

A recent review on optimisation methods applied to credit scoring models

Optimisation methods applied to CSMs

Elias Shohei Kamimura and Anderson Rogério Faia Pinto

Production Engineering Department, University of Araraquara, Araraquara, Brazil, and

Marcelo Seido Nagano

Production Engineering Department, São Carlos School of Engineering, University of São Paulo, São Carlos, Brazil

Received 28 September 2021

Revised 19 January 2022

31 October 2022

Accepted 5 February 2023

Abstract

Purpose – This paper aims to present a literature review of the most recent optimisation methods applied to Credit Scoring Models (CSMs).

Design/methodology/approach – The research methodology employed technical procedures based on bibliographic and exploratory analyses. A traditional investigation was carried out using the Scopus, ScienceDirect and Web of Science databases. The papers selection and classification took place in three steps considering only studies in English language and published in electronic journals (from 2008 to 2022). The investigation led up to the selection of 46 publications (10 presenting literature reviews and 36 proposing CSMs).

Findings – The findings showed that CSMs are usually formulated using Financial Analysis, Machine Learning, Statistical Techniques, Operational Research and Data Mining Algorithms. The main databases used by the researchers were banks and the University of California, Irvine. The analyses identified 48 methods used by CSMs, the main ones being: Logistic Regression (13%), Naïve Bayes (10%) and Artificial Neural Networks (7%). The authors conclude that advances in credit score studies will require new hybrid approaches capable of integrating Big Data and Deep Learning algorithms into CSMs. These algorithms should have practical issues considered consider practical issues for improving the level of adaptation and performance demanded for the CSMs.

Practical implications – The results of this study might provide considerable practical implications for the application of CSMs. As it was aimed to demonstrate the application of optimisation methods, it is highly considerable that legal and ethical issues should be better adapted to CSMs. It is also suggested improvement of studies focused on micro and small companies for sales in instalment plans and commercial credit through the improvement or new CSMs.

Originality/value – The economic reality surrounding credit granting has made risk management a complex decision-making issue increasingly supported by CSMs. Therefore, this paper satisfies an important gap in the literature to present an analysis of recent advances in optimisation methods applied to CSMs. The main contribution of this paper consists of presenting the evolution of the state of the art and future trends in studies aimed at proposing better CSMs.

Keywords Credit scoring, Literature review, Risk management, Optimization methods

Paper type Literature review

JEL Classification — C01, C44, C53, E51, G21

© Elias Shohei Kamimura, Anderson Rogério Faia Pinto and Marcelo Seido Nagano. Published in *Journal of Economics, Finance and Administrative Science*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence maybe seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

The research of the authors is partially supported by grant numbers 306075/2017-2, 430137/2018-4 312585/2021-7 from the Conselho Nacional de Desenvolvimento Científico and Tecnológico (CNPq), as well as by grant 2700441 from the Fundação Nacional de Desenvolvimento do Ensino Superior Particular (FUNADESP), Brazil.

Conflict of interest: The authors declare no competing interest.



Journal of Economics, Finance and
Administrative Science
Emerald Publishing Limited
2077-1886

DOI 10.1108/JEFAS-09-2021-0193

1. Introduction

Credit granting is an important element of financial transactions to provide liquidity for several economic activities (Doumpos *et al.*, 2018; Xia *et al.*, 2021). The problem with granting credit is a decision made under uncertain conditions in the face of the risk of borrowers not meeting their obligations. Furthermore, credit granting is usually regarded as a dynamic scenario (Xia *et al.*, 2021; Laborda and Ryoo, 2021). This makes it a complex decision-making issue, which may compromise the survival of an organisation. Thus, organisations are fundamentally responsible for assessing the risk of prospective borrowers before granting credit (Roy and Shaw, 2022). This risk consists of the possibility that the creditor incurs losses due to the non-fulfilment of obligations brought about by the borrower (Doumpos *et al.*, 2018; Li *et al.*, 2021). If the creditor can estimate the probability of a loss, then decision-making will be more reliable (Marqués *et al.*, 2013; Roy and Shaw, 2021a). These issues have become relevant topics in risk management for minimising financial losses for those who grant credit. Efficient credit risk management is a decisive factor for credit institutions, non-financial businesses and consumers (Andriopoulos *et al.*, 2019; Sariannidis *et al.*, 2020; Roy and Shaw, 2022). Most companies offer credit to customers (Doumpos *et al.*, 2018; Ashofteh and Bravo, 2021). Examples include banks, retailers, insurance companies, and micro and small businesses (Chen *et al.*, 2016; Li and Chen, 2020). The main components of credit risk modelling are i) *Probability of Default* (PD), ii) *Exposure at Default* and iii) *Loss Given Default*. For a theoretical background on these topics, refer to Andriopoulos *et al.* (2019), Breeden (2021), Salcedo (2021a), Salcedo (2021b) and Kozodoi *et al.* (2022).

The management and classification of the risk of a borrower or credit operation are made by employing Credit Scoring Models (CSMs). The CSMs aim to estimate the *default* risk by classifying the credit borrowers based on sociodemographic characteristics that allow them to be categorised as '*good*' or '*bad*' payers (refer to Louzada *et al.* (2016), Li and Chen (2020), Gunnarsson *et al.* (2021), Xia *et al.*, 2021; Kozodoi *et al.* (2022)). To establish superior CSMs, industry and academia mainly utilise the following two tools: *algorithms* and *data sources* (Trivedi, 2020; Breeden, 2021; Xia *et al.*, 2021). Over the last few decades, researchers have focused on developing improved CSMs. Emphasis has been placed on prediction methods including artificial intelligence algorithms and performance measures incorporated into CSMs (Lessmann *et al.*, 2015; Chen *et al.*, 2016). Most of these studies have used extensive databases with abundant variables to test the performance of CSMs. Furthermore, CSMs access these available data sources to extract, analyse and convert borrowers' information into risk measurement values (Rezáč, 2015; Kozodoi *et al.*, 2022). The indicator that quantifies the probability that a borrower to sustain a '*good*' risk is the Credit Score (CS). Letters, numbers or specific labels representing borrowers' idiosyncratic rate or quality may symbolise a CS (Li *et al.*, 2021). Thus, customers whose CS responds with a high probability of being a '*good*' payer would be accepted, and the others rejected by CSMs.

Recent studies have focused on improving the accuracy of CSMs for predicting payment default risk (refer to Lessmann *et al.*, 2015; Louzada *et al.*, 2016; Andriopoulos *et al.*, 2019; Kozodoi *et al.*, 2022). Important advances have been obtained, and practically all credit management areas (receipt, response, recovery, collection or risk measurement) use CSMs (Rezáč, 2015; Dastile and Celik, 2021). However, the literature has not yet presented a broader analysis of the CSM modelling process. Most approaches aim exclusively for the accuracy and results of optimisation methods and do not cover the entire context of building CSMs. Many studies have disregarded the real-world specificities (problem characteristics and customer databases) in applying CSMs. This study aims to present a literature review of the most recent optimisation methods applied to CSMs. This literature review included only papers published between January 2008 and May 2022. The delimitation of this study is based on the exponential growth of publications, as presented by Louzada *et al.* (2016). The scope of this study lies in providing theoretical lines which synthesise the state of the art and

cooperate with more promising studies for the better development of CSMs. This paper is structured as follows: **Section 2** provides a brief theoretical basis for credit operations, credit scoring and the main quantitative models used in CSMs; **Section 3** presents the literature review; **Section 4** demonstrates the research methodology; **Section 5** provides the results obtained; **Section 6** discusses the findings, and the paper ends with conclusions and future directions regarding CSMs.

2. Theoretical background

2.1 Credit operations

Credit operations express the delivery of goods or present value with the expectation of receiving a certain amount in the future (Marqués *et al.*, 2013; Trivedi, 2020). Such operations generate interest (Bravo *et al.*, 2013; Ashofteh and Bravo, 2021). These interests are charged for a predetermined period to minimise payment defaults (Doumpas *et al.*, 2018; Li and Chen, 2020; Trivedi, 2020). Given the criticality of risk assessment, a CSM's purpose is to classify potential borrowers as 'good' or 'bad' (Marqués *et al.*, 2013; Li and Chen, 2020). That is, those whose payment is expected on time and those whose payment is expected not to be complied with (Li and Chen, 2020). Traditionally, the models used for credit granting have been based on expert judgement. Expert judgement is primarily used when borrowers' historical data are missing or for special types of credit assessments (Andriopoulos *et al.*, 2019; Gunnarsson *et al.*, 2021). A common practice considers the 3, 4 or 5 C's qualitative model: character, capacity, capital, collateral and conditions (Marqués *et al.*, 2013). Nevertheless, as customer bases grew exponentially, financial institutions began to combine or replace credit granting decisions based on judgements with statistical models (Chen *et al.*, 2016; Gunnarsson *et al.*, 2021). In this respect, the Basel II Accord, which established a minimum capital requirement for financial institutions, was a watershed in the CS. These institutions resort to approaches based on internal classification, culminating in constant attempts to build CSMs (Chen *et al.*, 2016; Gunnarsson *et al.*, 2021). These CSMs make CS a primary tool for financial institutions to assess credit risk and make decisions on cash management and resource allocation (Marqués *et al.*, 2013; Gunnarsson *et al.*, 2021).

