Journal of Theoretical and Applied Information Technology
15th August 2024. Vol.102. No. 15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PREDICTION MODEL APPLYING MACHINE LEARNING TO FORECAST THE BANKRUPTCY OF COMPANIES. A SYSTEMATIC REVIEW OF THE LITERATURE.

DINO QUINTEROS-NAVARRO[1] , CIRO RODRIGUEZ[2]

[1]Universidad Nacional Mayor de San Marcos, Lima, Peru

[2]Universidad Nacional Mayor de San Marcos, Lima, Peru

E-mail: [1]dino.quinteros@unmsm.edu.pe, [2]crodriguezro@unmsm.edu.pe

ABSTRACT

Prediction models that are aimed at companies allow trends to be identified and generate a much broader picture, making decisions more effective and efficient. In that sense, the research study uses a literature review to identify the state of the art of how predictive models can interact and generate more accurate and reliable results to identify, forecast and eventually reduce the bankruptcy of companies. In addition, the systematic analysis of the contributions was carried out where the main models of supervised learning and other techniques were considered. The authors [25] specify that a systematic literature review is a primary tool for developing an evidence base by identifying, evaluating, and interpreting all available research relevant to a particular research question, thematic area, or phenomenon of interest. Likewise, questions were posed through three (3) stages: identification of parameters, calculation of the ACCP and determination of factors. For the calculation of the weighted precision metric (ACCP), the most significant values of each model were used and the results obtained were analyzed. In this sense, the review must contain the following: method of analysis, theoretical basis, classification of payment yield prediction models and analysis of the topics according to the questions asked.

Keywords: *Bankruptcy, Predictive, Models, Machine Learning, Making Decisions*

1. INTRODUCTION

In the last decade, companies have been using tools and business strategies with emerging technology that has allowed consolidation in a more constant and forceful way in an aggressive business environment [1]. In this context, it is specified that one of the emerging strategies is the implementation of applications generated with artificial intelligence, where proposals and initiatives are generated that include various metrics such as efficiency and optimization of a company's critical processes. Likewise, it has been identified that among the main proposals are subsistence activities to avoid bankruptcy. The authors [4] consider that the results required by a company must be aligned with institutional goals and strategic plans.

The authors [1] consider that a very recurrent problem but never marked or worked on is financial forecasting. Most of the most vulnerable companies (micro, small and medium-sized companies) are interested in venturing into initiatives to forecast financial statements that allow them to adapt to the current financial and commercial environment to avoid bankruptcy. They also consider that due to the business sector and nature of the orientation of the business bankruptcy problem, it is necessary to identify the best model but that it has grouping techniques to avoid problems of data inconsistency. Similarly, the authors [37] consider that the diagnosis of business crisis models on the risk of bankruptcy through financial analysis is essential to protect financial difficulties. They also specify that the construction of a stable and accurate model to predict the bankruptcy of a banking company is very important as a basis for a successful and effective result.

Companies face significant challenges due to severe financial difficulties, incomplete financial information, and capital constraints, all of which lead to experimenting with business bankruptcy activities [7]. They are also often excluded from traditional channels of financing for both investment and livelihood activities. In this sense, financial challenges have a negative influence on the health of companies and the sustainability of the development of these companies [35].

The authors [36] indicate that the model for identifying the risk of corporate default is important for the management of loan risks, credit scoring, bankruptcy risk, etc. Similarly, the authors [37] indicate that several studies require effective methods to build an efficient system to identify business bankruptcy, based on various modeling techniques. Machine learning methods offer an automatic and objective way to achieve a high level of prediction. In addition, several industrial sectors, including banking and business, may face the risk of bankruptcy triggered by financial difficulties. This risk forces each company to survive the various shocks that arise.

Various studies require models of credit rating, payment performance and business bankruptcy, which are very necessary and useful. In this sense, machine learning models provide techniques and determinants of aspects through specific patterns that allow the processing of information in order to generate predictions according to the selection criteria [23].

The authors [1] indicate that bankruptcy has a very relevant and transcendent effect on companies as a whole, both for the operational staff and for the strategic command. Similarly, the authors [37] mention that the sustainability of a company is possible with the implementation of performance indicators, which need to be of a financial nature and that constitute relevant factors in the internal and external environment. Financial indicators are essential to have a complete picture of the activities and processes carried out by the company. It is important that performance indicators are developed from the company's annual financial report. They also mention that proposals and strategies must be developed to predict extreme situations and mitigate the difficulties that arise through the implementation of indicators to reach the annual goals according to the company's strategic plan.

The bankruptcy of companies is a structural and legal situation that is generated from the failure to comply with payment obligations to creditors [1], which is why business activities are compromised to the point of the impossibility of generating continuity. Institutions do not achieve an adequate prediction to obtain information on the companies that have the highest incidence of corporate bankruptcies [4].

