

Preface

Actual global situation is forced to predict the condition of achievement and should be ready to decrease the probable risk in all areas. It is possible to do that with conventional human based approach but it takes long time and the compensation of damage becomes impossible. Improvement of information technologies reflects supervision as surveillance systems and early warning systems (EWS). Operational logic of early warning systems is based on finding unexpected and extraordinary behaviors in subject area. On the other hand, data mining is the way of uncover previously unknown, useful and valuable knowledge, patterns, relations from big amount of data via sophisticated evolutionary algorithms of classical techniques such as statistics, pattern recognition, artificial intelligence, machine learning. The definitions of EWS and data mining given lead an interesting similarity. Therefore, in this book, early warning systems has been taken into consideration in both theoretical and practical mean for mainly detecting risk in financial, economical, social, security, marketing issues.

This book is a multidisciplinary study. The main components of the book are data mining, early warning systems, Information Technologies and risk management. Furthermore, this book is included the intersection and combination of these components in problematique domains. The book is mainly structured on data mining. Data mining is the most efficient and evolutionary area in Business Intelligence domain. Usage of data mining has been diffusing all areas which have big amount of data. In the book, main application area of data mining methods is early warning systems. Early warning system is a mechanism which provides proactivation via identification of critical limit values, crucial indicators, utility functions, and important factors. It is possible to mention a lot of types of early warning systems but this book is focused on early warning sytems concerning risk detection. Therefore, risk management and risk analysis or shortly risk is the last key component for this book. Risk formally defines as the expected value of the loss function. There are mainly two ways to construct a risk measure. One of them is defining a loss or utility function and the other is finding indicators which will imply risk. In this book, both of the two ways is taken into consideration while early warning system design and risk functions or indicators obtained via data mining. Furthermore, all systems are designed in IT based or aided nature. In another words, the main objective of this book is drawing a frame for risk detection via data mining and then put it in a early warning systems via Information Technologies.

The key concepts covered in this book have a very large application area in real world. Because, increasing competition has been becoming a result of globalisation and its tendency showed that it will increase very rapidly day by day. Last global financial crisis and its extensions showed that all firms should be aware of their potancial risks and should be more proactive to produce their own solutions by themselves. On the other hand, one of the most important further reflections of financial crisis is social risk for governments. In addition, prevention of crime, proactivation for terrorist attacks can be the actual

application domains in security risk perspective. Furthermore, natural disasters such as earthquakes, storms etc. and disaster recovery systems can be count in the subject area, whether included in this book or not.

It can be seen that risk detection oriented early warning systems have a very large implementation domain from the picture given above. Furthermore, last generation Business Intelligence approach data mining accelerated the accuracy of those systems. Therefore, this book will be triggered new research domains such as finance, sociology, criminology, security, earth sciences etc. Although, there is no existing title concerning early warning systems based on data mining for risk detection. However, there are some studies (papers, proceedings etc.) on risk detection based on data mining but there is no book completely dedicated on risk detection from data mining point of view. In addition, there is no book specialized on early warning systems based on data mining. Therefore, this book is the first book in a manner of complete coverage with data mining, risk detection and early warning components.

The target audience of this book includes numerous individuals, students, academics, researchers, engineers, professionals from government and non-government institutions working in the field of data mining, knowledge discovery, risk management, IT based supervision, fraud detection, matters involving early warning in various disciplines such as information and communication sciences; insurance, banking and finance; health; social sciences; criminology; security; earth sciences; engineering etc. Therefore, this book can be identified as a practical guide of theoretical matters for solving real life problems with feasible potential systems, models and examples. From that view of point it can be used in public libraries, research organizations and academic institutes' libraries. In addition, individuals such as scientists, practitioners, managers and experts from government and non-government institutions involving risk detection will be interested in this book.

This book is divided into four sections. First section entitled 'Theoretical and Conceptual Approach to Early Warning Systems' introduces basic principals of data mining, early warning systems, risk evaluation and detection in multi dimensional structure.

Chapter 1 presents a formal definition of knowledge discovery in databases (KDD) process and DM, their functions and methods, used or likely to be used in early warning systems. It also presents a brief survey of overview and application papers and software in the early warning system literature.

