Machine Learning for Personal Credit Evaluation: A Systematic Review

CANO CHUQUI JORGE
Faculty of Engineering and Architecture
Universidad Privada César Vallejo
Av. Del Parque 640, San Juan de Lurigancho 15434
PERÚ

OGOSI AUQUI JOSÉ ANTONIO
Faculty of Engineering
Universidad Tecnológica del Perú
Av. Arequipa 265, Cercado de Lima 15046
PERÚ

GUADALUPE MORI, VICTOR HUGO
Faculty of Engineering
Universidad Privada San Juan Bautista
Ex Hacienda Villa, Av. José Antonio Lavalle s/n, Chorrillos 15067
PERÚ

OBANDO PACHECO, DAVID HUGO
Faculty of Engineering
Universidad Peruana de Ciencias Aplicadas
Prolongación Primavera 2390, Lima 15023
PERÚ

Abstract: - The importance of information in today's world as it is a key asset for business growth and innovation. The problem that arises is the lack of understanding of knowledge quality properties, which leads to the development of inefficient knowledge-intensive systems. But knowledge cannot be shared effectively without effective knowledge-intensive systems. Given this situation, the authors must analyze the benefits and believe that machine learning can benefit knowledge management and that machine learning algorithms can further improve knowledge-intensive systems. It also shows that machine learning is very helpful from a practical point of view. Machine learning not only improves knowledge-intensive systems but has powerful theoretical and practical implementations that can open up new areas of research. The objective set out is the comprehensive and systematic literature review of research published between 2018 and 2022, these studies were extracted from several critically important academic sources, with a total of 73 short articles selected. The findings also open up possible research areas for machine learning in knowledge management to generate a competitive advantage in financial institutions.

*Key-Words: m*achine learning, credit scoring, risk assessment, algorithms, artificial intelligence.

Received: March 25, 2021. Revised: April 14, 2022. Accepted: May 12, 2022. Published: July 1, 2022.

1 Introduction

Time analysis is critical because financial institutions consistently implement the credit scoring model over time. Therefore, due to the complexity and flexibility of the training process, ML methods can be more sensitive to overtime disturbances. It is susceptible to overfitting problems and may become unstable over time.

In addition, there is a lack of ability to account for time complexity in actual business operations. Everything depends on the model life cycle, from data collection, model development, and validation to the final model. This model can present new challenges and uncertainties to the generalization of

ML models over time. ML production representations are generally scheduled for the model to be applied at least one year after the development data period.

The motivation for this study stems from the fact that to access loans, applicants must remain enrolled in a regular socioeconomic dimension essential to repay the money. Applicants are concerned about whether they are good repayers and to ensure that they can meet the requirements they expect based on their client activities. The study contributes not only to the collection of citations and empirical evaluation of studies on personal credit evaluation but also addresses the measurement of Machine Learning methods to influence bank lending processes in a formalized financial institution with cash flow, which provides a particular interest rate and certain loan application requirements.

This survey includes empirical reviews of personal credit score surveys and extends the literature that incorporates information provided by banks or financial institutions regarding the services offered. This article is divided into Chapter 1-Introduction, Chapter 2-Methodology, Chapter 3-Results a, and Chapter 4-Conclusions for future work.

2 Methodology

This paper conducts a research study on the personal credit scoring process using machine learning. This work is called a preliminary survey because it is the first time it has been conducted. We believe that this review is part of an evidence-based personal credit assessment. It aims to collect research that can be used to implement machine learning methods, models, or approaches. The review articles are elaborated based on a methodology, which gives us very clear steps to make a review article, that methodology is called the "Systematic Literature Review", it is a methodology to make articles, and therefore, it is a starting point like all research, to establish the problems of something unknown or something that you want to optimize, and the objectives that you want to achieve. The review tells us what are the problems we have and what things we want to achieve as point number one, then, we look for good sources of information, because it is a systematic review of the literature, but in this case, the literature that we review, are scientific articles, we do not review web pages, books, or theses for this work, as point number two. Then, we have to do a systematic, orderly, and rigorous search with search equations. We can find hundreds, thousands, or hundreds of thousands, but it has to be done with a systematic, orderly and rigorous filter, and we will be left with 70 to 80 articles as point number three. After these articles are reviewed, we answer the questions as point number four. When we generate the answers to these questions, we will have the fundamental input to prepare the review article, which has all the parts of an article as point number five; its title, its authors, its affiliation, its summary, its keywords, the sources of information, likewise, it has the result of showing the answers raised in point number one. This is the overview of how review articles are developed, as shown in Figure 1.

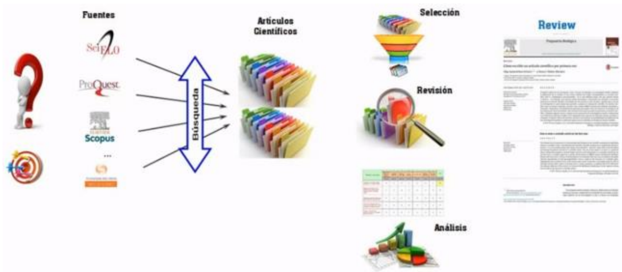

Fig. 1: SRL methodology.

2.1 Formulation of Questions and Objectives

There are two specific problems tracked in the meta-analysis presented in Table 1.

Table 1. Research questions

Research Question Motivation	Motivation
RQ1: What are the methods used to apply machine learning for personal credit assessment?	Demonstrate which methods are used to apply machine learning for personal credit evaluation.
RQ2: What complementary methods are used to apply machine learning for personal credit assessment?	Demonstrate which complementary methods apply machine learning for personal credit assessment.

2.2 Search Equation

2.2.1 Search Sources

Systematic literature searches take information from 11 search sources and use search terms to create general research equations.

- *Taylor & Francis*
- *ProQuest*
- *WOS*
- *IEEE Xplore*
- *ScienceDirect*
- *Scopus*
- *EBSCOhost*
- *Wiley Online Library*
- *IOP*
- *ERIC*
- *Google Scholar*

2.2.2 Search Equations

The search terms used in this study, applying the terms of use of each source, are shown in Table 2.