The authors [34] indicate that business bankruptcy has very significant consequences throughout the environment, whether internal or external. Complications are considered to be generated from the operational levels with instability and pressure. However, models with traditional algorithms were identified where the prediction is very basic and that does not contribute anything interesting to the bankruptcy forecast that is intended to be made, so these models do not represent something new or transcendental because they focus on forecasting the events that can be triggered at a social and economic level. Therefore, the study conducted by the authors [34] considered the integration of various perspectives and variables in the machine learning (ML) modeling process, which allowed to consider the different costs and expenses caused by subsistence events prior to bankruptcy. However, they did not consider the social impact that a serious financial problem such as bankruptcy will present.

The authors [16] consider insolvency to be a crucial problem for many companies because they have no way to identify or predict this problem. Therefore, this problem has direct or indirect effects on both the people who work in the financial business and the population. In addition, the ability to predict insolvency is in high demand in companies and there are several proposed efforts to predict insolvency using numerical methods and computer applications. Each country has its own data patterns for contextual reasons. Therefore, businesses can have different data patterns. Consequently, the chosen predictive model must be adapted to the dataset according to the business context of the country or region. In that sense, a company's bankruptcy can occur due to capital inadequate to withstand unforeseen losses, or due to insured incidents or fluctuations in the assets it owns. The activities of some companies may increase the risk of problems or losses and they are unable to regain their position, which is known as financial hardship or financial insolvency problems. Consequently, potential insolvencies should be identified as early as possible to identify actual insolvency.

Regarding insolvency, the authors [34] indicate that it is a recurring problem and a critical threat within a company where the size is not important. Insolvency limits the company's ability to develop and requires immediate support. When a company declares insolvency, it will influence the workers and put the interests of the shareholders at risk. Likewise, insolvency is a constant and recurring concern for the business sector, especially for small and micro enterprises. Therefore, it is important to implement indicators that provide a comprehensive

overview of the way forward and the difficulties that may arise in order not to fall into insolvency.

The authors [34] specify that business bankruptcy is an event that negatively influences all stakeholders such as customers, employees, suppliers, operational managers and strategic staff. In recent years, the application of various machine learning models and techniques has increased, in recent years, they became the standard for predicting bankruptcies, since statistical models such as logistic regressions and discriminant analysis have some disadvantages. Likewise, with the implementation of machine learning models, more accurate forecasts can be generated, but it is very necessary to incorporate variables that are adapted to the context of the company. Machine learning is a very useful tool for a company and is decisive when it comes to generating more accurate predictions with simple and complex variables.

Therefore, the authors [1] indicate that supervised learning and deep learning have gained the interest of researchers to analyze the financial field. However, data related to the financial status of companies is unbalanced.

The practical justification of the systematic literature review article refers to collecting articles from various sources that have relevant aspects on supervised learning and business bankruptcy, which has great importance for achieving the objectives of the review because it allows. It incorporates concepts, techniques and tools essential to answer the research questions. The methodological justification refers to specifying methods and techniques for the treatment of the articles studied with topics that generate new research approaches and future work.

This article specifies the study of prediction models using machine learning to forecast the bankruptcy of companies, which was developed through the method of analysis, classification of bankruptcy prediction models, analysis of topics based on the questions asked, discussion of results and conclusions.

2. METHOD OF ANALYSIS

2.1 Methodology

The methodology of this systematic literature review was carried out according to what was indicated by the authors [31] who specified the necessary criteria for the development of reviews and meta-analyses (PRISMA). A systematic literature review identifies, analyzes and interprets available evidence related to the research. Likewise, the process defined an eligibility criterion that identifies the objectives of the review, the inclusion and exclusion criteria applied. Good practices are proposals for the structured presentation of search, selection, analysis and synthesis activities of the various studies and bibliographic references compiled in a systematic literature review. The methodology specified with PRISMA allows bibliographic sources to be organized in a more efficient and effective way; motivating a better organization for the analysis of each article that is related to the research questions, which need to be developed. It is important to specify that PRISMA is essential in the investigation to guarantee the transparency of the documented sources and the subsequent conclusive analysis.

This section also identifies the databases used, the search strategy, the selection steps and the study characteristics shown by the articles selected in the review.

2.2 Research questions

To carry out the literature review on factors, prediction and explainability of prediction behavior based on machine learning, three research questions were proposed:

Q1: What are the prediction models that apply machine learning to forecast the bankruptcy of companies?

Q2: What is the ideal prediction model that machine learning uses to forecast business failure?

Q3: What are the factors that influence the prediction model that machine learning applies to predict the bankruptcy of companies?

2.3 Inclusion and exclusion criteria

The results obtained from the search, in accordance with the proposed strategy, were submitted to a selection process considering the established inclusion and exclusion criteria. According to the criteria indicated in Table 1.

