Decision Support Systems 107 (2018) 78–87

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Automatic feature weighting for improving financial Decision Support Systems

Yosimar Oswaldo Serrano-Silva [a], Yenny Villuendas-Rey [b,*], Cornelio Yáñez-Márquez [a,*]

[a] Centro de Investigación en Computación del Instituto Politécnico Nacional, Av. Juan de Dios Bátiz Esq. Miguel Othón de Mendizábal S/N, Nueva Industrial Vallejo, 07738 Gustavo A. Madero, CDMX, Mexico

[b] Centro de Innovación y Desarrollo Tecnológico en Cómputo del Instituto Politécnico Nacional, Av. Juan de Dios Bátiz Esq. Miguel Othón de Mendizábal S/N, Nueva Industrial Vallejo, 07738 Gustavo A. Madero, CDMX, Mexico

ARTICLE INFO

Article history:
Received 11 September 2017
Received in revised form 23 January 2018
Accepted 24 January 2018
Available online 31 January 2018

Keywords:
Credit risk
Bankruptcy prediction
Banknote authentication
Bank telemarketing
Feature weight
Decision support

ABSTRACT

We propose a novel methodology for improving financial Decision Support Systems (DSS) through automatic feature weighting. Using this methodology, we show that automatic feature weighting leads to a significant improvement in the performance of decision-making algorithms over financial data, which are the key of financial DSS. The statistical analysis carried out shows that metaheuristic algorithms are good for automatic feature weighting, and that Differential Evolution (DE) offers a good trade-off between decision-making performance and computational cost. We believe these results contribute to the development of novel financial DSS.

1. Introduction

In the world of financial companies, *risk* means the danger of loss [1]. Financial companies assume a certain credit risk in each one of the operations they carry out (loans, lines of credit, and guarantees, among others). These financial companies cannot know everything about of their clients and, on the other hand, a client's compliance with its obligations to a financial company depends on events in which no one can know whether they will occur. That is, there is an uncertainty associated to whether a customer will (or will not) pay their debt [2,3].

According to the Global Association of Risk Professionals (GARP), which is a non-for-profit association for world-class financial risk certification, the presence of risks in financial companies is relevant because the mismanagement of financial risks can lead to global financial crises, such as those registered in the world in a cyclical manner, since the origins of financial systems. In this context, financial companies have developed numerous methods to value credits and determine credit risks, as well as their management; all of this, in order to improve financial results and profitability. Some of these methods are closely linked to models of computer sensitivity, which indicates the importance of computational methods in financial risk management [4].

With the tremendous advance of technology, and specifically with the arrival in the world of powerful computer systems, in the 1980s financial institutions turned to see the possibilities that this type of technology offers them in their field of action. Thus, some researchers decided to venture into this type of topics, so that computing in its different manifestations, became part of the financing processes, including desktop computers.

According to [5], several types of DSS are being applied in the world of financial risks, such as neural networks and pattern recognition techniques. These approaches, among others, have provided a theoretical foundation for developing DSS (including expert systems) to estimate, for instance, the probability of bankruptcy and predict fraud.

While it is true that DSS are not the only option to solve problems related to financial risks, their importance as a good option has been growing in recent years [6]. They have been applied in quality control, financial forecasting, targeted marketing, bankruptcy prediction, optical character recognition, among other problems [2–9].

In the present work, we address the following problems: credits risk, bankruptcy prediction, banknote authentication and bank telemarketing. To do this, we focused on a key aspect of DSS: the improvement of the decision-making (or prediction) algorithm. Those algorithms are the kernel of DSS, due to they help the system to process knowledge and to "decide" or to predict the most likely situation in the future. For instance, the prediction of risks in the financial environment can be seen as a decision-making problem. That is, by deciding that a company is in bankruptcy risk, we can reduce the uncertainty for the potential

* Corresponding authors.
E-mail addresses: yvilluendasr@ipn.mx (Y. Villuendas-Rey), cyanez@cic.ipn.mx (C. Yáñez-Márquez).

https://doi.org/10.1016/j.dss.2018.01.005
0167-9236/

investors, and the risk of capital loss due to an investment in a company that may be in bankruptcy in short term. In this work we are focused on the improvement of an algorithm recently proposed by us, which was designed for the financial area: the Naïve Associative Classifier (NAC) [3]. While it is true that there are several algorithms in the literature that can be used by DSS to improve decision-making [10–15], we have chosen the NAC because that it is one of the topics on which we are currently directing our research efforts.

The NAC has obtained very promising preliminary results in financial decision-making, and we hypothesize that if we develop a methodology for the automatic obtaining of weights for the attributes or features that describe financial data, we can improve the performance of the NAC in financial decision-making. Thus, we aim at improving the decision-making process, by improving the learning algorithm that guides this process.

Clearly, attribute selection is not the most important process to solve financial risk problems. Its importance lies in that, by improving the NAC, we can apply this model to different problems, ensuring better efficacy in the results; therefore, by making a good feature selection the NAC improves, and consequently, its application in DSS to reduce financial risks will be more effective. On the other hand, it is not in our interest to analyze the problems that people find when selecting attributes: we focus on selecting attributes to improve the NAC, and consequently to solve and predict financial risks in different areas. The results of this work support this affirmation.