Table 2. Search Equations

Fuente	Generic research equation
Taylor & Francis	[All: machine learning] AND [All: personal credit evaluation] AND [[All: method] OR [All: methodology] OR [All: model]]
ProQuest	(Machine Learning) AND (Personal Credit Evaluation) AND ((method OR methodology OR model))
WOS	(Machine Learning) AND (Personal Credit Evaluation) AND (method OR methodology OR model)
IEEE Xplore	("All Metadata":Machine Learning) AND ("All Metadata":Personal Credit Evaluation) AND ("All Metadata":method OR "All Metadata":methodology OR "All Metadata":model)
ScienceDirect	("Machine Learning") AND ("Personal Credit Evaluation") AND ("method" OR "methodology" OR "model")
Scopus	ALL ((machine AND learning) AND (personal AND credit AND evaluation) AND ((method OR methodology OR model)))
EBSCOhost	TX Machine Learning AND TX Personal Credit Evaluation AND TX (method OR methodology OR model)
Wiley Online Library	"Machine Learning" anywhere and "Personal Credit Evaluation" anywhere and "method OR methodology OR model" anywhere
IOP	Machine Learning AND Personal Credit Evaluation AND (method OR methodology OR model)
ERIC	Machine Learning AND (Personal Credit Evaluation) AND (method OR methodology OR model)
Google Scholar	("Machine Learning" AND "Personal Credit Evaluation" AND method OR methodology OR model)

2.3 Studies Identified

2.3.1 Integrated Chart of Numbers of Results by Source

As shown in Figure 2, "n" was found in the 11-source search formula where (N = 164,793) articles were published in 11 sources. The overall results published in 11 sources representing the total number of pieces sampled were reported.

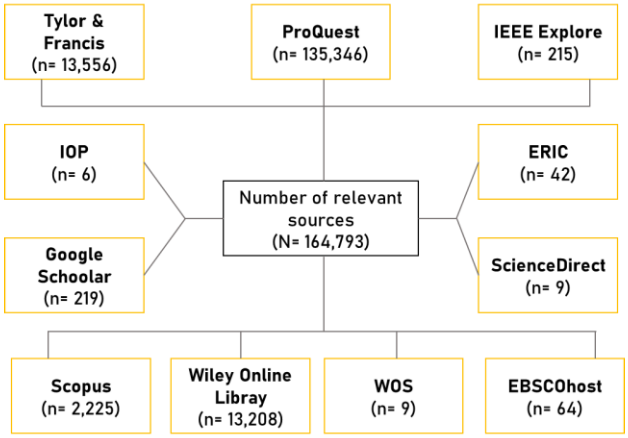

Fig. 2: Consolidated Graph of the Number of Results by Source.

2.3.2 Consolidated Matrix of Number of Results by Source

Table 3 shows two search terms for each source, and you can apply the search terms and their logical relationships to get the total for each reference. This will eventually show 11 stars with a total result of 164,793 for the articles found.

Table 3. Consolidated matrix of the resulting number by source

Source	Generic research equation	Total
Taylor & Francis	[All: machine learning] AND [All: personal credit evaluation] AND [[All: method] OR [All: methodology] OR [All: model]]	13,556
ProQuest	(Machine Learning) AND (Personal Credit Evaluation) AND ((method OR methodology OR model))	135,346
WOS	(Machine Learning) AND (Personal Credit Evaluation) AND (method OR methodology OR model)	9
IEEE Xplore	("All Metadata":Machine Learning) AND ("All Metadata":Personal Credit Evaluation) AND ("All Metadata":method OR "All Metadata":methodology OR "All Metadata":model)	215
ScienceDirect	("Machine Learning") AND ("Personal Credit Evaluation") AND ("method" OR "methodology" OR "model")	9
Scopus	ALL ((machine AND learning) AND (personal AND credit AND evaluation) AND ((method OR methodology OR model)))	2,225
EBSCOhost	TX Machine Learning AND TX Personal Credit Evaluation AND TX (method OR methodology OR model)	64
Wiley Online Library	"Machine Learning" anywhere and "Personal Credit Evaluation" anywhere and "method OR	13,208

	methodology OR model" anywhere	
IOP	Machine Learning AND Personal Credit Evaluation AND (method OR methodology OR model)	6
ERIC	Machine Learning AND (Personal Credit Evaluation) AND (method OR methodology OR model)	42
Google Scholar	("Machine Learning" AND "Personal Credit Evaluation" AND method OR methodology OR model)	219

2.4 Exclusion Criteria

2.4.1 Exclusion Criteria

To avoid selection bias and strictly apply the criteria, we created a list of objective exclusion criteria. As we apply these criteria, we are leaving out a large number of articles. When we finish the last criterion, we are left with a more manageable number of documents, to have around 80 articles that will serve to answer the questions of this research with distance and objectivity. These criteria depended on the study objectives, for which a total of 8 exclusion criteria were identified, as shown in Table 4 below.

Table 4. Exclusion Criteria

N	Descripción	Does it comply?
CE1	Articles are older than five years O	✓
CE2	Articles are not written in the English language OR	✓
CE3	The documents are not of article type.	✓
CE4	The titles and keywords of the articles are not very appropriate.	✓
CE5	The articles are not unique.	✓
CE6	The abstract of the articles is not very relevant O	✓
CE7	The reports do not mention a methodology, model, or method O	✓
CE8	The proposed solution does not apply to Customer Service OR Teaching_Learning, Recruitment, OR Political campaign management.	X

Of the eight exclusion criteria, five were selected to evaluate and investigate the articles found.
When applying the criteria:

- *CE1: excluded 897 items older than five years, resulting in 51,896 items using the first criterion.*
- *In CE2, we excluded 1,903 items that were not OR in English and used a second criterion to exclude 49,993 items.*
- *CE3: 38,238 items were excluded instead of article-type documents for 11,755 items when the third criterion was applied.*
- *CE4: 11,664 items were excluded, their titles and keywords were not very appropriate for the article, and 91 items were obtained using the fourth criterion.*
- *CE5: 18 items are excluded and ambiguous; applying the fifth criterion totals 73 items and processing them for research.*

2.4.2 The Aggregate Number of Results when Exclusion Criteria Apply

Table 5 shows the primary studies. The filters applied have two exclusions for each (Filter 1: CE1CE2, Filter 2: CE3CE4, and Filter 3: CE5), with the final filter being the element after maintaining the exclusion criteria.