Table 1: Inclusion and exclusion criteria

Inclusion criteria	Exclusion Criteria
CI1: Primary articles related to the area and topic	CE1: Bibliographic review articles, books, editorials, letters, errata, proceedings, posters.
CI2: Articles in English	CE2: Articles not related to the thematic area
CI3: Publication date from 2018 to 2023	CE3: Documents outside the established date

	range
IC4: Full-text and open access articles	CE4: Only paid or incomplete summaries
CI5: Articles that answer research questions	CE5: Articles that are not related to the research questions

2.4 Development of the review

To search for relevant information, articles published in journals indexed in Scopus and Scince Direct, in the period between January 2018 and June 2023, were reviewed in "title-abstract-key words", using the string:

TITLE-ABS-KEY(machine and learning and predictive and bankruptcy) AND (LIMIT-TO (OA,"all") AND (LIMIT-TO (SUBJAREA,"COMP")) AND (LIMIT-TO (DOCTYPE,"ar"))))

The following results were obtained in the search process:

Scopus: Initial result: 132,970 articles and Science Direct: 1,980 articles. The inclusion and exclusion criteria are then applied and obtained. Scopus: Final result: 22 articles and Science Direct: 25 articles, as illustrated in Figure 1.

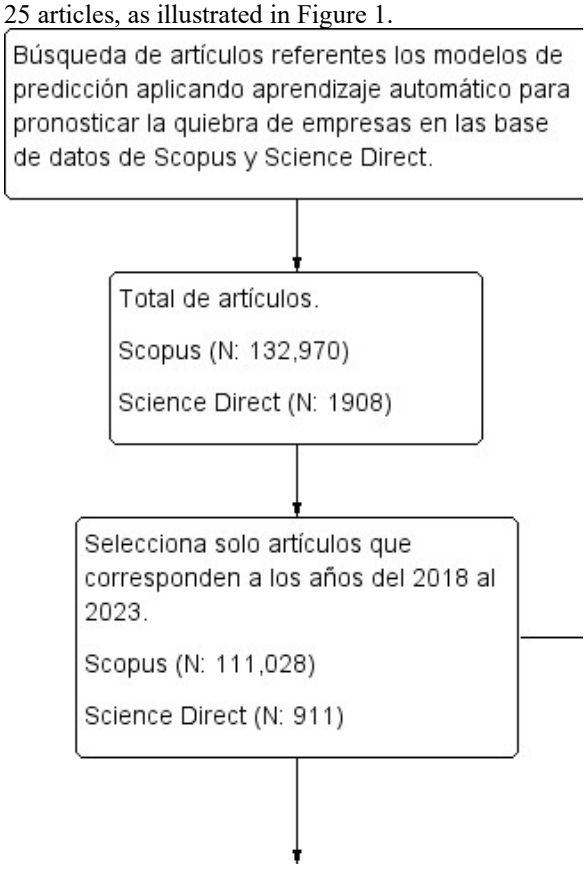

Figure 1: Prisma Flowchart

2.5 Dataset and approaches of the reviewed articles

The datasets mainly used by the reviewed articles correspond to data on small and medium-sized enterprises (MSEs) in Peru. In addition, there is some research that deals with medium-sized companies that are the product of adjacent franchises or conglomerates. Deep learning needs to be optimal for MSEs. The main objective of the technique is to determine the financial status of MSEs, which contains the design of the feature selection based on the convolution neural network and optimization algorithm (CNN-LSTM) that is used for data classification [27]. Accuracy plays an important role in defining the profitability and productivity of financial companies.

There is also a need to discuss the principles and application of general machine learning approaches. Traditional machine learning models for bankruptcy risk contain supervised learning (SVM) algorithms [24], K-Nearest Neighbor (KNN) [39], Random Forest (RF) [18], decision trees (DT) [9], AdaBoost (SGB) [23], Extreme Gradient Boost (XGBoost) [28], and LightGBM algorithm [12]. Neural network models usually belong to deep learning methods, where most of them include neural networks (ANNs) [7] [3]. In addition, big data models applied in the forecasting and prediction of yield [14] [29] [12] are presented.

3. CLASSIFICATION OF BANKRUPTCY PREDICTION MODELS

It is necessary to identify companies that use market dynamics and variability information to estimate investment probabilities. In addition, data sources are oriented towards bankruptcy probability models according to variables in the financial statements [6]. For example: stock market information, particularities of the sector in which the activity is carried out, regulation, ease of access to credit, etc.

Various studies have been carried out to forecast the performance of companies in terms of payment, investments or risk situations. Likewise, it was identified that a large percentage of these studies are generated from personalized, qualitative and criterion judgments, which consider the treatment of financial information, performance and management of the company. Similarly, it is specified that the study mentioned by the authors [20] considers that data processing is very important to generate advanced reports, but they require very high consistency. The authors have

also begun to use machine learning techniques to predict yields. For example, prediction models based on deep neural networks are once again attracting attention and are now widely used in the construction of prediction and classification models [20].

