with
EU’s SME criteria on basis of amount of annual turnover.
When classifying the firm’s annual turnovers thereof will
convert into Euro at average FX rate of Central Bank of the
Republic of Turkey (CBRT) and firms were classified into
four categories as follows:
1. Micro firms,
2. Small enterprises,
3. Medium enterprises, and
4. Large-scale firms [5].

After the classification firms conforming to the SME
definition of EU, financial ratios calculated as a financial risk
factors. Also, financial performance level of SMEs determined
by ratio analysis. Table I shows the ratios and definitions that
will be use in the first step of model.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>VARIABLES AND THEIR DEFINITIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratios</td>
<td>Definition</td>
</tr>
<tr>
<td>Return on Equity</td>
<td>Net Income / Total Assets</td>
</tr>
<tr>
<td>Return on Assets</td>
<td>Net Income / Total Equity</td>
</tr>
<tr>
<td>Profit Margin</td>
<td>Net Income / Total Margin</td>
</tr>
<tr>
<td>Equity Turnover Rate</td>
<td>Net Revenues / Equity</td>
</tr>
<tr>
<td>Total Assets Turnover Rate</td>
<td>Net Revenues / Total Assets</td>
</tr>
<tr>
<td>Inventories Turnover Rate</td>
<td>Net Revenues / Average Inventories</td>
</tr>
<tr>
<td>Fixed Assets Turnover Rate</td>
<td>Net Revenues / Fixed Assets</td>
</tr>
<tr>
<td>Tangible Assets to Total Liabilities</td>
<td>Tangible Assets / Long Term Liabilities</td>
</tr>
<tr>
<td>Days in Accounts Receivables</td>
<td>Net Accounts Receivable / (Net Revenues /365)</td>
</tr>
<tr>
<td>Current Assets Turnover Rate</td>
<td>Net Revenues / Current Assets</td>
</tr>
<tr>
<td>Fixed Assets to Long Term Liabilities</td>
<td>Fixed Assets / Long Term Liabilities</td>
</tr>
<tr>
<td>Tangible Assets to Equities</td>
<td>Tangible Assets / Equities</td>
</tr>
<tr>
<td>Long Term Liabilities to Constant Capital</td>
<td>Long Term Liabilities / Constant Capital</td>
</tr>
<tr>
<td>Long Term Liabilities to Total Liabilities</td>
<td>Long Term Liabilities / Total Liabilities</td>
</tr>
<tr>
<td>Current Liabilities to Total Liabilities</td>
<td>Current Liabilities / Total Liabilities</td>
</tr>
<tr>
<td>Total Debt to Equities</td>
<td>Total Debt / Equities</td>
</tr>
<tr>
<td>Equities to Total Assets</td>
<td>Total Equity / Total Assets</td>
</tr>
<tr>
<td>Debt Ratio</td>
<td>Total Debt / Total Assets</td>
</tr>
<tr>
<td>Current Account Receivables to Total Assets</td>
<td>Current Account Receivables / Total Assets</td>
</tr>
<tr>
<td>Inventories to Current Assets</td>
<td>Total Inventories / Current Assets</td>
</tr>
<tr>
<td>Absolute Liquidity</td>
<td>(Cash+Bank+ Marketable Sec.+ Acc. Rec.) / Current Liab.</td>
</tr>
<tr>
<td>Quick Ratio (Liquidity Ratio)</td>
<td>(Cash+ Marketable Sec.+ Acc. Rec.) / Current Liab.</td>
</tr>
<tr>
<td>Current Ratio</td>
<td>Current Assets / Current Liabilities</td>
</tr>
</tbody>
</table>

Operational data which couldn’t be access by balance
sheets and income statements such financial management
requirements of SMEs, training and skills of employees and
managers, process and systems in SMEs collected in phase 2.
To collect data information to be used as operational risk
factors data in model, a questionnaire designed and
information collected from SMEs via a field study.

In the third phase, qualitative and quantitative data to be
obtained through phases 1 and 2 will be analyzed with data
mining. The main approach for model is discovering different financial and operational risk factors, levels and early warning signs. Therefore, the model should focus segmentation methods. In the scope of the methods of data mining, we can be the principal methods, in addition to this several classification/segmentation methods can be mentioned [6], [7], [8], [9], [10]. However, during the development of model for SMEs, one of the basic objectives is to help SME administrators and decision makers, who do not have financial expertise, knowledge of data mining and analytic perspective, to reach easy to understand, easy to interpret, and easy to apply results about the risk condition of their enterprises. Therefore, decision tree algorithms that are one of the segmentation methods can be used because of their easy to understand and easy to apply visualization.

Decision tree algorithms were suitable for profiling because they are visual and easy-to-understand, easily interpretable, and they allow establishment of rules. With the series of rules obtained from decision trees would be possible to create profiles of firms and then classify firms in terms of levels of financial distress by using such profiles. Therefore, the most important risk indicators of financial distress signals as an early warning can be determined for each profile.

There are different decision tree algorithms. In the late 1970s J. Ross Quinlan introduced a decision tree algorithm named ID3. ID3 picks predictors and their splitting values based on the gain in information that the split or splits provide. ID3 was later enhanced in the version called C4.5. Classification and Regression Trees or CART, a relatively new and popular non-parametric analysis technique, was used after these algorithms. Another equally popular decision tree technology to CART is CHAID or Chi-Square Automatic Interaction Detector. CHAID is similar to CART in that it builds a decision tree but it differs in the way that it chooses its splits. Instead of the entropy or Gini metrics for choosing optimal splits the technique relies on the chi square test used in contingency tables to determine which categorical predictor is furthest from independence with the prediction values [6], [9].