Table 5. The aggregate number of results when exclusion criteria apply

Source	Initials	Filter 1: CE1-CE2	Filter 2: CE3-CE4	Filter 3: CE5
Taylor & Francis	13,556	1,732	1,545	6
ProQuest	135,346	46,459	8.146	16
WOS	9	9	8	4
IEEE Xplore	215	8	8	8
ScienceDirect	9	7	7	2
Scopus	2,225	1,204	702	25
EBSCOhost	64	18	18	7
Wiley Online Library	13,208	2,060	1,313	3
IOP	6	6	4	2
ERIC	42	3	2	0
Google Scholar	219	219	2	0
Total	164,793	51725	11753	73

2.5 Review of the Studies

2.5.1 The Aggregate Number of Results when Exclusion Criteria Apply

In this section, the answers that appear in each study are identified from the full-text studies so that the page on which the answers appear can be viewed promptly, where the conceptual clarification of the research topic can also be started, such as it is mentioned in Table 6.

Table 6. The aggregate number of results when exclusion criteria apply

N	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6
[1]	6,12, 14	1	1,2	1	10,14,1,8	11,16
[2]	1,4	1	1	1		2
[3]		1		1		2
[4]		1	1	1		3
[5]		1				
[6]	2,11	1	1	1	3	2,3,8
[7]		1		1		2
[8]		1	1	1		2

Cano Chuqui Jorge, Ogosi Auqui José Antonio,
Guadalupe Mori Victor Hugo,
Obando Pacheco David Hugo

[9]	3	1	2	1	6	8
[10]		1	1	1		
[11]		1	1	1		6
[12]				1		
[13]		1	1,2	1	13,14	4
[14]		1				2
[15]		1		1	1	
[16]		1		1	2	1,2
[17]	8	1	1	1		
[18]		1	1	1		
[19]		1	1	1	6	
[20]	5	1	1	1	1	3
[21]		1		1		1
[22]	2,3,4	1	1,2	2	6	1
[23]		1	1	1	1	
[24]	1	1	1	1		3
[25]		1		1	2,3	
[26]		1	1	1		2
[27]		1	1,2	1	4	4
[30]		1	1	1	6	3
[31]		1	1	1		5
[32]		2	2	2		
[33]		1	1	1		
[34]	10	1	1	1	6	12
[35]		1	1	1		
[36]	1,2	1	1	1		1,2
[37]		1				
[38]	2	1	1,2	1		2,3,4
[39]		1	1	1		
[40]		1	1	1		
[41]	2	1	1	1		
[42]		1	1	1		2
[43]	6	1	1,2	1	9	
[44]		1	1	1		2,3
[45]		1	1	1		9
[46]	1,2	1	1	1		2,3
[47]	2,3	1	1	1		4,5
[48]		1	1	1		
[49]		1	1	1	5	4
[50]		1	1,2	1	1,4,6,7	
[51]	8,12	1	1,2	1		6,8,10,11,
[52]	3,7,1,0	12,14				
[53]		1	1,2	1	5,7,11	7,10
[54]	16	1		2		4,16
[55]	4	1	1,2	1	6	4
[56]		1	1,2	1	6	14,19
[57]		1	1,2	1	16	17
[58]	8,9	1	1	1		1,2
[59]	1	2	1	1	9,12	8
[60]		1	1	1		
[61]	3,4	1		1,2		
[63]		1	1	1		
[64]	2	1	1	1		14,15
[65]		1	1	1		
[66]	2,3	1	1	1		2
[67]	4	1	1,2	1		4,8,9,10
[68]		1	1	1		2
[69]		1	1	1		1,2
[70]		1		1	1,2	4
[71]		1	1	1		4
[72]		1	1,2	1	5,6,7	
[73]		2	1	2		

3 Results

Answers to RQs.

RQ1: What are the methods used to apply machine learning for personal credit assessment?

Table 7. Definitions of most commonly used ML methods for credit evaluation

N	ML Method Definitions	References	Total number of articles
1	Supervised learning	[4] [5] [7] [16] [19] [33] [34] [37] [40] [41] [45] [46] [49] [53] [63] [68]	16 (22%)
2	Unsupervised learning	[1] [5] [7] [10] [13] [16] [27] [28] [31] [37] [39] [41] [46] [56] [61] [63] [66] [69] [70] [78]	20 (27%)

The definition of supervised learning is used in 20 (27%) items and also emphasizes complementary methods. Finally, the few articles mentioned it, totaling 16 (22%). The following are the most commonly used machine learning (ML) methods. The meanings used to define ML methods are unsupervised learning and unsupervised learning.

RQ2: What complementary methods are used to apply machine learning for personal credit assessment?