Among the studies that predict a company's corporate performance, there are few that build a prediction model using patent data and a deep learning algorithm. In this study, we propose a corporate performance prediction model based on deep neural networks that uses a company's financial and technical indicators as predictors [20].

An improved algorithm was identified to predict exchange rate movements, where a conjugate gradient method was applied in the training process [20]. In addition, in this study, a comparative analysis was performed to empirically verify the prediction performance of the proposed model using the same exchange rate time series data. Their results showed that the proposed model outperforms a neural network-based model [3].

The model, which uses prediction algorithms to identify defaults by French companies, allows the model's prediction performance to be compared with that of a prediction model based on supporting vectors [39]. The results showed that the performance of the prediction-based prediction model outperformed that of the supportive vector-based prediction model, [20].

The authors [12] specify that the LightGBM algorithm uses multidimensional data that has significant consistency. Likewise, they identified results of the precision metric with a value of 80.25%, which allows inferring that the processed data of the users is greater than 80%, it is considered that the consistency of the data must be high to obtain greater values of accuracy. The model identified with the LightGBM algorithm has the best prediction performance for business performance evaluations. It is important to note that the integration of the LightGBM algorithm and the linear combination with consistent values provides an accuracy value of 91.36%, and an accuracy of 84.36%. The model proposed with the LightGBM algorithm can improve the performance forecast and identify reliable and efficient predictions of activities that could cause perjury to the company in the future. This streamlined model differs from traditional machine learning models such as SVM algorithms or logistic regression.

4. ANALYSIS OF THE TOPICS ACCORDING TO THE QUESTIONS ASKED

4.1 Analysis of the questions asked

The topics presented in sections 2 and 3 are directly related to the questions asked in section 2 (2.2). In this sense, the weighted precision calculation model [20] was applied, which allowed answering the questions formulated through three (3) stages: identification of parameters, calculation of the ACCP and determination of factors. In addition, the following analysis was obtained:

Q1: What are the prediction models that apply machine learning to forecast the bankruptcy of companies?

Section 3 specifies the models that have obtained the greatest relevance at the time of application and obtaining results, where precision is highlighted as the indicator that gives greater confidence to the prediction. Likewise, the variables that are handled depend on the scope and the target companies. The following Table 2 specifies the parameters mentioned directly or indirectly in different studies related to the prediction of company bankruptcy. Likewise, the largest number of articles that mention business bankruptcy, credit risk, models, algorithms, machine learning, companies, was considered.

Table 2: PARAMETERS USED BY DIFFERENT JOBS

SOURCES	NRO	PARAMETERS	PURPOSES
[29]	1	Data mining	Data processing using data mining techniques
[7]	4	KNN Algorithm, Bankruptcy	Neural networks, prediction evaluations using a model
[16]	1	Alternative models	Alternative methods
[1]	1	Machine Learning Model	Machine Learning Model with Deep Learning Problems
[9]	4	Prediction model	Identification of prediction models
[41]	2	Prediction model	Identification of prediction

			models
[37]	4	Machine Learning Model	Identification of optimization techniques using proprietary models
[28]	5	Machine Learning Model	Data processing using the XGBoost algorithm
[15]	2	Data mining	Comparative analysis with prognostic methods
[32]	2	Machine Learning Model	Combining Variables in Prediction Models
[17]	2	Prediction model	Identification of prediction models
[23]	4	Prediction model	Identification of prediction models
[38]	5	Prediction model	Identification of prediction models
[34]	1	Machine Learning Model	Benchmarking using machine learning models
[8]	1	Machine Learning Model	Combining Variables in Prediction Models
[10]	1	Prediction model	Proposal of prediction models for application with SMEs
[33]	4	Machine Learning Model	Machine learning-based model to predict bankruptcies
[26]	1	Machine Learning Model	Data mining and machine learning proposals
[36]	1	Machine Learning	Data mining and machine
		Model	learning proposals
[3]	4	KNN Algorithm, Bankruptcy	Combining Variables in Prediction Models
[18]	4	Machine Learning Model	Proposal of selection methods to predict performance
[14]	4	Big Data	Prediction and Performance Models

It is observed that in Figure 2 the parameters can signify relationships and similarities with the treatment of the models that are proposed, all of which means that the scientific articles reviewed correspond not only to the context of predictability but also to the risk of bankruptcy and to the variables that can enrich each model studied. With this, it is specified that the parameters represent a direct relationship with the models proposed to predict the bankruptcy of companies.