2.2 Credit scoring

In general, the CS is used to assess the risk of payment default when granting credit. CS includes an estimation based on the probability model of a borrower showing behaviour considered undesirable for the future (Lessmann *et al.*, 2015; Gunnarsson *et al.*, 2021). Thereby CSMs deal with a generic market-originated denomination and aim to quantify risk using formulas to calculate the referred CS (Marqués *et al.*, 2013; Louzada *et al.*, 2016; Andriopoulos *et al.*, 2019). Most CSMs aspire to identify the characteristics that influence behaviours that lead to either payment or default in a way that a customer might be classified as a 'good' or 'bad' payer (Louzada *et al.*, 2016). Thus, those customers whose CSMs present a high probability of being 'good' payers are accepted, and those with low probability are declined (Finlay, 2009; Breeden, 2021; Carta *et al.*, 2021). Furthermore, the latest CSMs have been employed for issues such as profitability, the use of Big Data (BD), Deep Learning (DL), equity in analysis, and sustainability (refer to Bastani *et al.*, 2019; Kozodoi *et al.*, 2019; Ashofteh and Bravo, 2021; Dastile and Celik, 2021; Djeundje *et al.*, 2021; Kang *et al.*, 2021 and Kozodoi *et al.*, 2022).

2.3 Main quantitative models

In quantitative models, each data instance is described by various characteristics representing the level of risk of a loan or borrower (Laborda and Ryoo, 2021; Xia *et al.*, 2021). This score might

be associated with risk classification and the PD estimation (Marqués *et al.*, 2013; Andriopoulos *et al.*, 2019). The traditional statistical methods include Discriminant Analysis (DA), Logistic Regression (LR), Classification Tree (CT) and Multiple Discriminant Analysis. These methods are linear in form and have the advantage of being easily applied and interpreted (Andriopoulos *et al.*, 2019; Marqués *et al.*, 2013). To establish a CS, the Operational Research models, such as Linear Programming, Quadratic Programming and Multiple-Criteria Decision-Making are also used in CSMs (Marqués *et al.*, 2013; Roy and Shaw, 2021b). Evolutionary Computation, Artificial Intelligence, Data Mining and Machine-Learning techniques describe credit risk with greater precision (Breeden, 2021; Xia *et al.*, 2021; Kozodoi *et al.*, 2022). The most prominent are Fuzzy Logic, Markov Chain, Bayesian Networks, Genetic Algorithm (GA), Naive Bayes (NB), k-Nearest Neighbors (k-NN), Artificial Neural Networks (ANN), Support Vector Machine (SVM) and Case-based Reasoning (CBR). The disadvantages are that these techniques require a great computational endeavour and finance, and business analysts seldom know them (Marqués *et al.*, 2013; Breeden, 2021).

3. Literature review

The call for an analysis of credit granting came about as sales commerce under future payment compromises began (Louzada *et al.*, 2016). The statistical score which distinguishes 'good' and 'bad' applicants was possibly presented for the first time by Durand (1941). Durand (1941) approached the risk elements in customer payment instalment plans (Gunnarsson *et al.*, 2021). However, the first operational scoring model was only proposed after a reasonable amount of time by Altman (1968). This model is based on five indices selected from eight variables in corporate financial statements, with a linear combination of these indices that demonstrated a discriminant function Z . Next, Orgler (1970) developed a general model of CS for commercial loans that approached the issue of dependent and independent variables using Multivariate Regression (MR). Eisenbeis (1978) used DA techniques to analyse the methodological approaches and statistical problems associated with CSMs.

In the 1980s, Capon (1982) suggested a more critical view of the logical basis of systems and CSMs. This is because statistical issues may cause severe legal problems for creditors if they are not correctly implemented in CSMs. Subsequently, Leonard (1992) modelled the credit decision process using DA and LR. Leonard (1992) used loan requests from small businesses handled by a large Canadian bank. Nonetheless, since the 2000s, new types of approaches have emerged to better deal with CS. Baesens *et al.* (2003) used three rule-extraction techniques through an ANN (Neurorule, Trepan and Nefclass). These techniques were employed for credit risk assessment using three data sets, demonstrating a powerful management tool via ANN and decision tables. Sinha and Zhao (2008) compared the performance of seven classification methods: LR, ANN, k-NN, SVM, Data Mining, Decision Table and Decision Tree (DT). Antonakis and Sfakianakis (2009) scrutinised the efficiency of Bayes' Theorem as a method for building classification rules in the triage of credit applicants. In this study, the researchers used two sets of real data to compare the rule with NB, LR, ANN, k-NN, CT and Linear Discriminant (LD). Finlay (2009) used a GA to generate a set of linear-scoring models oriented towards individual measures of organisational interest. Šuštersič *et al.* (2009) developed a CSM for consumers with limited data by implementing an ANN; for variable selection, GA and Principal Component Analysis (PCA) were used. Ince and Aktan (2009) researched the performance of CSMs which applied traditional approaches and artificial intelligence, such as DA, LR, ANN and Classification and Regression Trees (CART).

In 2010, research on CS grew exponentially. Finlay (2010) models continuous financial measures such as default, revenue and contribution to profit. Liu and Bo (2011) used a Simulated Annealing algorithm together with a GA to select the ideal attributes of an NB

classifier in real databases. [Vukovic et al. \(2012\)](#) presented a system of four CBR models which use GA to select the functions of preference and define the value of the attributes. [Bravo et al. \(2013\)](#) presented a methodology for granting and monitoring credit to micro-entrepreneurs by applying LR and Knowledge Discovery in Databases (KDD). [Kruppa et al. \(2013\)](#) improved probability estimation using methods such as k-NN and Random Forests (RF) deployed along with LR in a data set from a company that produces appliances. [Řezáč \(2014\)](#) proposes a new ESIS2 algorithm that estimates the information value and assesses the discriminatory power of the CSMs. [Verbraken et al. \(2014\)](#) adapted the Expected Maximum Profit (EMP) measure to find the compensation between expected losses and losses by default. [Kozeny \(2015\)](#) partially fills a gap in the usage of GA in CS, as these algorithms play a supporting role in other techniques, such as NN. [Lessmann et al. \(2015\)](#) updated [Baesens et al. \(2003\)](#) by comparing 41 classifiers in real-world databases. This study examined the extent to which alternative scoring card assessments differed between predictive indicators. Furthermore, [Lessmann et al. \(2015\)](#) compared other ensemble, hybrid system and single-model approaches. For a theoretical foundation regarding these modelling types, refer to [Louzada et al. \(2016\)](#) and [Andriopoulos et al. \(2019\)](#).

By the second half of the 2010s, studies were not limited to predicting payment default probability. [Serrano-Cinca and Gutiérrez-Nieto \(2016\)](#) proposed a system of support for a profit-scoring decision oriented to a Person-to-Person (P2P) loan based on MR and using the Internal Rate of Return (IRR). [Maldonado et al. \(2017\)](#) developed a structure based on profit to select models and attributes using a linear SVM. They also present a detailed cost-benefit analysis, including the calculation of financial losses for non-compliant payers. [Krichene \(2017\)](#) deployed an NB classifier to predict payment defaults on short-term loans in a commercial bank in Tunisia. [Bastani et al. \(2019\)](#) proposed a two-step approach that focuses on the lending market fund allocation process for P2P lending. This study integrated credit and profit scores based on Learning Algorithms (LA). [Sariannidis et al. \(2020\)](#) compared the prediction accuracy of seven methods: LR, NB, DT, k-NN, RF, Support Vector Clustering (SVC) and Linear Support Vector Clustering (LSVC). The precision of the resulting method ranged from 70% to 83%. [Kozodoi et al. \(2019\)](#) used the EMP measure and number of attributes as two adequate functions for selecting characteristics based on coverage to tackle both profitability and interpretability. [Çiğşar and Ünal \(2019\)](#) identified and used Data Mining classification algorithms to prevent default risk. They used NB, the J48 algorithm, a multivariate perceptron, six classification algorithms, and regression using WEKA 3.9 Data Mining (<https://waikato.github.io/weka-wiki>).

Moreover, researchers have combined more than one technique. [Trivedi \(2020\)](#) presented a prediction model and CSM using the NB, RF, DT and SVM classifiers. [Nalić and Martinovic \(2020\)](#) proposed a high-performance custom CSM based on credit history with real data and deployed the Generalised Linear Classification algorithm and SVM. [Li and Chen \(2020\)](#) conducted experiments and discussions in which a credit risk prediction model was used in a comparative assessment of four sets of algorithms: RF, AdaBoost, XGBoost and LightGBM. They combined piling with four traditional algorithms: ANN, LR, DT and SVM.