Due to the advanced technology associated with big data, data availability, and computational power, most banks or credit unions are innovating their business models [1]. Credit scoring analytics is an effective technique for assessing credit risk and is one of the major research areas in the banking industry [2]. Neural networks are one of the most widely used methods to obtain credit scores [3]. Using machine learning through modeling and prediction, both models were trained on real credit card transaction datasets [4]. Currently, these functions are performed manually and are subject to expert evaluation [5]. Available multiclass classifiers, such as random forest algorithms, can perform this task very well, using available customer data [6]. He suggested using this credit scoring from unconventional data sources for online lenders to enable them to investigate and detect changes in customer behavior over time and target unsecured customers based on their claims data [7]. In a rapidly developing economy, credit plays a very important role. Several prediction models have been developed to predict credit risk with many different variables [8]. The information provided by the candidates constitutes the variables of our analysis [9]. Machine learning algorithms have dominated many different industries [10]. Several aspects need to be taken into account to make credit scoring models understandable and to provide a framework for making "black box" machine learning models transparent, audible, and solvable [11]. With the rapid development of corporate lending in China, the creation of a high-risk credit scoring system has become an important measure of financial guarantees [12]. Non-performing loans are a serious problem in the banking sector. Credit rating models used logistic regression and linear discriminant analysis to identify potential defaulters [13]. Financial institutions are faced with the need to assess the creditworthiness of the borrower applying for the loan. The best results are observed for randomization [14]. The best prediction results are obtained using conventional synthesis techniques, namely, packing, random forest, and Bo improvement [15]. Replacing subjective analysis with objective credit analysis using deterministic models would benefit Brazilian credit unions [16]. The model quadruples the accepted default rate to break even from 8% to 32% [17]. Several 'rejection inference' methods attempt to exploit the available data for candidates that were rejected during the learning process [18]. It then quantifies each indicator and defines criteria for evaluating the assessment results [19]. It then quantifies each indicator and defines criteria for evaluating the assessment results [20]. A good credit rating decision support system allows telecom operators to measure the creditworthiness of subscribers in detail [21]. Current research on credit scoring in microfinance is limited to genetic and regression algorithms, which excludes newer machine learning algorithms [22]. General psychometric modeling is effective in predicting lifetime consumer mortgage behavior [23]. The development of accurate analytical credit scoring models has become an important goal of financial institutions [24]. Choosing the optimal techniques, whether attribute selection techniques, attribute assignment techniques, or ML resampling mechanisms and classifiers to support the coverage decision is challenging and doable. The integrated SVM-Logistic model is complementary and has a high evaluation density [26]. For domain adaptation problems, transfer learning techniques are often used; however, it is very difficult to run accurate predictions of unknown domain datasets in CSM because name distributions may be different depending on domain properties [27]. The Dempster-Shafer synthesis method allows accurate labeling by exploiting the advantages of both methods [28]. Rural credit is one of the most important inputs for agricultural production in the world. However, it is the banking or non-banking institutions that will decide how to apply this advanced technology to reduce human biases in the credit decision-making process [29]. This particular model performs better than multilayer backpropagation networks, probabilistic neural networks, radial basis functions, and regression trees, as well as other advanced classifiers [30]. A credit default prediction model based on a complex graph network can reflect nonlinear relationships between borrower characteristics and default risk and higher-order relationships between borrowers [31]. The XGB model has obvious advantages in both feature selection and classification performance over logistic regression and the other three tree-based models [32]. Application of radial-based neural network model along with optimal segmentation algorithm in credit scoring model of personal loans to banks or other financial institutions [33]. In a nonlinear method that eliminates the obvious subjective and artificial factors, the evaluation results are more objective and effective [34]. The multiple averaging methods can effectively reduce the diversity of the results, and the accuracy will not be significantly reduced by the different proportions of training and prediction sets [35]. In emerging markets, there is a gap between having a credit rating or credit score and having no credit history [36]. It assists borrowers in the fundraising process, allowing any number and size of lenders to participate [37]. The training of the model will be performed using machine learning-based algorithms such as; Random Forest, Extreme Gradient, Boost, Mild Gradient Boost, Adaboost, and ExtraTrees [38]. Measuring credit risk is essential for financial institutions due to the high level of risk associated with bad credit decisions. The recent Basel Accord specifies that reserve requirements have increased according to risk [39]. A credit score is a central component of a corporation's lending. The combination of credit scoring and machine learning can integrate a relatively complete functionality into the credit scoring process [40]. The model structure is determined by hyperparameters, aiming to address the time-consuming and labor-intensive manual adjustment problem, and the optimization method is used for adjustment [41]. The SCSRF model parameters were optimized by grid search [42]. Credit ratings are becoming increasingly important in the financial sector. By changing the number of nodes in a Spark cluster, the execution time of these algorithms is compared and the analysis of variance compares the execution time of each algorithm with an increasing number of nodes [43]. To reduce the negative impact of the unbalanced dataset on the performance of the credit rating model, the SMOTE technique was used to rebalance the target training dataset [44]. There are relatively few transparency models that take interpretability and clarity into account [45]. To identify eligible end customers from defaulters, a credit scoring model is used to reliably screen the credit data using a combination of Min-Max normalization and linear regression. [Forty-six]. The support vector machine is the most widely used classifier for credit rating and, although the system performs well, it does not apply collateral approaches [47]. Corporate insolvency has significant negative effects on the economy. The RF algorithm shows utility in credit risk management [49]. Maximum machine learning is used as a scoring tool for the credit risk assessment model [50]. Intensive machine learning enables multilayer neural networks to perform operations to facilitate operations and business dynamics [51]. The proposed P2P personal credit scoring model is superior to both individual models and other sets of criteria [52]. The influence of controlling shareholder characteristics on corporate risk has been a popular topic of debate in academic and theoretical circles [53]. Online personal loans are a new form of lending [54]. The method assigns terms to the embedding space, groups the linguistically related terms into semantic clusters, and then selects the flexible semantic items corresponding to the semantic clusters [55]. Object selection techniques or object selection techniques should be incorporated in predictive model building to improve prediction performance [56]. According to the application scenario of credit scoring of personal credit data, the test data set is cleaned, the separated data is coded as HOT and the data is normalized [57]. The incorporation of macroeconomic variables can improve the performance of existing models [58]. To confirm the effectiveness of the proposed credit rating model, experiments were conducted with realistic credit data sets for comparison [59]. The accuracy and kappa values for all four methods exceeded 90 %, and RF outperformed other rating models [60]. Three hybrid AI models have been studied, including decision tree: artificial neural network, decision tree: logistic regression, and decision tree: dynamic Bayes network [61]. Online financial institutions lack effective methods to evaluate personal credit, which seriously hinders the development of personal credit business [62]. More and more attention is paid to the use of machine learning algorithms to predict people's credit ratings in the era of artificial intelligence [63]. Because there are many unusual users in these data, they are "real but fake data" of personal credit rating [64]. The model built there performed well on some of the scales used to compare it to other commonly used raters [65]. Blockchain, decision trees, and other technologies can effectively improve the transparency of personal credit information in the field of Internet finance [66]. Based on the seed neural network, a predictive model of the probability of granting formal credit to farmers was built [67]. Geospatial data collection from location-based services can provide location evidence during spatial information analysis [68]. Due to the rapid increase in the number of personal loan applications, the importance of credit risk assessment for practitioners and researchers is increasing [69]. Using blockchain, decision trees, and other technologies, this paper designs the credit rating process and establishes the individual credit rating technology [70]. CNN is used to create a model to predict individual credit default, and ACC and AUC are taken as performance indicators for the model [71]. The model includes five dimensions participation, positivity, frequency, eligibility rate, and impact [72]. The feasibility analysis of the selected models is carried out through rigorous experiments with real data describing the client's ability to repay loans [73].