Chapter 2 introduces the comparison of laws on data privacy protection. In this chapter, the comparison of EU comprehensive laws model and US sectoral laws model that arise from different cultural and historical background have been presented. The main objectives are to compare the current state of consumer's privacy protection in EU and USA, discuss legal frameworks, propose some best practice implications, and summarize perceived future trends.

Chapter 3 deals with the divergence in statistical estimations from statistical learning point of view. In this chapter some of the approaches presented which open possibilities for the reduction of the huge gaps in modern statistical estimations of the same phenomena and its linkage with statistical learning. In addition, a solution has been given for create a single number of standards of economical information and economical indicators based on total conventional decisions via datawarehouse and data mining logic for clean, comparable and standardized definitions instead of directed ones for acceptable estimations and reliable conclusions.

Chapter 4 gives a review of recent developments in sequential surveillance and modeling of default probabilities of corporate and retail loans, and relates them to the development of early warning or quick detection systems for managing the risk associated with the so-called "black swans" or their close relatives, the black-necked swans.

The second section entitled ‘Early Warning Systems for Finance’ introduces early warning systems for detection and prevention of financial crisis, stock market crashes and bankruptcies.

Chapter 5 introduces a financial early warning system that all enterprises in need which detects signs to warn against risks and prevent from financial crisis. Before the global financial crisis that began 2008, small and medium-sized enterprises (SMEs) have already fought with important financial issues. The global financial crisis and the ensuring flight away from risk have affected SMEs more than larger enterprises. When these effects considered, besides the issues of poor business performance, insufficient information and insufficiencies of managers in finance education, it is clear that early warning systems (EWS) are vital for SMEs for risk detection and prevention from financial crisis. The aim of this study is to develop a financial EWS for risk detection via data mining. For this purpose, data of SMEs listed in Istanbul Stock Exchange (ISE) is processed with Chi-Square Automatic Interaction Detector (CHAID) Decision Tree Algorithm. By using this EWS, the risk profiles and risk signals have been determined for risk detection and road maps have been developed for risk prevention from financial crisis.

Chapter 6 focuses on building a financial early warning system (EWS) to predict stock market crashes by using stock market volatility and rising stock prices. The relation of stock market volatility with stock market crashes is analyzed empirically. Also, Istanbul Stock Exchange (ISE) national 100 index data used to achieve better results from the view point of modeling purpose. A risk indicator of stock market crash is computed to predict crashes and to give an early warning signal. Various data mining classifiers are compared to obtain the best practical solution for the financial early warning system. Adaptive neuro fuzzy inference system (ANFIS) model was proposed to forecast stock market crashes efficiently. Also, ANFIS was explained in detail as a training tool for the EWS.

Chapter 7 introduces a comparison of bankruptcy prediction performances of new and advanced machine learning and statistical techniques. It is very important for financial institutions which are capable of accurately predicting business failure. In literature, numbers of bankruptcy prediction models have been developed based on statistical and machine learning techniques. In particular, many machine learning techniques, such as neural networks, decision trees, etc. have shown better prediction performances than statistical ones. However, advanced machine learning techniques, such as classifier ensembles and stacked generalization have not been fully examined and compared in terms of their bankruptcy prediction performances. The aim of this chapter is to compare two different machine learning techniques, one statistical approach, two types of classifier ensembles, and three stacked generalization classifiers over three related datasets.

Chapter 8 introduces a surveillance system for bankruptcy risk of Romanian SMEs. The small and medium enterprises (SMEs) represent the backbone of the economy, playing a major economic and social role in the process of developing a dynamic economy. But the recent evolutions in the financial markets, the international financial crisis, and the increased competition on markets, the lack of financial resources and the insufficient adaptation of many firms to the requests of the European market are new threats which can determine the bankruptcies of the Romanian SMEs. In this context, starting from the necessity to design an early warning system, authors elaborated a new model for analysis of bankruptcy risk for the Romanian SMEs that combine two main categories of indicators: financial ratios and non-financial indicators. Analysis based on data mining techniques (CHAID) in order to identify the firms’ categories accordingly to the bankruptcy risk levels. Through the proposed analysis model authors tried to offer a real surveillance system for the Romanian SMEs which can allow an early signal regarding the bankruptcy risk.