Recently, CSMs have addressed BD use, DL, and issues such as equity, profitability, sustainability, fraud prevention and economic variables. [Ashofteh and Bravo \(2021\)](#) presented a two-step method based on an initial Kruskal-Wallis non-parametric statistical analysis to formulate a conservative CSM. This CSM is based on the Machine Learning (ML) method for the default prediction of high-risk branches or customers. Thus, the RF, ANN, SVM and LR with Ridge penalty were used for the learning and evaluation of the referred CSM. [Carta et al. \(2021\)](#) proposed an ensemble stochastic criterion that operates in a discretised feature space and is extended to some meta-features to build an efficient CSM. This approach uses a real-world data set with different data imbalance configurations to apply the following classification algorithms: RF, DT, Adaptive Boosting, Multilayer

Perceptron and Gradient Boosting (GB). The stochastic criteria applied to a new feature space obtained by a twofold preprocessing technique perform the final classification of the CSM. [Dastile and Celik \(2021\)](#) provided a CSM using DL that converted tabular data sets into images to allow the application of 2D CNNs. Each pixel in an image corresponds to a feature bin in the tabular data set. The predictions of the 2D CNNs were explained using state-of-the-art CSM methods. [Djeundje et al. \(2021\)](#) evaluated the predictive performance of using psychometric variables and/or the characteristics of email use to predict consumer default probabilities. Researchers have applied a wide range of classification methods including LR, DL, PCA, XGBoost, Ridge Regression (RR) and Least Absolute Shrinkage and Selection Operator (LASSO). Instead, they are used to predict the credit risk of a new account and evaluate the predictive accuracy of CSMs. [Kang et al. \(2021\)](#) proposed a CSM to address the Rejection Inference (RI) issue. It considers an imbalanced data distribution for the consumer CS. Different classifiers were studied to propose the CSM; RF, DT, XGBoost, LightGBM and Modified Synthetic Minority Oversampling Technique (Borderline-SMOTE). Thus, the researchers' conduction of imbalanced learning using a Borderline-SMOTE and a graph-based semi-supervised LA called Label Spreading is applied to solve the RI.

[Kozodoi et al. \(2022\)](#) examined ML applications in the retail credit market. The researchers revisit(ed) statistical fairness criteria and examined their adequacy for CS. They then catalogued algorithmic options to incorporate fairness goals into the development of ML-based CSMs. Ergo empirically compared different fairness processors in a profit-oriented CS context using real-world data through the EMP. The fairness pre- and post-processors, as well as an unconstrained scorecard, use four base classifiers: LR, RF, ANN and XGBoost. The corresponding code is available on GitHub (<https://github.com/>). [Laborda and Ryoo \(2021\)](#) presented a methodology for selecting key variables to establish a CSM. In this study, LR, RA, SVM and k-NN were proposed to separate the data into two classes and identify the candidates that are likely to default on this CS. [Li et al. \(2021\)](#) presented a CSM that captures defaulting borrowers on an online lending platform using Multi-Layer Structured Gradient Boosted Decision Trees with Light Gradient Boosting Machines (ML-LightGBM). [Roa et al. \(2021\)](#) presented the impact of alternative data originating from an app-based marketplace (in contrast to traditional bureau data) on CSMs. Researchers have applied EMP measures and Stochastic Gradient Boosting (SGB). Furthermore, the Tree-based SHapley Additive explanation method was used for the SGB interpretation. [Roy and Shaw \(2021a\)](#) proposed a low-cost CSM for financial institutions that focused on Small and Medium Enterprises (SMEs) CS. The researchers integrated the Analytic Hierarchy Process (AHP) and the Technique for Order Preferences by Similarity to an Ideal Solution (TOPSIS) for AHP-TOPSIS. [Roy and Shaw \(2021b\)](#) developed a system to predict SMEs' credit risk by introducing a multi-criteria model formulated using a hybrid method that combines TOPSIS and Best-Worst Method (BWM). [Xia et al. \(2021\)](#) devised a CSM in which the data frequency and delays from Multilevel Macroeconomic Variables (MVs) are associated with app data for CS. Moreover, [Xia et al. \(2021\)](#) proposed a Bayesian selection and lag optimisation method to handle highly correlated MVs and capture flexible lag effects. [Roy and Shaw \(2022\)](#) filled a gap in the literature by proposing a multi-criteria Sustainability Credit Score System. This approach considers environmental and social aspects besides financial and managerial issues by combining BWM and TOPSIS.

4. Method

This research focuses on analysing the database characteristics and optimisation methods applied to CSMs. Therefore, this paper presents a literature review of technical procedures based on bibliographic exploratory research (refer to [Louzada et al., 2016](#); [Watson and Webster, 2020](#); [Lim et al., 2022](#)). The selection and classification of scientific publications

included the following steps: i) database search, ii) selection of published papers and iii) classification of the selected papers. [Figure 1](#) illustrates the steps of the research methodology.

First, a search for publications was performed using the Scopus, ScienceDirect and Web of Science databases ([Paul and Criado, 2020](#); [Donthu et al., 2021](#)). The keywords clusters used in advanced search are as follows: 'credit', 'review', 'scoring', 'modelling' and 'profitability'. These keywords were combined with the Boolean operator 'AND'. Publications were published between January 1968 and May 2022. The search only considered papers published in online journals in English. After the exclusion of duplicate studies, 647 publications were included. Therefore, a preliminary analysis resulted in the segmentation of 321 publications using several approaches to CS and CSM.

In the second step, publications were selected through a careful evaluation of the purpose of the study regarding CSMs. Notably, for this literature review, we selected papers published between January 2008 and May 2022. The paper selection period was based on the exponential growth of publications, as presented by [Louzada et al. \(2016\)](#). This step resulted in a final selection of 46 papers containing literature review approaches and solution methods proposed for CSMs. Papers published between January 1968 and December 2007 were used as theoretical frameworks for CS and CSMs. The remaining papers were discarded because they did not fit the established protocols for approaches inherent to CSMs. Books and abstracts addressing CSMs were excluded. The keywords used by the 46 selected papers are illustrated in the cloud map shown in [Figure 2](#) and generated by VOSviewer version 1.6.16 (<http://www.vosviewer.com/>).

[Figure 2](#) shows that the total number of keywords listed by all selected papers was 224, as generated by VOSviewer. Furthermore, the map shows the main keywords related to the theme group, represented by 'Credit Scoring'. The keywords most used by the selected papers from 2008 to 2022 are Credit Scoring (31), Data Mining (6), Classification (6), Machine Learning (6), Genetic Algorithm (4) and Genetic Algorithms (4). The relations between the intensity and occurrence of keywords indicate that the selected papers are pertinent to a literature review of CSMs.

In the third step, the selected papers were classified into two groups: Solution Methods and Literature Reviews. Thus, the number of papers that proposed solution methods (including

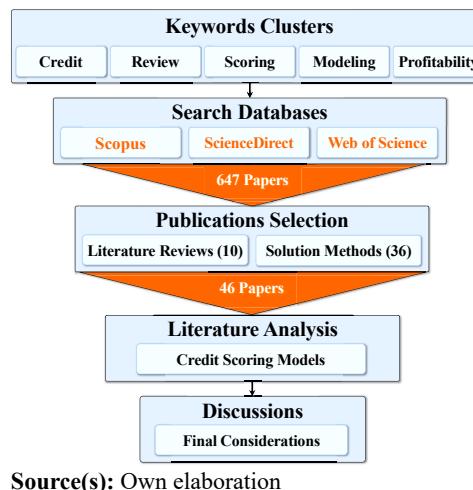


Figure 1.
Diagram showing the
research
methodology steps

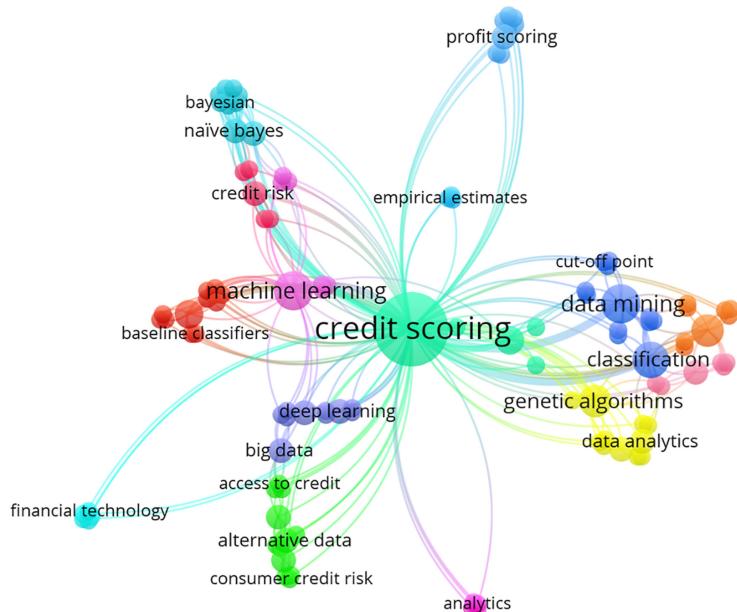


Figure 2.
Keyword relations map

Source(s): Own elaboration

modelling, profitability and database selection) used by CSMs was 36 (78%). The literature reviews ten (22%) papers that provide a theoretical framework for CSMs. Furthermore, the literature review incorporates innovations and analyses related to recent publications proposing new solution methods for CSMs. Thus, these papers were selected based on their relevance in transferring historical information to update and improve state-of-the-art CSMs. The literature review observed two types of approaches: narrative and systematic reviews. The classifications and research methodologies used in the literature review are presented in Table 1.

Table 1 demonstrates that the number of papers referring to narrative reviews (5) is identical to that referring to systematic reviews (5). The first exposes the state of the art with a theoretical or contextual focus, and the second answers questions using specific methods to locate, select and technically evaluate studies (refer to [Paul and Criado, 2020](#); [Donthu et al., 2021](#); [Lim et al., 2022](#)). Table 2 presents the bibliometric indicators of the Scopus and Web of Science databases referring to journals that published papers classified as Solution Methods (36) and Literature Reviews (10). The graph in Figure 3 illustrates journals with two or more publications, while the rest are classified as 'Other Journals with Just One Publication'.