One of the most important differences between CHAID and the other methods is tree generating. ID3, C 4.5 and CART generate binary trees, whereas CHAID can generate nonbinary trees. CHAID works with all types of continuous or categorical variables. However, continuous predictor variables automatically categorized for the purpose of the analysis. By means of Chi-Square metrics CHAID is able to separately segment the groups classified in terms of level of relations. Therefore, leaves of the tree have not binary branches but as much branches as the number of different variables in the data (Berson et al., 2000) [11]. Hence; the method of CHAID is used in the scope of this study.

In phase 4, the fitness and availability of the model for necessities of SMEs will be tested. One of the most important reasons of financing problem that SMEs encounter in Turkey is shortcomings in financial management. Therefore this phase has a vital importance to design model in a manner suitable for use of SMEs’ managers.

In the last phase, design of the model will be revised according to the findings of the phase 4. After revising the model will be finalized. Our model can be used to detect financial and operational risk indicators of SMEs. It also gives the early warning signs for financial distress.

III. IMPLEMENTATION

Implementation of the study is realized according to the data flow diagram of the risk model which is given below in Fig. 1. Implementation is given below by phases.

In Phase 1 predefined indicators was computed from database of Central Bank of The Republic of Turkey (CBRT). CBRT database is included financial data of 143.594 companies belong to years 1992-2006 which is shown in Table II below. Table II is showed the distribution of companies belong to years 1992-2006 which is shown in CBRT database of Central Bank of The Republic of Turkey (CBRT). There are 96.179 large scale, 29.829 medium sized, 10.319 small sized, 7.267 micro sized companies’ financial data belong to years 1992-2006 years in CBRT database. Predefined financial indicators were computed in sectoral level.

In Phase 2 operational data which couldn’t be access by balance sheets and income statements such financial management requirements of SMEs, training and skills of employees and managers, process and systems et. To collect data to be used as operational risk factors data in model, a questionnaire designed and information collected from SMEs was used in the scope of this study.

In the last phase, design of the model will be revised according to the findings of the phase 4. After revising the model will be finalized. Our model can be used to detect financial and operational risk indicators of SMEs. It also gives the early warning signs for financial distress.

In Phase 2 operational data which couldn’t be access by balance sheets and income statements such financial management requirements of SMEs, training and skills of employees and managers, process and systems et. To collect data to be used as operational risk factors data in model, a questionnaire designed and information collected from SMEs via a field study. Field survey was conducted as a Pilot study and it was implemented in OSTIM Organized Industrial Zone in Ankara-Turkey. There were 6.185 companies in OSTIM and the field survey was designed in complete coverage. Therefore, 6.185 companies were taken into consideration in field survey and 6.110 companies were gave response the
questionnaire with face to face interviews. Response rate for the field survey 98.8 % and field survey was statistically significant for the representation of OSTIM.

In Phase 3 financial and operational data will be combined together in central database. Financial indicators will be added to operational records in sectoral level. Therefore, all company records will have both financial and operational risk indicators together. CHAID Decision Tree Algorithm will be used on that combined data for development of risk model.

As mentioned in part 2 above Phase 1 and Phase 2 was completed and Phase 3 is still on going. After completion of all Phases it is planned that the risk model defined will be turned into software for SMEs.

IV. CONCLUSION

As mentioned in part 2 above Phase 1 and Phase 2 was completed and Phase 3 is still on going. After completion of all Phases it is planned that the risk model defined will be turned into software for SMEs.

REFERENCES


Ali Serhan Koyuncugil, M.Sc., Ph.D. is working as a statistician for Capital Markets Board of Turkey. He had his licence, M.Sc. and Ph.D. degrees in Statistics from Ankara University Department of Statistics.

His current research interests are design and development of fraud detection, risk management, early warning, surveillance, information, decision-support and classification systems, design and development of datawarehouses and statistical databases, development of indicators, models and algorithms, conducting analysis on capital markets, finance, health, SME’s, large scale statistical researches (e.g. census), population and development, socio-economic and demographic affairs based on data mining, statistics, quantitative decision making, operational research, optimization, mathematical programming, fuzzy set, technical demography theory and applications. He took part in a lot of international and national projects (UN, IBRD, EU etc.).

Dr. Koyuncugil took part in a lot of international and national conferences as an organizer, reviewer and advisor. He is member of the IASC and IASS sections of International Statistical Institute, Turkish Statistical Association, Turkish Informatics Society and was former vice head of Turkish Statisticians Association.

Nermin Ozgulbas, M.Sc., Ph.D. is associate professor of finance at Baskent University in Turkey.

She taught finance, financial analysis and cost accounting at the Departments of Health Care Management, Finance, Banking, Accounting, Risk Management and International Trade of Baskent University. She also taught finance and cost accounting at distance learning programs of Administration of Health Care Organizations in Anadolu University and Health Care Management in Ministry of Health. Her research and publication activities include finance, accounting, cost accounting, risk analysis, in health care organizations, capital markets and SMEs.

Dr. Ozgulbas has publications, presentations and projects in subject areas including the topics mentioned above. Some of the journals published her articles are: The International Journal of Health Planning and Management The Business Review Cambridge, Journal of Economy, Business and Finance, Journal of Productivity, The Health Care Manager, Journal of Accounting and Finance, World of the Accounting and Finance Journal, Journal of Health and Society. She took part in a lot of international and national conferences as an organizer, reviewer and advisor.