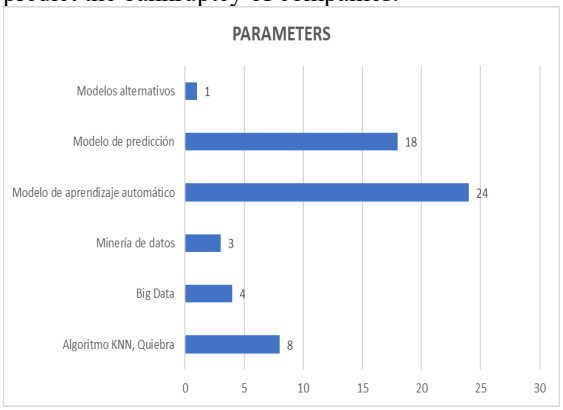

Figure 2: Parameters used

Q2: What is the ideal prediction model that machine learning uses to forecast business failure?

This question is aimed at knowing the ideal prediction model to know the bankruptcy forecast of companies, which is achieved by calculating the ACC formula (Precision), which corresponds to what is obtained in each model proposed in this systematic review, in this sense Table 3 is available.

4.2 Evaluation Metrics

For the evaluation, the metric proposed by the authors [20] was used, which determined the weighted precision metric ACCP (1) as the main study metric, which considers the evaluation of the

performance of the model and is suitable for bibliographic reviews. In this sense, accuracy is calculated through the classified values (precision) of the model divided by the number of models applied. That is, the average of the sum of the precision value of each model and the number of models applied.

$$ACCP = \frac{\sum ACC\ Model}{\sum models} \quad (1)$$

The ACCP metric [20] identifies the best (highest) values of each model and provides an important overview towards the proposed goal. Overall, it is an ideal indicator for optimizing measurements and mapping out quantifiable products according to the context.

4.3 Classification of techniques

The method applied to obtain the ACCP metric [20] is mainly discussed in literature reviews where it is necessary to configure an ideal model according to the similarity of the model studied. A similar result is also obtained from models with their own parameters and variables.

Table 3 shows the eight most representative and relevant models of the literature review, where the accuracy of each model is indicated.

Table 3: MODEL ACCURACY FOR PREDICTION

SOURCES	MODEL	ACC (Precision)
[7] [3]	KNN Algorithm, Bankruptcy	97.80%
[9] [21] [22]	Prediction model: hybrid ML algorithm	86.00%
[28]	XGBoost algorithm	91.70%
[23] [32] [38]	Prediction model	89.05%
[33]	DSS Model	93.94%
[18] [24]	Clustering algorithm	87.10%
[14] [29]	Big Data Model	93.00%

In this sense, Figure 3 shows that the model with the ANN algorithm referring to business bankruptcy is the model with the highest accuracy (97.80%), which was achieved through the proposed calculation of ACCP.

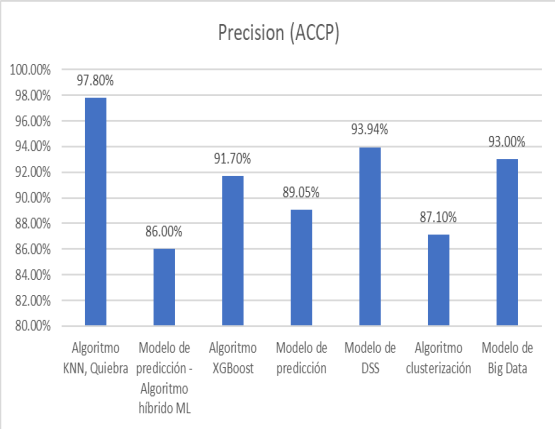

Figure 3: Accuracy (ACCP) of prediction models)

Q3: What are the factors that influence the prediction model that machine learning applies to predict the bankruptcy of companies?

As for the factors that have the greatest predominance or influence, it is specified that they are generated from the exhaustive review of each article, where several factors that are relevant for the calculation of the accuracy of the model were identified. In this sense, the process to determine the predominant or relevant factor was carried out by adding the number of times the factor was used as a calculation variable in each research question for the model of the reviewed article. Table 4 shows the number of references to the various factors presented in the articles reviewed and studied. It is also necessary to identify factors for further treatment in prediction models.

Table 4: FACTOR IDENTIFICATION

FACTORS	NRO	REFERENCES
Growth rate	16	[7] [9] [28] [3] [18]
Precision	13	[7] [9] [28] [33] [38]
Years of operation	6	[16] [1] [17]
Seed capital	8	[7] [23] [3] [37]
Personnel	8	[3] [18] [14] [32] [21]
Category	12	[3] [33] [8] [15]
Member of the Chamber of Commerce	12	[7] [9] [28] [29]
Entry into the foreign market	6	[10] [34] [8] [36]

Figure 4 also shows the factors that have the greatest impact on the articles reviewed and studied. These are identified as relevant factors and variables to be studied in the proposals for

optimized models for the prediction of company bankruptcies.