Table 2 lists the selected papers published in 28 journals. Most of these journals were published in Europe (23 journals, 41 papers). The other journals were from America (3 journals, 3 papers), Africa (1 journal, 1 paper) and Asia (1 journal, 1 paper). In terms of the number of journals, Europe (82%) was superior to America (11%), Africa (4%) and Asia (4%). Almost all the selected papers came from Europe (89%), and the rest from America (7%), Africa (2%) and Asia (2%). The countries where most journals were based were the United Kingdom (10 journals, 19 papers), the Netherlands (9 journals, 17 papers), the United States (3 journals, 3 papers) and Germany (2 journals, 2 papers). Ergo, 85% of the selected papers

| References Authors | Literature reviews Credit scoring models | Review type NR | SR | Optimisation methods applied to CSMs |
|--|--|-------------------|----|---|
| Abdou and Pointon (2011) | 214 books/theses/papers (Application in different areas) | | ✓ | |
| Marqués et al. (2013) | Journals and conference papers: 2000–2012 | | ✓ | |
| Lessmann et al. (2015) | 41 classifiers in 8 real-world data sets | ✓ | | |
| Chen et al. (2016) | Not specified | ✓ | | |
| Louzada et al. (2016) | 437 papers (Reaxys, Scopus, Science Direct and Engineering Information: 1992–2015) | | ✓ | — |
| Onay and Ozturk (2018) | 299 papers (ProQuest and Emerald Research Bases: 1976–2017) | | ✓ | |
| Andriopoulos et al. (2019) | Not specified | ✓ | | |
| Goh and Lee (2019) | 75 papers (Science Direct, Google Scholar and IEEE Xplore: 1997–2018) | | ✓ | |
| Breedon (2021) | Not specified | ✓ | | |
| Gunnarsson et al. (2021) | Not specified | ✓ | | |

Note(s): Referenced abbreviations: NR – Narrative Review; SR – Systematic Review
Source(s): Own elaboration

Table 1.
Approaches proposing literature review

were concentrated in journals in the United Kingdom (41%), the Netherlands (37%) and the United States (7%). [Figure 3](#) shows that the journals with more than one selected paper were Expert Systems with Applications (20%), European Journal of Operational Research (13%), Decision Support Systems (9%), Journal of the Operational Research Society (4%) and Mathematics (4%), amounting to 23 (50%). The remaining journals were grouped as ‘Other Journals with Just One Publication’.

5. Results

The literature review shows that during the last decade, there has been a constancy in research proposing CSMs. A significant increase was observed in 2019. Their findings show that CSMs are usually formulated using financial analysis, ML, statistical techniques, operational research and data-mining algorithms. The analysis identified 48 methods used by researchers for construction, performance tests and comparisons between CSMs. These studies and the solution methods used by CSMs are presented in [Table 3](#). Next, the graph in [Figure 4](#) illustrates the methods with two or more applications while the rest are classified as ‘Other Methods’.

[Figure 4](#) shows the amounts and percentages of the solution methods applied to the CSMs. The most used methods were LR (13%), NB (10%) and ANN (7%). Furthermore, according to [Louzada et al. \(2016\)](#) and [Andriopoulos et al. \(2019\)](#), three methodological schemes can be identified for constructing CSMs: ensemble, hybrid system and single-model approaches. The distribution of the methodological schemes applied to each study is shown in [Table 4](#). [Figure 5](#) displays the modelling types used in the studies according to the classification presented by [Louzada et al. \(2016\)](#).

[Figure 5](#) demonstrates that the most commonly used modelling types are Hybrid Systems (72%), followed by single-model approaches (25%) and ensembles (3%). Single-model approaches propose CSMs using only one method ([Andriopoulos et al., 2019](#)). Hybrid Systems combine diverse techniques and modelling schemes in different ways to improve CSMs’ performance ([Louzada et al., 2016](#); [Andriopoulos et al., 2019](#)). Although many

JEFAS

| Journals Denomination | Journals countries | Journal H-index | Journal CiteScore | Impact factor | Highest percentile | Highest quartile | References (Authors) |
|---|--------------------|-----------------|-------------------|---------------|--------------------|------------------|--|
| Expert Systems with Applications | United Kingdom | 225 | 12.7 | 6.954 | 98% | Q1 | Finlay (2009), Šusteršić <i>et al.</i> (2009), Vukovic <i>et al.</i> (2012), Kruppa <i>et al.</i> (2013), Kozeny (2015), Bastani <i>et al.</i> (2019), Ashofteh and Bravo (2021), Djeundje <i>et al.</i> (2021) and Roa <i>et al.</i> (2021) |
| European Journal of Operational Research | Netherlands | 274 | 9.5 | 5.334 | 97% | Q1 | Finlay (2010), Bravo <i>et al.</i> (2013), Verbraken <i>et al.</i> (2014), Lessmann <i>et al.</i> (2015), Gunnarsson <i>et al.</i> (2021) and Kozodoi <i>et al.</i> (2022) |
| Decision Support Systems | Netherlands | 161 | 10.5 | 5.795 | 98% | Q1 | Sinha and Zhao (2008), Serrano-Cinca and Gutiérrez-Nieto (2016), Maldonado <i>et al.</i> (2017) and Kozodoi <i>et al.</i> (2019) |
| Journal of the Operational Research Society | United Kingdom | 115 | 4.1 | 2.860 | 87% | Q1 | Marqués <i>et al.</i> (2013) and Andriopoulos <i>et al.</i> (2019) |
| Mathematics | Switzerland | 43 | 2.2 | 2.258 | 80% | Q1 | Li and Chen (2020) and Laborda and Ryoo (2021) |
| IEEE Access | United States | 158 | 6.7 | 3.476 | 90% | Q1 | Dastile and Celik (2021) |
| Financial Innovation | Germany | 25 | 6.7 | 6.793 | 92% | Q2 | Roy e Shaw (2021b) |
| Procedia Engineering | Netherlands | 88 | 4.0 | 1.880 | 80% | Q1 | Liu and Bo (2011) |
| Technology in Society | United Kingdom | 58 | 4.2 | 4.192 | 90% | Q1 | Trivedi (2020) |
| Journal of Credit Risks | United States | 11 | 1.3 | 0.226 | 36% | Q3 | Breeden (2021) |
| Scientific Programming | Egypt | 36 | 2.0 | 1.025 | 41% | Q3 | Çiğşar and Ünal (2019) |

Table 2.
Bibliometric indicators
of journals

(continued)

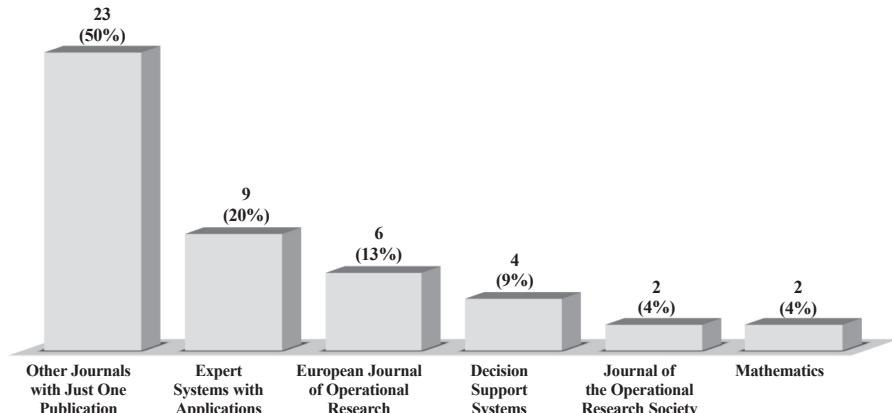
| Journals Denomination | Journals countries | Journal H-index | Journal CiteScore | Impact factor | Highest percentile | Highest quartile | References (Authors) | Optimisation methods applied to CSMs |
|--|--------------------|-----------------|-------------------|---------------|--------------------|------------------|--|--------------------------------------|
| Applied Soft Computing | Netherlands | 156 | 12.4 | 8.263 | 92% | Q1 | Kang <i>et al.</i> (2021) | |
| Computational Economics | Netherlands | 43 | 2.3 | 1.876 | 72% | Q2 | Řez áč (2015) | |
| Knowledge-Based Systems | Netherlands | 135 | 12.0 | 8.139 | 92% | Q1 | Li <i>et al.</i> (2021) | |
| Journal of Applied Statistics | United Kingdom | 63 | 1.9 | 1.404 | 62% | Q2 | Antonakis and Sfakianakis (2009) | |
| Artificial Intelligence Review | Netherlands | 86 | 10.4 | 8.139 | 99% | Q1 | Chen <i>et al.</i> (2016) | |
| Annals of Operations Research | Netherlands | 111 | 5.2 | 4.854 | 83% | Q1 | Sariannidis <i>et al.</i> (2020) | |
| Progress in Artificial Intelligence | Germany | 22 | 5.4 | 2.254 | 67% | Q2 | Carta <i>et al.</i> (2021) | |
| Advances in Operations Research | United States | 17 | 2.9 | 3.579 | 55% | Q2 | Goh and Lee (2019) | |
| International Journal of Finance and Economics | United Kingdom | 41 | 2.1 | 0.420 | 55% | Q2 | Roy and Shaw (2021a) | |
| Journal of Financial Regulation and Compliance | United Kingdom | 20 | 1.6 | 0.761 | 40% | Q3 | Onay and Ozturk (2018) | |
| Electronic Commerce Research and Applications | Netherlands | 82 | 10.0 | 5.622 | 92% | Q1 | Xia <i>et al.</i> (2021) | |
| Journal of Business Economics and Management | Lithuania | 41 | 3.5 | 2.445 | 78% | Q1 | Ince and Aktan (2009) | |
| Journal of Economics Finance and Administrative Science | United Kingdom | 16 | 1.5 | 1.270 | 62% | Q2 | Krichene (2017) | |
| Surveys In Operations Research and Management Science | United Kingdom | 26 | 7.0 | 4.008 | 93% | Q1 | Louzada <i>et al.</i> (2016) | |
| Intelligent Systems in Accounting Finance and Management | United Kingdom | 14 | 4.1 | 5.500 | 89% | Q1 | Abdou and Pointon (2011) | |