In Section II, a systematic literature search methodology was performed. Here you can see if the search was performed for the entire year (2018) onwards specified by the exclusion criteria (Table 4). We obtained a total of 73 journal articles and conference proceedings. The year with the most journals published was 2021, with 27 records and Scopus leads in quantity in this range of years with 25 publications, as shown in Table 8.

Table 8. Publication over the years

Source	**Last five years**					**Total**
	2018	**2019**	**2020**	**2021**	**2022**	
Taylor & Francis	1	1		1	3	6
ProQuest		1	5	10		16
WOS			2	2		4
IEEE Xplore	1	1	4	2		8
ScienceDirect		1			1	2
Scopus	1	4	11	7	2	25
EBSCOhost	1	1	2	1	2	7
Wiley Online Library				2	1	3
IOP				2		2
ERIC						0
Google Scholar						0
Total	4	9	24	27	9	73

Specific machine learning algorithms for classifying the quality of knowledge in the system (in this case, a decision tree algorithm using a training model [58]) work as follows: The algorithm recognizes whether the knowledge quality is high, medium, or low [8]. The knowledge quality attribute [63] is satisfied. This is determined by the training model itself and must be identified when creating the training model dataset [17]. Subsequently, the algorithm analyzes the knowledge. The algorithm continues with the following perspective when a perspective receives a score until all views receive a knowledge quality score [57].

For practitioners, the direct application of the results provided in the study should be done carefully using some of the techniques found (supervised and unsupervised learning). This evidence suggests that the field of study is developing promptly, so directly applying the techniques and tools used in the study will achieve the desired effect of implementing machine learning in financial institutions. In particular, we agree with [26], that there are several methods and theoretical models to apply data mining technology. We also believe that careful evaluation of the context in which the research is conducted is important to assess the generalizability of the findings to change in other potential contexts.

4 Conclusion

In conclusion, this study uses a systematic literature review (SLR). This iterative process combines all existing literature on a particular topic or research question. The goal of SLR is to solve a specific problem by examining and integrating the results of all state-of-the-art surveys that address two or more survey questions. The studies found are the best available in our time on the subject of the study. The most commonly used criterion is "unsupervised learning" to determine its basic effectiveness. This method identifies how far the current research on the use of machine learning has progressed. The review work has a total of 73 articles between journals and congresses. Similarly, it was found that the most suitable study varied between 16 and 17 pages to answer the highest number of RQs, being the publication medium Risk, as well as Expert Syst. Appl. and Math. Probl. Eng. being the journals with the highest production in the area of machine learning. It is claimed that RQ1 provides the basis for future comments on the applicability of personal credit assessments. Here are some machine learning methods to show how machine learning can be applied to an individual's credit score. This method identifies how far current research on machine learning has progressed. Future work will be devoted to improving the literature by also taking into account questions that have answers by adding a wider variety of configurations.

References:

[1] P. M. Addo, D. Guegan, and B. Hassani, “Credit risk analysis using machine and deep learning models,” *Risks*, vol. 6, no. 2, 2018, DOI: 10.3390/risks6020038.

[2] V. Aithal and R. D. Jathanna, “Credit risk assessment using machine learning techniques,” *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 1, pp. 3482–3486, 2019, DOI: 10.35940/ijitee.A4936.119119.

[3] S. Akkoç, “Exploring the nature of credit scoring: A neuro fuzzy approach,” *Fuzzy Econ. Rev.*, vol. 24, no. 1, pp. 3–24, 2019, DOI: 10.25102/fer.2019.01.01.

[4] M. Ala’raj, M. F. Abbod, M. Majdalawieh, and L. Jum’a, “A deep learning model for behavioural credit scoring in banks,” *Neural Comput. Appl.*, vol. 34, no. 8, pp. 5839–5866, 2022, DOI: 10.1007/s00521-021-06695-z.

[5] G. S. Alzamora, M. R. Aceituno-Rojo, and H. I. Condori-Alejo, “An Assertive Machine Learning Model for Rural Micro Credit Assessment in Peru,” *Procedia Comput. Sci.*, vol. 202, pp. 301–306, 2022, DOI: https://doi.org/10.1016/j.procs.2022.04.040.

[6] A. Ampountolas, T. N. Nde, P. Date, and C. Constantinescu, “A machine learning approach for micro-credit scoring,” *Risks*, vol. 9, no. 3, 2021, DOI: 10.3390/risks9030050.

[7] A. Ashofteh and J. M. Bravo, “A conservative approach for online credit scoring,” *Expert Syst. Appl.*, vol. 176, 2021, DOI: 10.1016/j.eswa.2021.114835.

[8] U. Bhuvaneswari and S. Sophia, “A predictive analysis of credit risk evaluation and the quality decision making using different predictive models,” *Int. J. Recent Technol. Eng.*, vol. 8, no. 1, pp. 1965–1969, 2019, [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068001945&partnerID=40&md5=63ad0475a06fe201cfa9d754158756c7.

[9] Z. Boz, D. Gunnec, S. I. Birbil, and M. K. Öztürk, “Reassessment and Monitoring of Loan Applications with Machine Learning,” *Appl. Artif. Intell.*, vol. 32, no. 9–10, pp. 939–955, 2018, DOI: 10.1080/08839514.2018.1525517.

[10] . L. Breeden, "A survey of machine learning in credit risk," *J. Credit Risk*, vol. 17, no. 3, pp. 1–62, 2021, DOI: 10.21314/JCR.2021.008.

[11] M. Bücker, G. Szepannek, A. Gosiewska, and P. Biecek, "Transparency, auditability, and explainability of machine learning models in credit scoring," *J. Oper. Res. Soc.*, vol. 73, no. 1, pp. 70–90, 2022, DOI: 10.1080/01605682.2021.1922098.