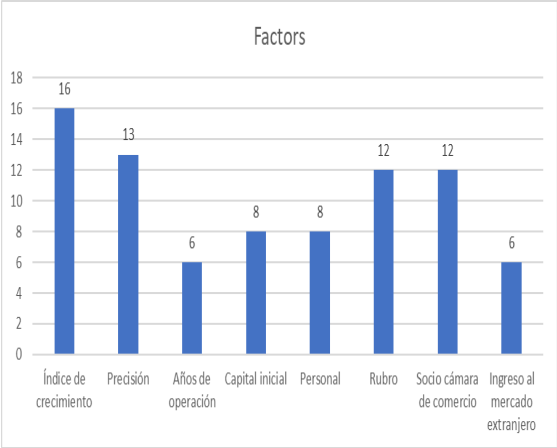

Figure 4: Factor Identification

5. DISCUSSION OF THE RESULTS

In order to identify the most relevant and determining models, it was carried out through the treatment and analysis of each model used in the respective article. In this sense, it was identified that the models that have the greatest relevance to answer the first research question are: machine learning model, prediction model and ANN algorithm. The models identified were oriented to data mining, big data, KNN algorithms, simple, complex and alternative supervised learning models, which were found to be associated with predictive models. Likewise, the authors of the reviewed articles specify that the models need to be treated with data aimed at a specific and personalized context; that is, that they incorporate suitable and specific variables. The authors [29] consider the data mining model for the prediction treatment with companies and where ACCP metrics of 93% were obtained; Similarly, the authors [14] indicate that the variables used needed a large volume of data with big data models to achieve an adequate result. In this context, the authors [33] used DSS models for decision-making; however, the prediction had inconsistent variables. In this sense, the present systematic literature review identified that KNN models are the most suitable to predict the bankruptcy of companies with an ACCP of 97% when considering the incorporation of precise and personalized variables.

It was identified that the authors [5] in their review: "Machine learning classification techniques to detect the Impact of Outcomes on Commercial Bank Performance", specified supervised learning techniques that allowed predicting performance with the bank, which considered an accuracy of 93 %; Likewise, the authors [11] in their study: "A Survey on Machine Learning and Statistical Techniques in Predicting Bankruptcy" considered techniques to predict business bankruptcy, of which 4 techniques were identified with results between 90% and 93.5% where the average of 91.6%. The authors [33] in their study: "The evaluation of bankruptcy prediction models based on socio-economic costs" considered prediction models with data on socio-economic sequences of companies up to 2 years of initiated operations. Where the results of the metrics used generated an average of 95.4%. This indicates that the aforementioned studies have a similar procedure for calculating the average precision. In this sense, it is stated that this procedure is important and relevant to know techniques and models suitable for predicting business bankruptcy. Similarly, the authors [10] in their study: "Rethinking SME default prediction: a systematic review of the literature and future perspectives" considered the evaluation of various models for prediction. In the aforementioned study, variables with different contexts were considered and not all of them responded to the eligible criteria for predicting business bankruptcy. In this sense, in this review, suitable factors were presented to be considered as study variables, which would allow generating correlations between relevant variables and establishing a methodology with clear and precise procedures according to the context of the company. In addition to this, a procedure was considered for calculating the variables that allowed generating precise results to answer the questions posed.

To determine the ideal model, the precision formula proposed in this review article was used. In this sense, the ideal prediction model applying machine learning was the ANN algorithm (referring to bankruptcy), which is evidenced in 3 articles related to company bankruptcy, where an ACCP of 97.80% was obtained. Likewise, data greater than 90% were obtained with XGBoost algorithms, DSS model, Big Data models and other supervised learning prediction models.

To identify the most relevant factors that generate the greatest influence on the prediction model, they were determined by the number of times the factor was used as a calculation variable in each research question. In this sense, it was identified that the growth rate (CI) is the most relevant and predominant factor in prediction models that use machine learning, which were identified in the articles [7] [9] [28] [3] [18]. The authors specified that this factor significantly influences not only decision-making but also

present and future strategies that could guide the incursion of investments or proposals for improvement both for their employees and for external events, the construction of control boards and the implementation of management and performance indicators is essential. However, it was identified that the predominant factor has variables that support the calculation treatment such as the initial capital, item, credit valuation or if it is a partner in a business conglomerate. Likewise, the articles that refer to the growth index (CI) have various variables depending on the size of the company under study and the market segment. After that, there are various factors such as: Accuracy, years of operation, initial capital, personnel, industry, member of the chamber of commerce and entry into foreign trade; which must be treated according to the business context and the way of processing the data. Therefore, it is important to consider variables through detailed studies that lead to improvements in all the company's processes and activities.

6. CONCLUSIONS

Businesses require access to credit to achieve their business goals and financial success. In this sense, it is important to obtain access to financial services and an adequate credit evaluation can facilitate the initiation of financing and investments.

Recent advances in artificial intelligence emerged from a new wave of machine learning-powered credit risk models that led to forecasts with near-100% accuracy.