(continued)

Table 2.

| | Journals Denomination | Journals countries | Journal H-index | Journal CiteScore | Impact factor | Highest percentile | Highest quartile | References (Authors) |
|--|---|--------------------|-----------------|-------------------|---------------|--------------------|------------------|-----------------------------|
| | International Journal of Sustainable Development and World Ecology | Singapore | 36 | 2.3 | 1.470 | 48% | Q3 | Roy and Shaw (2022) |
| | International Journal of Software Engineering and Knowledge Engineering | United Kingdom | 48 | 7.1 | 3.716 | 93% | Q1 | Nalić and Martinovic (2020) |

Table 2.

Figure 3.
Publications per journal

Source(s): Own elaboration

techniques have been explored for Hybrid Systems, only one is typically implemented in the final prediction (Chen *et al.*, 2016). Lin *et al.* (2012) presented three approaches to construct a Hybrid System: cascade, integration and clustering combination modes. A summary of these approaches is provided in Table 5.

Ensembles combine different models developed using one or more algorithms to obtain better classifiers (Louzada *et al.*, 2016; Andriopoulos *et al.*, 2019). A literature analysis shows that the most used ensemble models are Piling, Bagging and Impulsing. These models' performance depends on the diversity of the methods used to reduce their bias (Louzada *et al.*, 2016; Andriopoulos *et al.*, 2019; Breeden, 2021). Analyses of the studies also confirmed that researchers used large and diverse databases with many variables to apply CSMs. These databases can be synthesised into three categories: i) Banks (22%), ii) Other Databases (53%) and iii) UCI Repositories and Others (25%). The frequencies of the databases used in these studies are shown in Figure 6.

Figure 6 indicates that most researchers have used Other Databases to develop CSMs. These are the main Other Databases: i) Lending Club in the United States; ii) US and China

| References Authors | Solution methods Credit scoring models | Optimisation methods applied to CSMs |
|---|--|---|
| Sinha and Zhao (2008) | LR, DT, NB, k-NN, ANN, SVM and Decision Table | |
| Antonakis and Sfakianakis (2009) | CT, NB, LD, LR, k-NN and ANN | |
| Finlay (2009) | GA | |
| Ince and Aktan (2009) | DA, LR, CART and ANN | |
| Šuštersić <i>et al.</i> (2009) | GA, LR, EBP and ANN | |
| Finlay (2010) | GA | |
| Liu and Bo (2011) | SA, GA and NB | |
| Vukovic <i>et al.</i> (2012) | GA, k-NN and CBR | |
| Bravo <i>et al.</i> (2013) | LR and KDD | |
| Kruppa <i>et al.</i> (2013) | RF, LR and k-NN | |
| Verbraken <i>et al.</i> (2014) | LR, EMP and ANN | |
| Kozeny (2015) | GA | |
| Řez áč (2015) | MCS and ESIS2 Algorithm | |
| Serrano-Cinca and Gutié rrez-Nieto (2016) | DT, MR and IRR | |
| Krichene (2017) | NB and ANN | |
| Maldonado <i>et al.</i> (2017) | SVM | |
| Bastani <i>et al.</i> (2019) | LR, IHT, IRR and SMOTE | |
| Çi ̄ş ar and Ünal (2019) | LR, RF, NB, MP, J48 Algorithm and Bayesian Networks | |
| Kozodoi <i>et al.</i> (2019) | EMP and NSGA-II Algorithm | |
| Sariannidis <i>et al.</i> (2020) | LR, NB, DT, RF, SVC, k-NN and LSVC | |
| Li and Chen (2020) | DT, LR, RF, NB, ANN, SVM, XGBoost, AdaBoost and LightGBM | |
| Nalić and Martinovic (2020) | GLC and SVM | |
| Trivedi (2020) | RF, DT, NB and SVM | |
| Ashofteh and Bravo (2021) | LR, RF, ANN and SVM | |
| Carta <i>et al.</i> (2021) | GB, AB, RF, DT and MP | |
| Dastile and Celik (2021) | CNNs | |
| Djeundje <i>et al.</i> (2021) | LR, RR, PCA, XGBoost and LASSO Regression | |
| Kang <i>et al.</i> (2021) | RF, DT, XGBoost, LightGBM and Borderline-SMOTE | |
| Laborda and Ryoo (2021) | LR, RA, SVM and k-NN | |
| Li <i>et al.</i> (2021) | DT, LR, RF, GB and ML-LightGBM | |
| Roa <i>et al.</i> (2021) | EMP and SGB | |
| Roy and Shaw (2021a) | AHP and TOPSIS | |
| Roy and Shaw (2021b) | BWM and TOPSIS | |
| Xia <i>et al.</i> (2021) | LR, RF, CatBoost and XGBoost | |
| Kozodoi <i>et al.</i> (2022) | LR, ANN, RF, XGBoost and EMP | |
| Roy and Shaw (2022) | BWM and TOPSIS | |

Note(s): Referenced abbreviations: NB – Naive Bayes; DT – Decision Trees; RF – Random Forests; VS – Variable Selection; RR – Ridge Regression; GB – Gradient Boosting; GA – Genetic Algorithm; AB – Adaptive Boosting; CT – Classification Trees; LR – Logistic Regression; LD – Linear Discriminant; SA – Simulated Annealing; DA – Discriminant Analysis; MP – Multilayer Perceptron; k-NN – k-Nearest; Neighbors; IRR – Internal Rate of Return; BWM – Best-Worst Method; CBR – Case-Based Reasoning; MP – Multilayered Perceptron; MR – Multivariate Regression; SVM – Support Vector Machine; MCS – Monte Carlo Simulations; EMP – Expected Maximum Profit; SVC – Support Vector Clustering; ANN – Artificial Neural Networks; SGB – Stochastic Gradient Boosting; AHP – Analytic Hierarchy Process; IHT – Instance Hardness Threshold; PCA – Principal Component Analysis; XGBoost – Extreme Gradient Boosting; GLC – Generalised Linear Classification; CNNs – Convolutional Neural Networks; CatBoost – Categorical Gradient Boosting; KDD – Knowledge Discovery in Databases; ML-LightGBM – Light Gradient Boosting Machines; CART – Classification and Regression Trees; SMOTE – Synthetic Minority Oversampling Technique; LASSO – Least Absolute Shrinkage and Selection Operator; TOPSIS – Technique for Order of Preference by Similarity to Ideal Solution; Borderline-SMOTE – Modified Synthetic Minority Oversampling Technique

Source(s): Own elaboration

Table 3.
Approaches proposing
solution methods

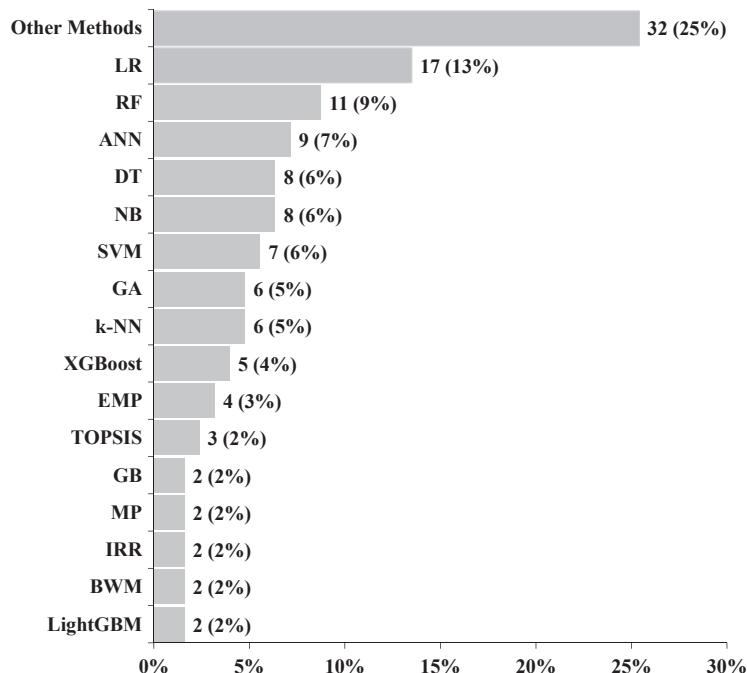


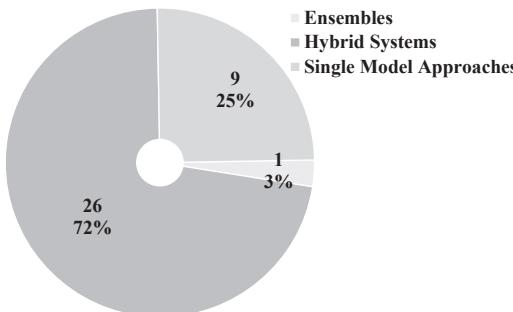
Figure 4.
Methods used in credit scoring models