Our main objective is to contribute to the development of DSS in the financial field, through the design of a methodology for the automatic adjustment of the attribute weights of financial data, by means of metaheuristic algorithms [7–9]. This methodology should improve the decision-making process in the financial area.

To test our hypothesis, we designed an experimental framework, taking into consideration the characteristics of financial data (such as data imbalance, presence of missing values, and presence of both numeric and categorical attributes), the functioning of metaheuristic algorithms, as well as the functioning of the NAC. We intended to compare the algorithms over the same conditions, and then to test the performance of the proposed automatic feature weighting methodology.

This paper offers the following contributions. First, we demonstrate how we can automatically obtain attribute weights in financial datasets, and how to improve the decision-making process in a Decision Support System of the financial field. Second, we examine the effect of metaheuristic algorithms for feature weighting, in the context of the NAC. The study's third contribution concerns the determination of the best metaheuristic for attribute weighting in the financial decision-making scenario. Our results show that DE is the most adequate algorithm for financial feature weighting.

The remainder of this paper is structured as follows. In Section 2, we review the related scientific literature on some of the problems that involve some type of risk within the financial environment. Then, in Section 3 we briefly describe the NAC, while in Section 4 the most relevant part of this article is profusely explained: the proposed methodology for financial feature weighting. After that, we present the results obtained, and discuss them in Section 5. The article ends with the conclusions in Section 6.

2. Related scientific literature

This section includes, and briefly describes, some of the research works that have addressed some of the four problems that we address in the present work: credits risk, bankruptcy prediction, banknote authentication and bank telemarketing. In this sense, the relationship of each of these works with our proposal is that the purpose is common: solve any of the four mentioned problems, using a computational method based on DSS. It is worth mentioning that, as far as we know, our proposal is the first to address financial problems by the improvement of a DSS through automatic feature weighting using metaheuristics. This means that in none of the works described in this section, feature selection, or feature weights, or metaheuristics are relevant. In the context of this article, the relevance of the works described here lies in the problems that intend to resolve, which are similar to those we approach with our proposal. In [10], the problem of credit risk and credit approval is faced. The authors propose a model to improve corporate credit rating systems that help financial experts to make decisions. Another different approach, based on social media is shown in [11], where it is analyzed whether the different opinions shared by users about different financial issues in these media are relevant or not for the prediction of credit risk to future of financial companies. To carry out this study, the authors analyzed the opinions expressed in two of the social networks for most popular financial investors in China, in addition to various articles published by financial analysts.

The second problem that we have selected, bankruptcy prediction, consists of to determine whether a company will be bankrupt in the near future [12]. Both financial institutions and investors need to reduce the possible risk of not obtaining the expected dividends or even of not recovering their capital if certain company declares bankruptcy. These types of situations can be addressed as a classification problem, in which a company can be classified as bankrupt or not in a short term, according to its current characteristics [6]. In this context, it is also important for financial firms to determine whether a person applying for a loan can file for bankruptcy in the short term, which would result in losses to such companies. It is well known that financial companies to grant some type of bank credit carry out a study of the client requesting such credit, i.e. they need to know if its economic solvency is good to be able to pay their obligations. For this type of problem, a model is proposed in [14] to help predict good clients (those who will meet their obligations), and bad clients (those who are likely to fail to fulfill their obligations for bankruptcy). With regard to banknote authentication problem, in [15] a system developed for discriminating fake notes from genuine ones and apply it to Indian banknotes is described. Image processing and pattern recognition techniques are used to design the overall approach.

The last of the problems we have selected is telemarketing. Today marketing campaigns can be considered as one of the most important business strategies, because different marketing strategies help companies to meet different goals and target specific sectors of customers [16]. However, the task of selecting the largest set of potential customers, for example, those clients most likely to hire a service or banking product, is considered NP-hard [17], so different solutions have been proposed to this problem in particular.

In one of these solutions, published in [18], a Decision Support System based on a Data Mining approach is proposed, which can automatically predict the outcome of a telephone call in which a banking service is offered. This DSS is intended to support the managers of financial companies in the selection and prioritization of customers to contact by telephone in the telemarketing banking campaigns and, consequently, reduce the cost and time of them. Another proposal is presented in [19], which is very similar to the previous proposal, since both seek to optimize the subscriptions to a banking service through telephone calls using a DSS, but in this case the aforementioned DSS uses Artificial Neural Networks.

To conclude this section, it is necessary to mention what concerns to algorithms used for decision-making, against which our proposal will compete. In the design and operation of the DSS are used several of these algorithms, among which the Artificial Neural Networks, the Support Vector Machines, the C4.5, the rule induction, the Instance-based learning algorithms and those based on the Bayes Theorem stand out [20–25].

The results obtained in [16] were taken into account in our experiments. In those results, the performance of the NAC was significantly better than the models RIPPER, MLP, SMO and 1NN, so we discarded those models, and only C4.5 and Naïve Bayes (NB) were considered in our experiments, due to they do not had significant difference in