[12] M. Chen and X. Ma, "An Optimized BP Neural Network Model and Its Application in the Credit Evaluation of Venture Loans.*," Comput. Intell. Neurosci.*, pp. 1–9, 2022, [Online]. Available: https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=156652121&lang=es&site=eds-live.

[13] A. Chopra and P. Bhilare, "Application of Ensemble Models in Credit Scoring Models," *Bus. Perspect. Res.*, vol. 6, no. 2, pp. 129–141, 2018, DOI: 10.1177/2278533718765531.

[14] A. Coşer, M. M. Maer-Matei, and C. Albu, "Predictive models for loan default risk assessment," *Econ. Comput. Econ. Cybern. Stud. Res.*, vol. 53, no. 2, pp. 149–165, 2019, DOI: 10.24818/18423264/53.2.19.09.

[15] J. R. de Castro Vieira, F. Barboza, V. A. Sobreiro, and H. Kimura, "Machine learning models for credit analysis improvements: Predicting low-income families' default," *Appl. Soft Comput. J.*, vol. 83, 2019, DOI: 10.1016/j.asoc.2019.105640.

[16] D. A. V. de Paula, R. Artes, F. Ayres, and A. M. A. F. Minardi, "Estimating credit and profit scoring of a Brazilian credit union with logistic regression and machine-learning techniques," *RAUSP Manag. J.*, vol. 54, no. 3, pp. 321–336, 2019, DOI: https://doi.org/10.1108/RAUSP-03-2018-0003.

[17] B. Dushimimana, Y. Wambui, T. Lubega, and P. E. McSharry, "Use of Machine Learning Techniques to Create a Credit Score Model for Airtime Loans," *J. Risk Financ. Manag.*, vol. 13, no. 8, p. 180, 2020, DOI: https://doi.org/10.3390/jrfm13080180.

[18] A. Ehrhardt, C. Biernacki, V. Vandewalle, P. Heinrich, and S. Beben, "Reject inference methods in credit scoring," *J. Appl. Stat.*, vol. 48, no. 13–15, pp. 2734–2754, 2021, DOI: 10.1080/02664763.2021.1929090.

[19] W. Fengpei, S. Xiang, S. O. Young, and W. Zhiying, "Personal Credit Risk Evaluation Model of P2P Online Lending Based on AHP," *Symmetry (Basel).*, vol. 13, no. 1, p. 83, 2021, DOI: https://doi.org/10.3390/sym13010083.

[20] A. Filchenkov, N. Khanzhina, A. Tsai, and I. Smetannikov, "Regularization of autoencoders for bank client profiling based on financial transactions," *Risks*, vol. 9, no. 3, 2021, DOI: 10.3390/risks9030054.

[21] H. Gao, H. Liu, H. Ma, C. Ye, and M. Zhan, "Network-aware credit scoring system for telecom subscribers using machine learning and network analysis," *Asia Pacific J. Mark. Logist.*, vol. 34, no. 5, pp. 1010–1030, 2022, DOI: 10.1108/APJML-12-2020-0872.

[22] A. Gicić and A. Subasi, "Credit scoring for a microcredit data set using the synthetic minority oversampling technique and ensemble classifiers," *Expert Syst.*, vol. 36, no. 2, 2019, DOI: 10.1111/exsy.12363.

[23] W. Gui, L. Wang, H. Wu, X. Jian, D. Li, and N. Huang, "Multiple psychological characteristics predict housing mortgage loan behavior: A holistic model based on machine learning," *PsyCh J.*, vol. 11, no. 2, pp. 263–274, 2022, DOI: https://doi.org/10.1002/pchj.521.

[24] B. R. Gunnarsson, S. vanden Broucke, B. Baesens, M. Óskarsdóttir, and W. Lemahieu, "Deep learning for credit scoring: Do or don't?," *Eur. J. Oper. Res.*, vol. 295, no. 1, pp. 292–305, 2021, DOI: 10.1016/j.ejor.2021.03.006.

[25] M. Hanafy and R. Ming, "Classification of the Insureds Using Integrated Machine Learning Algorithms: A Comparative Study," *Appl. Artif. Intell.*, vol. 0, no. 0, pp. 1–32, 2022, DOI: 10.1080/08839514.2021.2020489.

[26] J. Hu, C. Chen, and K. Zhu, "Application of Data Mining Combined with Fuzzy Multicriteria Decision-Making in Credit Risk Assessment from Legal Service Companies," *Math. Probl. Eng.*, vol. 2021, 2021, DOI: 10.1155/2021/2499948.

[27] K. Iwai, M. Akiyoshi, and T. Hamagami, "Bayesian Network Oriented Transfer Learning Method for Credit Scoring Model," *IEEJ Trans. Electr. Electron. Eng.*, vol. 16, no. 9, pp. 1195–1202, 2021, DOI: https://doi.org/10.1002/tee.23417.

[28] A. Kim and S.-B. Cho, "An ensemble semi-supervised learning method for predicting defaults in social lending," *Eng. Appl. Artif. Intell.*, vol. 81, pp. 193–199, 2019, DOI: 10.1016/j.engappai.2019.02.014.

[29] A. Kumar, S. Sharma, and M. Mahdavi, "Machine Learning (ML) Technologies for Digital Credit Scoring in Rural Finance: A Literature Review," *Risks*, vol. 9, no. 11, p.

192, 2021, DOI: https://doi.org/10.3390/risks9110192.

[30] S. Lahmiri and S. Bekiros, "Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design," *Quant. Financ.*, vol. 19, no. 9, pp. 1569–1577, 2019, DOI: 10.1080/14697688.2019.1588468.

[31] J. W. Lee, W. K. Lee, and S. Y. Sohn, "Graph convolutional network-based credit default prediction utilizing three types of virtual distances among borrowers," *Expert Syst. Appl.*, vol. 168, 2021, DOI: 10.1016/j.eswa.2020.114411.

[32] H. Li, Y. Cao, S. Li, J. Zhao, and Y. Sun, "XGBoost Model and Its Application to Personal Credit Evaluation," *IEEE Intell. Syst.*, vol. 35, no. 3, pp. 52–61, May 2020, DOI: 10.1109/MIS.2020.2972533.