Source(s): Own elaboration

| Modelling types | References Authors |
|-------------------------|---|
| Ensembles | Li and Chen (2020) |
| Hybrid Systems | Antonakis and Sfakianakis (2009), Ince and Aktan (2009), Šuštersič <i>et al.</i> (2009), Liu and Bo (2011), Vukovic <i>et al.</i> (2012), Bravo <i>et al.</i> (2013), Kruppa <i>et al.</i> (2013), Verbraken <i>et al.</i> (2014), Řezáč (2015), Serrano-Cinca and Gutiérrez-Nieto (2016), Bastani <i>et al.</i> (2019), Çiğşar and Ünal (2019), Kozodoi <i>et al.</i> (2019), Nalić and Martinovic (2020), Ashofteh and Bravo (2021), Carta <i>et al.</i> (2021), Djeundje <i>et al.</i> (2021), Kang <i>et al.</i> (2021), Laborda and Ryoo (2021), Li <i>et al.</i> (2021), Roa <i>et al.</i> (2021), Roy and Shaw (2021a), Roy and Shaw (2021b), Xia <i>et al.</i> (2021), Kozodoi <i>et al.</i> (2022) and Roy and Shaw (2022) |
| Single-Model Approaches | Sinha and Zhao (2008), Finlay (2009), Finlay (2010), Kozeny (2015), Krichene (2017), Maldonado <i>et al.</i> (2017), Sariannidis <i>et al.</i> (2020), and Trivedi (2020) and Dastile and Celik (2021) |

Table 4.
Modelling types used in credit scoring models

P2P Platform; iii) Credit Bureau Germany and Australia; iv) PAKDD; v) GMSC; vi) Homecredit and vii) Financial Institutions Platforms in Benelux and the UK. The literature also demonstrates that these studies used 22 databases to formulate CSMs. Thus, eight researchers used data from banks across several countries. Another six studies used databases available in the UCI Repository of the Machine Learning Database. In another three studies, the researchers dealt with UCI databases and other platforms, such as the Greek banks PAKDD and Kaggle, and financial institutions from Benelux and the UK.



Source(s): Own elaboration

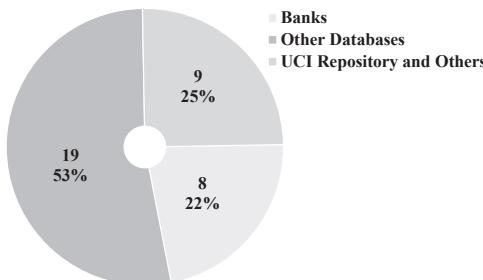
Optimisation methods applied to CSMs

Figure 5.
Modelling types used in credit scoring models

| Research approaches | Applied Techniques |
|-----------------------------|---|
| Cascade Mode | Different classifiers in cascade, where the exit of the first-level classifier feeds the second-level classifier as entry |
| Integration Mode | Heuristic techniques are integrated into classification models to optimise the prediction performance from various perspectives |
| Clustering Combination Mode | Clustered storage is used as a stage of pre-processing classification to enhance the prediction precision |

Source(s): Own elaboration

Table 5.
Approaches to construction of hybrid systems



Source(s): Own elaboration

Figure 6.
Databases used in credit scoring models

6. Discussion

A literature review demonstrates the use of different techniques and approaches for formulating CSMs. The analysed papers present CSMs formulated upon applying different techniques and methods to solve various problems present in various contexts and realities that configure the CS. We demonstrated that CS approaches are directly related to the context and characteristics of the problems, together with the choices of the most appropriate methods for CSMs. The most recent CSM studies are based on profit and loan profitability estimates instead of focusing only on payment default probability. This is because researchers concluded that the causes of profitability differ from the reasons for default. Customers with a high probability of payment non-compliance may also be profitable (Serrano-Cinca and Gutiérrez-Nieto, 2016; Onay and Ozturk, 2018). Thus, CSMs based on

distinguishing payment delinquents and constructing a loan profit and profitability score resort to approaches such as the IRR, Game Theory, Statistical Techniques and Artificial Intelligence. There is a growing trend towards complex ML algorithms (Xia *et al.*, 2021; Kozodoi *et al.*, 2022). For the theoretical framework, refer to Bravo *et al.* (2013), Řezáč (2014), Verbraken *et al.* (2014), Serrano-Cinca and Gutiérrez-Nieto (2016), Onay and Ozturk (2018) and Kozodoi *et al.* (2019). Recent studies have demonstrated that BD prompts disruptive changes in CSMs. The incorporation of a greater volume and variety of data linked to the need for higher speed in collecting and storing these data has become a challenge for CSMs (Ashofteh and Bravo, 2021; Kang *et al.*, 2021). This requires a broader approach, not only of the recorded history of borrowers' payments and receipts but also data from social networks, information from apps and the so-called digital footprints (Roa *et al.*, 2021). Therefore, BD enables credit quality assessment for potential borrowers with a limited financial history (Onay and Ozturk, 2018). Recent studies address the application of DL and alternative data using psychometric variables and/or email-use characteristics to predict consumer default probabilities (Dastile and Celik, 2021; Djeundje *et al.*, 2021; Roa *et al.*, 2021). Banks, fintech companies, credit bureaus and other non-banking providers of financial services use BD to achieve a higher level of precision in their services. However, this new reality has introduced regulatory challenges in preventing discrimination and consumer rights (Onay and Ozturk, 2018). Current studies include themes such as equity in customer classification (Kozodoi *et al.*, 2022), sustainability issues in CS (Roy and Shaw, 2022) and the incorporation of macroeconomic variables that can directly affect CSMs (Xia *et al.* 2021).

7. Conclusions

This study presents a literature review of the most recent optimisation methods applied to CSMs. The Scopus, ScienceDirect and Web of Science databases were used (from 2008 to 2022). This investigation led to the selection of 36 papers proposing CSMs. These CSMs are used to assess the risk of payment default when granting credit, namely Credit Scoring (CS). Their findings show that CSMs are usually formulated using financial analysis, ML, statistical techniques, operational research and data-mining algorithms. The analysis identified 48 methods used by researchers for construction, performance tests and comparisons between CSMs. The most commonly used methods were LR (13%), NB (10%) and ANN (7%). Most models were formulated using three methodological schemes called Hybrid Systems (72%), followed by single-model approaches (25%) and ensembles (3%). Analyses of the studies also confirmed that researchers used large and diverse databases with many variables to apply CSMs. These databases can be synthesised into three categories: i) Other Databases (53%), ii) UCI Repositories and Others (25%) and iii) Banks (22%). The databases are as follows: i) banks from various nations (8); ii) UCI Repository of Machine Learning Database (6); iii) UCI, Greek banks PAKDD and Kaggle, financial institutions from Benelux and the United Kingdom (2), and Lending Club in the United States (2). These journals were Decision Support Systems, Expert Systems with Applications, European Journal of Operational Research and Journal of the Operational Research Society. Other studies have also used different databases to apply CSMs.

This study also demonstrated that recent studies have focused on the loan yield and profit-scoring theme of CSMs. Therefore, estimating only the PD is no longer the primary objective of all CSMs. The researchers' shift in focus sheds light on a new perspective on maximising the financial results of loans in analyses that include CS. The main contribution of this study is to present the evolution of the state of the art and future trends in research aimed at proposing better CSMs. The results of this study can guide researchers and provide considerable practical implications for the application of CSMs. We also encourage researchers to consider legal and ethical issues and conduct studies aimed at micro- and small-sized companies for instalment

sales and commercial credit through improvements or new CSMs. We conclude that advances in CS studies require new hybrid approaches that can integrate BD and DL algorithms into CSMs. These algorithms must consider practical issues to improve the level of adaptation and performance required for CSMs. Suggestions for future research are i) the Use of BD and DL in CSMs; ii) equity issues in credit ratings; iii) formulating CSMs focused on sustainability; iv) providing decision support tools for credit sales; v) improving CSMs for default risk and investment, instalment and credit sales decisions and viii) implementing legal and ethical issues in CSMs based on the General Data Protection Regulation.