The third section entitled ‘Early Warning Systems for Detection and Prevention of Fraud, Crime, Money Laundering and Terrorist Financing’ introduces early warning systems for proactivation, detection and prevention of security, criminal and terrorist issues.

Chapter 9 introduces a fraud detection system in social aids for social risk mitigation which has a poverty map construction facility. One of the most important concerns of social policies is social risk mitigation and fight against poverty and social aids as its extensions. In general, measurements of social events have been mostly based on subjective statements. More specifically, targeting mechanisms have been using for determination of potential social aid owners. Most popular targeting mechanisms are subjective ones as well. In this chapter, an objective targeting mechanism model and a fraud detection system model have been developed via data mining for social aids as an identifier of poverty levels which includes early warning signals for inappropriate applications. Then, these models have been used for development of a poverty map. Developed new targeting mechanism which has been based on rating approach will be an alternative to Means Test and Proxy Means Test. In addition, social aid fraud detection system has automatic update property with Intelligent System approach and the poverty map computation facility which can be used for absence of detailed data. Furthermore, Millenium Development Goals, Targeting Mechanisms, Poverty and Poverty Maps concepts have been reviewed from an analytical and objective point of view.

Chapter 10 introduces the generations of video surveillance systems and their applications in potential risk and crime detection. Moreover, as the surveillance video and data for safety and security are very important for all kinds of risk and crime detection, the system is required not only to data protection of the message transmission over Internet, but also to further provide reliable transmission to preserve the visual quality-of-service (QoS).

Chapter 11 discusses the application of data mining in the field of economic crime, or corporate fraud. The classification external versus internal fraud is explained and the major types of fraud within these classifications are given. Aside from explaining these classifications, some numbers and statistics are provided. After this thorough introduction into fraud, an academic literature review concerning data mining in combination with fraud is given, along with the current solutions for corporate fraud in business practice. At the end, a current state of data mining applications within the field of economic crime, both in the academic world and in business practice, is given.

Chapter 12 explores the operational data related to transactions in a financial organisation to find out the suitable techniques to assess the origin and purpose of these transactions and to detect if they are relevant to money laundering. Authors’ purpose is to provide an AML/CTF compliance report that provides AUSTRAC with information about reporting entities’ compliance with the *Anti-Money Laundering and Counter-Terrorism Financing Act 2006*. Authors’ aim is to look into the Money Laundering activities and try to identify the most critical classifiers that can be used in building a decision tree. The tree has been tested using a sample of the data and passing it through the relevant paths/scenarios on the tree. It is proposed that a decision tree using the classifiers identified in this chapter can be incorporated into financial applications to enable organizations to identify the High Risk transactions and monitor or report them accordingly.

The fourth section entitled ‘Early Warning Systems for Customer Services and Marketing’ introduces early warning systems for customer satisfaction and promotions.

Chapter 13 introduces The Learning Management System data and the subsequent Customer Interaction System data can help to provide “early warning system data” for risk detection in enterprises. This chapter provides data from an international research project investigating on customer satisfaction in

services to persons of public utility, like (education) training services and health care services, by means of explorative multivariate data analysis tools as Ordered Multiple Correspondence Analysis, Boosting regression, Partial Least Squares regression and its generalizations.

Chapter 14 introduces two models analysis – quantity (SCAN*PRO) and market share (MCI) and their power for explanatory and forecasting research using POS data for price promotions. Having dealt with more than 30 brand categories within a wider research, authors conclude that the models developed are usable for a fast decision making process within a company.

It can be seen that risk detection oriented early warning systems have a very large implementation domain from the picture given above. Furthermore, last generation Business Intelligence approach data mining accelerated the accuracy of those systems. Therefore, this book will be triggered new research domains such as data mining, statistical data analysis, statistical learning, business intelligence, information technologies, finance, banking, economics, governance, sociology, criminology, security, marketing etc. and their intersections. It is intended that this book will be a primary reference in all areas which need supervision and early warning approach for risk reduction via latest Business Intelligence method as Data Mining.

Ali Serhan Koyuncugil
Nermin Ozgulbas
Editors