Questions were posed through three (3) stages: identification of parameters, calculation of the ACCP and determination of factors. For the calculation of the weighted precision metric (ACCP) as the main study metric, which considers the evaluation of the performance of the model and is suitable for bibliographic reviews. It identifies the best (highest) values of each model and provides an ideal indicator to optimize measures that allow quantifiable products to be drawn according to the context.

The scientific articles reviewed correspond not only to the context of credit predictability but also to the risk of bankruptcy and the variables that enrich each model. With this, it is specified that the parameters represent a direct relationship with the models proposed to predict the bankruptcy of companies. It is specified that simple, complex, and alternative supervised learning models represent the largest number of references on predictability studies to forecast the bankruptcy of companies of various sectors, segments, and sizes.

The eight most representative and relevant models of the literature review were presented, indicating the accuracy of each model. In this sense, the model enhanced with deep learning is the model that has the highest accuracy (97.80%). It is considered that the optimization models are based on generic supervised learning models, so it is important to incorporate internal and external variables through exhaustive and detailed studies, with this it will be possible to obtain greater and better metrics.

It was identified that the factors for better and greater accuracy respond to the company's own and personalized data according to its context, where the predominant factor was the growth index (CI). Likewise, the predominant factor is essential to know the orientation and application of strategies, in addition to anticipating market situations.

The strength of this review is a solid base of studies collected to answer the research questions. Likewise, the development of a formula to calculate the average precision of the prediction models was proposed. With this, it was possible to obtain metrics and also factors suitable for incorporating relevant variables on company bankruptcy. Regarding limitations, it was possible to identify that it is necessary to incorporate articles and sources of information where other metrics are required that allow obtaining complementary results to those obtained in the present review.

7. FUTURE WORK

This systematic literature review will allow us to incorporate a broader overview of the bankruptcy of companies at a general and specific level, where the identification of the factors that lead to an imbalance in business activities is vital and of very high consideration when implementing control mechanisms and indicators for decision-making. It also constitutes a solid basis for future research on the valuation of factors that generate a business breakdown and subsequent bankruptcy. Regarding the identification of factors, it is specified that those that constitute fundamental actions for the business must be considered; Therefore, each subsequent study will be a great advance for both the business field and for computational science.

It is important to note that the results for each research question posed will allow studies to be carried out on supervised learning models that provide better results in terms of metrics and methodologies more appropriate to the business context. In addition to this, it is important to

promote pre-test and pos-test studies on the incorporation and application of relevant variables such as the growth index (CI) and others that can be identified.

It is necessary to incorporate studies that identify metrics for other types of variables and respond more broadly to the research questions posed. Likewise, it is important that the studies are directed towards the forecast of companies in various types of environments and business contexts.

Finally, future studies are needed to identify the supervised learning models that are ideal for each company, which require the incorporation of variables according to the context of the company. The models identified in the studies require optimization work to improve each of the metrics, thus motivating the deployment of better supervised learning models for the company.

REFERENCES:

[1] Aljawazneh H., Mora A.M., García-Sánchez P., Castillo-Valdivieso P.A. "Comparison of the performance of deep learning methods to predict the financial failure of companies". 2021. 10.1109/ACCESS.2021.3093461

[2] Al-Milli N., Hudaib A., Obeid N. "Population diversity control of the genetic algorithm using a new injection method for the bankruptcy prediction problem". 2021. 10.3390/math9080823

[3] Ansari A., Ahmad I.S., Bakar A.A., Yaakub M.R. "A hybrid metaheuristic method in training artificial neural networks for bankruptcy prediction." 2020. 10.1109/ACCESS.2020.3026529

[4] Antunes F., Ribeiro B., Pereira F. "Probabilistic modeling and visualization for bankruptcy prediction". 2017. 10.1016/j.asoc.2017.06.043

[5] Atiku S.O., Obagbuwa I.C. "Machine Learning Classification Techniques to Detect the Impact of HR Outcomes on Commercial Bank Performance." 2021. 10.1155/2021/7747907

[6] Ben-Sassi, D. "A Rough Approach to Competitive Intelligence Based on Sets to Anticipate Competitor Action." 2022. Canada. https://doi.org/10.1016/j.eswa.2022.117523

[7] Brenes R.F., Johannssen A., Chukhrova N. "An intelligent model of bankruptcy prediction using a multilayer perceptron". 2022. 10.1016/j.iswa.2022.200136

[8] Chen H. "Prediction and analysis of the behavior of loans due to financial default based on a machine learning model". 2022. 10.1155/2022/7907210

[9] Chi D.-J., Shen Z.-D. "Using hybrid AI and machine learning technologies for sustainability in the prediction of going concern." 2022. 10.3390/SU14031810

[10] Ciampi F., Giannozzi A., Marzi G., Altman E.I. "Rethinking SME default prediction: a systematic review of the literature and future prospects". 2021 10.1007/S11192-020-03856-0