References

Abdou, A.J. and Pointon, H.A. (2011), "Intelligent systems in accounting, finance and management", *Intelligent Systems in Accounting, Finance and Management*, Vol. 16 Nos 1-2, pp. 21-31, available at: <https://onlinelibrary.wiley.com/doi/10.1002/isaf.325>

Altman, E.I. (1968), "American finance association", *The Journal of Finance*, Vol. 29 No. 1, pp. 312-312, available at: <https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1968.tb00843.x>

Andriopoulos, D., Doumpos, M., Pardalos, P.M. and Zopounidis, C. (2019), "Computational approaches and data analytics in financial services: a literature review", *Journal of the Operational Research Society*, Vol. 70 No. 10, pp. 1581-1599, Taylor & Francis, available at: <https://www.tandfonline.com/doi/full/10.1080/01605682.2019.1595193>

Antonakis, A.C. and Sfakianakis, M.E. (2009), "Assessing Naïve Bayes as a method for screening credit applicants", *Journal of Applied Statistics*, Vol. 36 No. 5, pp. 537-545, doi: [10.1080/02664760802554263](https://doi.org/10.1080/02664760802554263).

Ashofteh, A. and Bravo, J.M. (2021), "A conservative approach for online credit scoring", *Expert Systems with Applications*, Vol. 176, 114835, doi: [10.1016/j.eswa.2021.114835](https://doi.org/10.1016/j.eswa.2021.114835).

Baesens, B., Setiono, R., Mues, C. and Vanthienen, J. (2003), "Using neural network rule extraction and decision tables for credit-risk evaluation", *Management Science*, Vol. 49 No. 3, pp. 312-329, doi: [10.1287/mnsc.49.3.312.12739](https://doi.org/10.1287/mnsc.49.3.312.12739).

Bastani, K., Asgari, E. and Namavari, H. (2019), "Wide and deep learning for peer-to-peer lending", *Expert Systems with Applications*, Vol. 134, pp. 209-224, Elsevier, doi: [10.1016/j.eswa.2019.05.042](https://doi.org/10.1016/j.eswa.2019.05.042).

Bravo, C., Maldonado, S. and Weber, R. (2013), "Granting and managing loans for micro-entrepreneurs: new developments and practical experiences", *European Journal of Operational Research*, Vol. 227 No. 2, pp. 358-366, doi: [10.1016/j.ejor.2012.10.040](https://doi.org/10.1016/j.ejor.2012.10.040).

Breeden, J.L. (2021), "Survey of machine learning in credit risk", *SSRN Electronic Journal*, Vol. 17 No. 3, pp. 1-60, doi: [10.2139/ssrn.3616342](https://doi.org/10.2139/ssrn.3616342).

Capon, N. (1982), "Credit scoring systems: a critical analysis", *Journal of Marketing*, Vol. 46, pp. 82-91, No. Spring, doi: [10.2307/3203343](https://doi.org/10.2307/3203343).

Carta, S., Ferreira, A., Reforgiato Recupero, D. and Saia, R. (2021), "Credit scoring by leveraging an ensemble stochastic criterion in a transformed feature space", *Progress in Artificial Intelligence*, Vol. 10 No. 4, pp. 417-432, doi: [10.1007/s13748-021-00246-2](https://doi.org/10.1007/s13748-021-00246-2).

Chen, N., Ribeiro, B. and Chen, A. (2016), "Financial credit risk assessment: a recent review", *Artificial Intelligence Review*, Vol. 45 No. 1, pp. 1-23, doi: [10.1007/s10462-015-9434-x](https://doi.org/10.1007/s10462-015-9434-x).

Çiğşar, B. and Ünal, D. (2019), "Comparison of data mining classification algorithms determining the default risk", *Scientific Programming*, Vol. 2019 No. 8706505, pp. 1-9, doi: [10.1155/2019/8706505](https://doi.org/10.1155/2019/8706505).

Dastile, X. and Celik, T. (2021), "Making deep learning-based predictions for credit scoring explainable", *IEEE Access*, Vol. 9, pp. 50426-50440, doi: [10.1109/ACCESS.2021.3068854](https://doi.org/10.1109/ACCESS.2021.3068854).

Djeundje, V.B., Crook, J., Calabrese, R. and Hamid, M. (2021), "Enhancing credit scoring with alternative data", *Expert Systems with Applications*, Vol. 163, 113766, doi: [10.1016/j.eswa.2020.113766](https://doi.org/10.1016/j.eswa.2020.113766).

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. (2021), "How to conduct a bibliometric analysis: an overview and guidelines", *Journal of Business Research*, Vol. 133, pp. 285-296, doi: [10.1016/j.jbusres.2021.04.070](https://doi.org/10.1016/j.jbusres.2021.04.070).

Doumpos, M., Lemonakis, C., Niklis, D. and Zopounidis, C. (2018), *Analytical Techniques in the Assessment of Credit Risk: an Overview of Methodologies and Applications*, 1st ed., Springer, Switzerland, AG.

Durand, D. (1941), *Risk Elements in Consumer Instalment Financing*, National Bureau of Economic Research, Cambridge, MA, No. dura41-1, pp. 189-201, available at: <https://www.nber.org/books-and-chapters/risk-elements-consumer-instalment-financing>

Eisenbeis, R.A. (1978), "Problems in applying discriminant analysis in credit scoring models", *Journal of Banking and Finance*, Vol. 2 No. 3, pp. 205-219, doi: [10.1016/0378-4266\(78\)90012-2](https://doi.org/10.1016/0378-4266(78)90012-2).

Finlay, S. (2009), "Are we modelling the right thing? The impact of incorrect problem specification in credit scoring", *Expert Systems with Applications*, Vol. 36 No. 5, pp. 9065-9071, doi: [10.1016/j.eswa.2008.12.016](https://doi.org/10.1016/j.eswa.2008.12.016).

Finlay, S. (2010), "Credit scoring for profitability objectives", *European Journal of Operational Research*, Vol. 202 No. 2, pp. 528-537, doi: [10.1016/j.ejor.2009.05.025](https://doi.org/10.1016/j.ejor.2009.05.025).

Goh, R.Y. and Lee, L.S. (2019), "Credit scoring: a review on support vector machines and metaheuristic approaches", *Advances in Operations Research*, Vol. 2019 No. 8706505, pp. 1-31, doi: [10.1155/2019/1974794](https://doi.org/10.1155/2019/1974794).

Gunnarsson, B.R., vanden Broucke, S., Baesens, B., Óskarsdóttir, M. and Lemahieu, W. (2021), "Deep learning for credit scoring: do or don't?", *European Journal of Operational Research*, Vol. 295 No. 1, pp. 292-305, doi: [10.1016/j.ejor.2021.03.006](https://doi.org/10.1016/j.ejor.2021.03.006).

Ince, H. and Aktan, B. (2009), "A comparison of data mining techniques for credit scoring in banking: a managerial perspective", *Journal of Business Economics and Management*, Vol. 10 No. 3, pp. 233-240, doi: [10.3846/1611-1699.2009.10.233-240](https://doi.org/10.3846/1611-1699.2009.10.233-240).

Kang, Y., Jia, N., Cui, R. and Deng, J. (2021), "A graph-based semi-supervised reject inference framework considering imbalanced data distribution for consumer credit scoring", *Applied Soft Computing*, Vol. 105, 107259, doi: [10.1016/j.asoc.2021.107259](https://doi.org/10.1016/j.asoc.2021.107259).

Kozeny, V. (2015), "Genetic algorithms for credit scoring: alternative fitness function performance comparison", *Expert Systems with Applications*, Vol. 42 No. 6, pp. 2998-3004, doi: [10.1016/j.eswa.2014.11.028](https://doi.org/10.1016/j.eswa.2014.11.028).

Kozodoi, N., Jacob, J. and Lessmann, S. (2022), "Fairness in credit scoring: assessment, implementation and profit implications", *European Journal of Operational Research*, Vol. 297 No. 3, pp. 1083-1094, doi: [10.1016/j.ejor.2021.06.023](https://doi.org/10.1016/j.ejor.2021.06.023).

Kozodoi, N., Lessmann, S., Papakonstantinou, K., Gatsoulis, Y. and Baesens, B. (2019), "A multi-objective approach for profit-driven feature selection in credit scoring", *Decision Support Systems*, Vol. 120, pp. 106-117, doi: [10.1016/j.dss.2019.03.011](https://doi.org/10.1016/j.dss.2019.03.011).

Krichene, A. (2017), "Using a naive Bayesian classifier methodology for loan risk assessment: evidence from a Tunisian commercial bank", *Journal of Economics, Finance and Administrative Science*, Vol. 22 No. 42, pp. 3-24, doi: [10.1108/JEFAS-02-2017-0039](https://doi.org/10.1108/JEFAS-02-2017-0039).

Kruppa, J., Schwarz, A., Arminger, G. and Ziegler, A. (2013), "Consumer credit risk: individual probability estimates using machine learning", *Expert Systems with Applications*, Vol. 40 No. 13, pp. 5125-5131, doi: [10.1016/j.eswa.2013.03.019](https://doi.org/10.1016/j.eswa.2013.03.019).

Laborda, J. and Ryoo, S. (2021), "Feature selection in a credit scoring model", *Mathematics*, Vol. 9 No. 7, pp. 1-22, doi: [10.3390/math9070746](https://doi.org/10.3390/math9070746).

Leonard, K.J. (1992), "Credit-scoring models for the evaluation of small-business loan applications", *IMA Journal of Management Mathematics*, Vol. 4 No. 1, pp. 89-95, doi: [10.1093/imaman/4.1.89](https://doi.org/10.1093/imaman/4.1.89).

