# Loan Payment Prediction Using Rough Sets

BAHADTIN RUZGAR[1], NURSEL SELVER RUZGAR[2]
[1]Banking and Insurance School, Actuaries Department, Marmara University
Goztepe Campus, Kadikoy, 34722, Istanbul, TURKEY
bruzgar@marmara.edu.tr, http://mimoza.marmara.edu.tr/~bruzgar/

[2]Vocational School of Social Sciences, Marmara University
Beyazit Campus, Eminonu, Istanbul, TURKEY
nruzgar@marmara.edu.tr

*Abstract:* - In small business loans, there is always a risk for nonpayment or non-refunding of loans though very detailed examinations are made about the company. In this study, behaviors that increase the risk in loans or causing non-refunding are tried to be determined by using the rough set approach. This paper will demonstrate that rough sets model is applicable to a wide range of practical problems pertaining to loan payment prediction. Moreover, the results show that the rough sets model is a promising alternative to the conventional methods for financial prediction.

*Key-Words:* - Loan payment failure, Prediction, Rough sets, Classification

## 1 Introduction

The financial system plays a crucial role in economic development as responsible for the allocation of resources over time and among different alternatives of investment by pricing the postposition of consumption and pricing the risk [1]. Decompositions of balances on financial systems affect markets badly. Therefore, these may force sound banks to go to bankrupt, large or small business to go to failure, or loans payments to go to failure. In this paper, we propose an approach to predict small business loan payment failure based on genetic programming (GP) techniques [2].

Business failure prediction is a scientific field which many academic and professional people have been working for, at least, the three last decades. Also, financial organizations, such as banks, credit institutions, clients, etc., need these predictions for firms in which they have an interest (of any kind) [3]. Many financial decisions involve the classification of a set of observations (firms, stocks) into one or several groups of categories, what leads many researches to apply operational research methods to management problems. There exists an extensive literature devoted to the study of classification problems in the financial field, i.e. the work by Chen and Yeh [4] dealing with financial management or the works by Dimitras et al. [5] and Tam [6] in business failure [2]. Most papers deal with insurance audits, purchase intentions, purchase channel studies, methodologies for investigating customer purchasing intentions, and customer satisfaction [8]. Many research papers have quantified the problem in order to simplify the parameters, such as social parameters, and use statistical tools to analyze data. This approach, however, is only good for crisp types of data sets and certain data values. If the value of data is continuous or uncertain we must apply fuzzy theory [9]. In this paper, the Rough Set approach (RS) is used to provide a set of rules able to discriminate between healthy and failing firms in order to predict business loan payments [3]. The Rough Set Theory (RST) has been successfully applied in many real-life problems in medicine, pharmacology, engineering, banking, financial and market analysis and others. The RS methodology has found many real-life applications [10].

RST overlaps with many other theories, such as fuzzy sets, evidence theory, and statistics. Nevertheless, it can be viewed in its own right as an independent, complementary, and noncompeting discipline [11]. The most important problems that can be solved by RST are: finding description of sets of objects in terms of attribute values, checking dependencies (full or partial) between attributes, reducing attributes, analyzing the significance of attributes, and generating decision rules [12, 13]. The application of the RS approach in business failure prediction was investigated by Slowinski et al. [14] and Dimitras et al. [3]. In the past a large number of methods have been proposed to predict business failure; however, the special characteristics of the insurance sector have made most of them unfeasible. Up to date, just a few of them have been applied to this sector. Most approaches applied to insurance companies are statistical methods such as discriminant or logit analysis [15, 16] and use financial ratios as explicative variables.

The rest of the paper is structured as follows: Section 2 gives determinants of loan problems and Section 3 introduces RST briefly and explains methodology