[33] X. Li and Y. Sun, "Application of RBF neural network optimal segmentation algorithm in credit rating," *Neural Comput. Appl.*, 2020, DOI: 10.1007/s00521-020-04958-9.

[34] L. Liu, "A Self-Learning BP Neural Network Assessment Algorithm for Credit Risk of Commercial Bank," *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, DOI: 10.1155/2022/9650934.

[35] S. Liu, R. Wang, and Y. Han, "Research on Personal Credit Evaluation Based on Machine Learning Algorithm," in *2021 6th International Symposium on Computer and Information Processing Technology (ISCIPT)*, 2021, pp. 48–52, DOI: 10.1109/ISCIPT53667.2021.00016.

[36] Y. Liu, A. Ghandar, and G. Theodoropoulos, "A Metaheuristic Strategy for Feature Selection Problems: Application to Credit Risk Evaluation in Emerging Markets," in *2019 IEEE Conference on Computational Intelligence for Financial Engineering Economics (CIFEr)*, May 2019, pp. 1–7, DOI: 10.1109/CIFEr.2019.8759117.

[37] V. Moscato, A. Picariello, and G. Sperli, "A benchmark of machine learning approaches for credit score prediction," *Expert Syst. Appl.*, vol. 165, Mar. 2021, DOI: 10.1016/j.eswa.2020.113986.

[38] I. Nica, D. B. Alexandru, S. L. P. Crăciunescu, and Ștefan Ionescu, "Automated Valuation Modelling: Analysing Mortgage Behavioural Life Profile Models Using Machine Learning Techniques," *Sustainability*, vol. 13, no. 9, p. 5162, 2021, DOI: https://doi.org/10.3390/su13095162.

[39] A. Pérez-Martín, A. Pérez-Torregrosa, A. Rabasa, and M. Vaca, "Feature selection to optimize credit banking risk evaluation decisions for the example of home equity loans," *Mathematics*, vol. 8, no. 11, pp. 1–16, 2020, DOI: 10.3390/math8111971.

[40] J. Qi, R. Yang, and P. Wang, "Application of explainable machine learning based on Catboost in credit scoring," *J. Phys. Conf. Ser.*, vol. 1955, no. 1, Jun. 2021, DOI: https://doi.org/10.1088/1742-6596/1955/1/012039.

[41] C. Qin, Y. Zhang, F. Bao, C. Zhang, P. Liu, and P. Liu, "XGBoost optimized by adaptive particle swarm optimization for credit scoring," *Math. Probl. Eng.*, vol. 2021, 2021, DOI: 10.1155/2021/6655510.

[42] C. Rao, M. Liu, M. Goh, and J. Wen, "2-stage modified random forest model for credit risk assessment of P2P network lending to 'Three Rurals' borrowers," *Appl. Soft Comput. J.*, vol. 95, 2020, DOI: 10.1016/j.asoc.2020.106570.

[43] B. Ruyu, H. Mo, and L. Haifeng, "A Comparison of Credit Rating Classification Models Based on Spark- Evidence from Lending-club," *Procedia Comput. Sci.*, vol. 162, pp. 811–818, 2019, DOI: https://doi.org/10.1016/j.procs.2019.12.054.

[44] F. Shen, X. Zhao, Z. Li, K. Li, and Z. Meng, "A novel ensemble classification model based on neural networks and a classifier optimisation technique for imbalanced credit risk evaluation," *Phys. A Stat. Mech. its Appl.*, vol. 526, 2019, DOI: 10.1016/j.physa.2019.121073.

[45] M. Soui, I. Gasmi, S. Smiti, and K. Ghédira, "Rule-based credit risk assessment model using multi-objective evolutionary algorithms," *Expert Syst. Appl.*, vol. 126, pp. 144–157, 2019, DOI: 10.1016/j.eswa.2019.01.078.

[46] S. Sreesouthry, A. Ayubkhan, M. M. Rizwan, D. Lokesh, and K. P. Raj, "Loan Prediction Using Logistic Regression in Machine Learning," *Ann. Rom. Soc. Cell Biol.*, vol. 25, no. 4, pp. 2790–2794, 2021, [Online]. Available: https://www.proquest.com/scholarly-journals/loan-prediction-using-logistic-regression-machine/docview/2582236651/se-2?accountid=37408.

[47] G. Teles, J. J. P. C. Rodrigues, K. Saleem, and S. A. Kozlov, "Classification Methods Applied to Credit Scoring with Collateral," *IEEE Syst. J.*, vol. 14, no. 3, pp. 4557–4566, 2020, DOI: 10.1109/JSYST.2019.2937552.

[48] G. Teles, J. J. P. C. Rodrigues, R. A. L. Rabêlo, and S. A. Kozlov, "Comparative study of support vector machines and random forests machine learning algorithms on credit operation," *Softw. Pract. Exp.*, vol. 51, no. 12, pp. 2492–2500, 2021, DOI: https://doi.org/10.1002/spe.2842.

[49] L. Ting, F. Qinwei, K. Qian, and N. Lei, "Extreme Learning Machine Based on Firefly Adaptive Flower Pollination Algorithm Optimization," *Processes*, vol. 8, no. 12, p. 1583, 2020, DOI: https://doi.org/10.3390/pr8121583.

[50] D. Tripathi, D. R. Edla, V. Kuppili, and A. Bablani, "Evolutionary Extreme Learning Machine with novel activation function for credit scoring," *Eng. Appl. Artif. Intell.*, vol. 96, 2020, DOI: 10.1016/j.engappai.2020.103980.

[51] de C. S. P. Vitor and T. L. C. Bambirra, "Extreme Wavelet Fast Learning Machine for Evaluation of the Default Profile on Financial Transactions," *Comput. Econ.*, vol. 57, no. 4, pp. 1263–1285, Apr. 2021, DOI: https://doi.org/10.1007/s10614-020-10018-0.

[52] C. Wang, Z. Hu, R. Chiong, S. Dhakal, Y. Chen, and Y. Bao, "A PSO-Based Ensemble Model for Peer-to-Peer Credit Scoring," in *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, 2018, pp. 412–418, DOI: 10.1109/FSKD.2018.8687154.