Lessmann, S., Baesens, B., Seow, H.V. and Thomas, L.C. (2015), "Benchmarking state-of-the-art classification algorithms for credit scoring: an update of research", *European Journal of Operational Research*, Vol. 247 No. 1, pp. 124-136, doi: [10.1016/j.ejor.2015.05.030](https://doi.org/10.1016/j.ejor.2015.05.030).

Li, Y. and Chen, W. (2020), "A comparative performance assessment of ensemble learning for credit scoring", *Mathematics*, Vol. 8 No. 10, pp. 1-19, doi: [10.3390/math8101756](https://doi.org/10.3390/math8101756).

Li, Z., Zhang, J., Yao, X. and Kou, G. (2021), "How to identify early defaults in online lending: a cost-sensitive multi-layer learning framework", *Knowledge-Based Systems*, Vol. 221, 106963, doi: [10.1016/j.knosys.2021.106963](https://doi.org/10.1016/j.knosys.2021.106963).

Lim, W.M., Kumar, S. and Ali, F. (2022), "Advancing knowledge through literature reviews: 'what', 'why', and 'how to contribute'", *The Service Industries Journal*, Vol. 42 Nos 7-8, pp. 481-513, doi: [10.1080/02642069.2022.2047941](https://doi.org/10.1080/02642069.2022.2047941).

Lin, W.Y., Hu, Y.H. and Tsai, C.F. (2012), "Machine learning in financial crisis prediction: a survey", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 42 No. 4, pp. 421-436, doi: [10.1109/TSMCC.2011.2170420](https://doi.org/10.1109/TSMCC.2011.2170420).

Liu, J. and Bo, S. (2011), "Naive Bayesian classifier based on genetic simulated annealing algorithm", *Procedia Engineering*, Vol. 23, pp. 504-509, doi: [10.1016/j.proeng.2011.11.2538](https://doi.org/10.1016/j.proeng.2011.11.2538).

Louzada, F., Ara, A. and Fernandes, G.B. (2016), "Classification methods applied to credit scoring: systematic review and overall comparison", *Surveys in Operations Research and Management Science*, Vol. 21 No. 2, pp. 117-134, doi: [10.1016/j.sorms.2016.10.001](https://doi.org/10.1016/j.sorms.2016.10.001).

Maldonado, S., Bravo, C., López, J. and Pérez, J. (2017), "Integrated framework for profit-based feature selection and SVM classification in credit scoring", *Decision Support Systems*, Vol. 104, pp. 113-121, doi: [10.1016/j.dss.2017.10.007](https://doi.org/10.1016/j.dss.2017.10.007).

Marqués, A.I., García, V. and Sánchez, J.S. (2013), "A literature review on the application of evolutionary computing to credit scoring", *Journal of the Operational Research Society*, Vol. 64 No. 9, pp. 1384-1399, doi: [10.1057/jors.2012.145](https://doi.org/10.1057/jors.2012.145).

Nalić, J. and Martinovic, G. (2020), "Building a credit scoring model based on data mining approaches", *International Journal of Software Engineering and Knowledge Engineering*, Vol. 30 No. 2, pp. 147-169, doi: [10.1142/S0218194020500072](https://doi.org/10.1142/S0218194020500072).

Onay, C. and Ozturk, E. (2018), "A review of credit scoring research in the age of big data", *Journal of Financial Regulation and Compliance*, Vol. 32 No. 10, pp. 91-100, doi: [10.1108/JFRC-06-2017-0054](https://doi.org/10.1108/JFRC-06-2017-0054).

Orgler, Y.E. (1970), "Scoring model for commercial loans", *Journal of Money, Credit and Banking*, Vol. 2 No. 4, pp. 435-445, doi: [10.2307/1991095](https://doi.org/10.2307/1991095).

Paul, J. and Criado, A.R. (2020), "The art of writing literature review: what do we know and what do we need to know?", *International Business Review*, Vol. 29 No. 4, 101717, doi: [10.1016/j.ibusrev.2020.101717](https://doi.org/10.1016/j.ibusrev.2020.101717).

Řežáč, M. (2015), "ESIS2: information value estimator for credit scoring models", *Computational Economics*, Vol. 45 No. 2, pp. 303-322, doi: [10.1007/s10614-014-9424-0](https://doi.org/10.1007/s10614-014-9424-0).

Roa, L., Correa-Bahnsen, A., Suárez, G., Cortés-Tejada, F., Luque, M.A. and Bravo, C. (2021), "Super-app behavioral patterns in credit risk models: financial, statistical and regulatory implications", *Expert Systems with Applications*, Vol. 169, 114486, doi: [10.1016/j.eswa.2020.114486](https://doi.org/10.1016/j.eswa.2020.114486).

Roy, P.K. and Shaw, K. (2021a), "A credit scoring model for SMEs using AHP and TOPSIS", *International Journal of Finance and Economics*, No. December 2020, pp. 1-20, doi: [10.1002/ijfe.2425](https://doi.org/10.1002/ijfe.2425).

Roy, P.K. and Shaw, K. (2021b), "A multicriteria credit scoring model for SMEs using hybrid BWM and TOPSIS", *Financial Innovation*, Vol. 7 No. 1, pp. 1-27, doi: [10.1186/s40854-021-00295-5](https://doi.org/10.1186/s40854-021-00295-5).

Roy, P.K. and Shaw, K. (2022), "Modelling a sustainable credit score system (SCSS) using BWM and fuzzy TOPSIS", *International Journal of Sustainable Development and World Ecology*, Vol. 29 No. 3, pp. 195-208, doi: [10.1080/13504509.2021.1935360](https://doi.org/10.1080/13504509.2021.1935360).

Salcedo, N.U. (2021a), "Editorial: review and roadmap from the last 10 years (2010-2020)", *Journal of Economics, Finance and Administrative Science*, Vol. 26 No. 51, pp. 2-6, doi: [10.1108/JEFAS-06-2021-271](https://doi.org/10.1108/JEFAS-06-2021-271).

Salcedo, N.U. (2021b), "Editorial: an upcoming 30th anniversary encouraging the papers' publication", *Journal of Economics, Finance and Administrative Science*, Vol. 26 No. 52, pp. 178-181, doi: [10.1108/JEFAS-11-2021-329](https://doi.org/10.1108/JEFAS-11-2021-329).

Sariannidis, N., Papadakis, S., Garefalakis, A., Lemonakis, C. and Kyriaki-Arkyro, T. (2020), "Default avoidance on credit card portfolios using accounting, demographical and exploratory factors: decision making based on machine learning (ML) techniques", *Annals of Operations Research*, Vol. 294 Nos 1-2, pp. 715-739, Springer US, doi: [10.1007/s10479-019-03188-0](https://doi.org/10.1007/s10479-019-03188-0).

Serrano-Cinca, C. and Gutiérrez-Nieto, B. (2016), "The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending", *Decision Support Systems*, Vol. 89, pp. 113-122, doi: [10.1016/j.dss.2016.06.014](https://doi.org/10.1016/j.dss.2016.06.014).

Sinha, A.P. and Zhao, H. (2008), "Incorporating domain knowledge into data mining classifiers: an application in indirect lending", *Decision Support Systems*, Vol. 46 No. 1, pp. 287-299, doi: [10.1016/j.dss.2008.06.013](https://doi.org/10.1016/j.dss.2008.06.013).

Šuštersič, M., Mramor, D. and Zupan, J. (2009), "Consumer credit scoring models with limited data", *Expert Systems with Applications*, Vol. 36 No. 3, pp. 4736-4744, PART 1, doi: [10.1016/j.eswa.2008.06.016](https://doi.org/10.1016/j.eswa.2008.06.016).

Trivedi, S.K. (2020), "A study on credit scoring modeling with different feature selection and machine learning approaches", *Technology in Society*, Vol. 63 No. 2017, 101413, doi: [10.1016/j.techsoc.2020.101413](https://doi.org/10.1016/j.techsoc.2020.101413).

Verbraken, T., Bravo, C., Weber, R. and Baesens, B. (2014), "Development and application of consumer credit scoring models using profit-based classification measures", *European Journal of Operational Research*, Vol. 238 No. 2, pp. 505-513, doi: [10.1016/j.ejor.2014.04.001](https://doi.org/10.1016/j.ejor.2014.04.001).

Vukovic, S., Delibasic, B., Uzelac, A. and Suknovic, M. (2012), "A case-based reasoning model that uses preference theory functions for credit scoring", *Expert Systems with Applications*, Vol. 39 No. 9, pp. 8389-8395, doi: [10.1016/j.eswa.2012.01.181](https://doi.org/10.1016/j.eswa.2012.01.181).

Watson, R., T. and Webster, J. (2020), "Analysing the past to prepare for the future: writing a literature review a roadmap for release 2.0", *Journal of Decision Systems*, Vol. 29, pp. 129-147, doi: [10.1080/12460125.2020.1798591](https://doi.org/10.1080/12460125.2020.1798591).

Xia, Y., Li, Y., He, L., Xu, Y. and Meng, Y. (2021), "Incorporating multilevel macroeconomic variables into credit scoring for online consumer lending", *Electronic Commerce Research and Applications*, Vol. 49 No. 9, 101095, doi: [10.1016/j.elerap.2021.101095](https://doi.org/10.1016/j.elerap.2021.101095).

Corresponding author

Marcelo Seido Nagano can be contacted at: drnagano@usp.br