[53] D. Wang, Z. Wu, and B. Zhu, "Controlling Shareholder Characteristics and Corporate Debt Default Risk: Evidence Based on Machine Learning," *Emerg. Mark. Financ. Trade*, vol. 0, no. 0, pp. 1–16, 2022, DOI: 10.1080/1540496X.2022.2037416.

[54] H. Wang, G. Kou, and Y. Peng, "Multi-class misclassification cost matrix for credit ratings in peer-to-peer lending," *J. Oper. Res. Soc.*, vol. 72, no. 4, pp. 923–934, 2021, DOI: 10.1080/01605682.2019.1705193.

[55] Z. Wang, C. Jiang, H. Zhao, and Y. Ding, "Mining Semantic Soft Factors for Credit Risk Evaluation in Peer-to-Peer Lending," *J. Manag. Inf. Syst.*, vol. 37, no. 1, pp. 282–308, 2020, DOI: 10.1080/07421222.2019.1705513.

[56] C.-H. Weng and C.-K. Huang, "A Hybrid Machine Learning Model for Credit Approval," *Appl. Artif. Intell.*, vol. 35, no. 15, pp. 1439–1465, 2021, DOI: 10.1080/08839514.2021.1982475.

[57] Y. Wu and Y. Pan, "Application Analysis of Credit Scoring of Financial Institutions Based on Machine Learning Model," *Complexity*, vol. 2021, 2021, DOI: https://doi.org/10.1155/2021/9222617.

[58] Y. Xia, L. He, Y. Li, Y. Fu, and Y. Xu, "A dynamic credit scoring model based on survival gradient boosting decision tree approach," *Technol. Econ. Dev. Econ.*, vol. 27, no. 1, pp. 96–119, 2021, DOI: 10.3846/tede.2020.13997.

[59] D. Xu, X. Zhang, J. Hu, and J. Chen, "A Novel Ensemble Credit Scoring Model Based on Extreme Learning Machine and Generalized Fuzzy Soft Sets," *Math. Probl. Eng.*, vol. 2020, 2020, DOI: 10.1155/2020/7504764.

[60] J. Xu, L. Zekai, and Y. Xie, "Loan default prediction of Chinese P2P market: a machine learning methodology," *Sci. Reports (Nature Publ. Group)*, vol. 11, no. 1, 2021, DOI: https://doi.org/10.1038/s41598-021-98361-6.

[61] Y.-Z. Xu, J.-L. Zhang, Y. Hua, and L.-Y. Wang, "Dynamic credit risk evaluation method for e-commerce sellers based on a hybrid artificial intelligence model," *Sustain.*, vol. 11, no. 19, 2019, DOI: 10.3390/su11195521.

[62] G. Yang, "Research on Financial Credit Evaluation and Early Warning System of Internet of Things Driven by Computer-Aided Technology," *Comput. Aided. Des. Appl.*, vol. 19, no. S6, pp. 158–169, 2022, DOI: 10.14733/cadaps.2022.S6.158-169.

[63] J. You, G. Li, and H. Wang, "Credit Grade Prediction Based on Decision Tree Model," in *2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, Nov. 2021, pp. 668–673, DOI: 10.1109/ISKE54062.2021.9755326.

[64] X. Yu, Q. Yang, R. Wang, R. Fang, and M. Deng, "Data Cleaning for Personal Credit Scoring by Utilizing Social Media Data: An Empirical Study," *IEEE Intell. Syst.*, vol. 35, no. 2, pp. 7–15, 2020, DOI: 10.1109/MIS.2020.2972214.

[65] Y. Yuan, X. Gong, M. Guo, Y. Han, and H. Li, "Research on Personal Credit Evaluation of Commercial Banks Under Ensemble Learning Framework," in *2020 2nd International Conference on Applied Machine Learning (ICAML)*, Oct. 2020, pp. 29–38, DOI: 10.1109/ICAML51583.2020.00015.

[66] Z. Yuhao, "Research on personal credit evaluation of internet finance based on blockchain and decision tree algorithm," *EURASIP J. Wirel. Commun. Netw.*, vol. 2020, no. 1, Dec. 2020, DOI: https://doi.org/10.1186/s13638-020-01819-w.

[67] L. Zhang and J. Fan, "Formal Credit-Assisted New Agricultural Business A Multifactor

Analysis Based on BP Neural Network.," *Mob. Inf. Syst.*, pp. 1–11, 2022, [Online]. Available: httpssearch.ebscohost.comlogin.aspxdirect=true&db=iih&AN=156288423&lang=es&site=eds-live.

[68] S. Zhang, Y. Hu, and Z. Tan, "Research on borrower's credit classification of P2P network loan based on LightGBM algorithm," *Int. J. Embed. Syst.*, vol. 11, no. 5, pp. 602–612, 2019, DOI: 10.1504/IJES.2019.102435.

[69] T. ZHANG, W. ZHANG, W. XU, and H. HAO, "Multiple instance learning for credit risk assessment with transaction data," *Knowledge-Based Syst.*, vol. 161, pp. 65–77, 2018, DOI: 10.1016/j.knosys.2018.07.030.

[70] Y. Zhao, "Research on Personal Credit Evaluation of Internet Finance Based on Blockchain and Decision Tree Algorithm," *EURASIP Journal on Wireless Communications and Networking*. Research Square, Durham, Aug. 13, 2020, DOI: https://doi.org/10.21203/rs.3.rs-34146/v2.

[71] X. Zhou, W. Zhang, and Y. Jiang, "Personal credit default prediction model based on convolution neural network," *Math. Probl. Eng.*, vol. 2020, 2020, DOI: 10.1155/2020/5608392.

[72] J. Zhu et al., "A Deep Learning-Oriented Personal Credit Model for Online Learning Platform," in *2020 15th International Conference on Computer Science Education (ICCSE)*, 2020, pp. 838–843, DOI: 10.1109/ICCSE49874.2020.9202397.

[73] P. Ziemba, J. Becker, A. Becker, A. Radomska-Zalas, M. Pawluk, and D. Wierzba, "Credit Decision Support Based on Real Set of Cash Loans Using Integrated Machine Learning Algorithms," *Electronics*, vol. 10, no. 17, p. 2099, 2021, DOI: https://doi.org/10.3390/electronics10172099.