[11] Devi S.S., Radhika Y. "A Survey on Machine Learning and Statistical Techniques in Predicting Bankruptcy." 2018. 10.18178/ijmlc.2018.8.2.676

[12] Gao, B., & Balyan, V. "Construction of a financial default risk prediction model based on the LightGBM algorithm". 2022. Journal of Intelligent Systems, 31(1), 767-779. https://doi.org/10.1515/jisys-2022-0036

[13] García V., Marqués A.I., Sánchez J.S. "Exploring the synergistic effects of sample types on the behavior of sets for the prediction of credit risk and company bankruptcy".2019. 10.1016/j.inffus.2018.07.004

[14] Guha A., Veeranjaneyulu N. "Prediction of bankruptcy using big data analysis based on the fuzzy C-means algorithm". 2019. 10.11591/ijai.v8.i2.pp168-174

[15] Karminsky A.M., Burekhin R.N. "Comparative Analysis of Bankruptcy Forecasting Methods of Russian Construction Companies". 2019. 10.17323/1998-0663.2019.3.52.66

[16] Khalil A.A., Liu Z., Salah A., Fathalla A., Ali A. "Prediction of the insolvency of insurance companies in the Egyptian market using bagging techniques and boosting of sets ". 2022.10.1109/ACCESS.2022.3210032

[17] Le T., Lee M.Y., Park J.R., Baik S.W. "Oversampling techniques for bankruptcy prediction: novel features of a transaction dataset". 2018. 10.3390/sym10040079

[18] Le T., Son L.H., Vo M.T., Lee M.Y., Baik S.W. "A cluster-based momentum algorithm for bankruptcy prediction in a highly unbalanced dataset." 2018. 10.3390/sym10070250

[19] Le T., Vo M.T., Vo B., Lee M.Y., Baik S.W. "A hybrid approach using the oversampling technique and cost-sensitive learning for bankruptcy prediction." 2019. 10.1155/2019/8460934

[20] Lee, J., Jang, D., & Park, S. "Deep Learning-Based Corporate Performance Prediction Model Taking into Account Technical Capability." 2018. Sustainability, 9(6), 899. https://doi.org/10.3390/su9060899

[21] Li Z., Crook J., Andreeva G. "Dynamic prediction of financial difficulties using Malmquist DEA". 2017. 10.1016/j.eswa.2017.03.017

[22] Lin W.-C., Lu Y.-H., Tsai C.-F. "Selection of characteristics in bankruptcy prediction models based on single and joint learning". 2019. 10.1111/exsy.12335

[23] Lombardo G., Pellegrino M., Adosoglou G., Cagnoni S., Pardalos P.M., Poggi A. "Machine Learning for Prediction of Bankruptcies in the U.S. Stock Market: Datasets and Benchmarks." 2022. 10.3390/FI14080244

[24] Ma Y., Liu H., Zhai G., Huo Z. "Early warning of financial risk based on wireless network communication and the optimal fuzzy SVM artificial intelligence model." 2021. 10.1155/2021/7819011

[25] Mengist, W. Method for Conducting a Systematic Review of the Literature and Meta-Analysis for Research in Environmental Sciences. 2019. 10.1016/j.mex.2019.100777

[26] Miao J., Zhu W. "Precision-Recovery Curve (PRC) Classification Trees". 2022. 10.1007/s12065-021-00565-2

[27] Muthukumaran, K. and Hariharanath, K. "Deep Learning Enabled Financial Crisis Prediction Model for Small-Medium Sized Industries." 2022. India. https://doi.org/10.32604/iasc.2023.025968

[28] Muslim M.A., Dasril Y. "Company Bankruptcy Prediction Framework Based on the Most Influential Characteristics Using XGBoost and Stacking Ensemble Learning ". 2021. 10.11591/ijece.v11i6.pp5549-5557

[29] Nießner T., Nießner S., Schumann M. "Is it worth the effort? Considerations on text mining in AI-based corporate failure prediction". 2023. 10.3390/info14040215

[30] Onan A. "Subsampling Approach Based on Consensus Clusters for Unbalanced Learning." 2019. 10.1155/2019/5901087

[31] Page et al., "The 2020 PRISMA Statement: An Updated Guide to Reporting Systematic Reviews," BMJ, vol. 372, March 2021, doi: 10.1136/bmj.n71.

[32] Park M.S., Son H., Hyun C., Hwang H.J. "Explainability of Machine Learning Models for Bankruptcy Prediction." 2021. 10.1109/ACCESS.2021.3110270

[33] Perboli G., Arabnezhad E. "A Machine Learning-based DSS for medium- and long-term business crisis prediction". 2021. 10.1016/j.eswa.2021.114758

[34] Radovanovic J., Haas C. "The Evaluation of Bankruptcy Prediction Models Based on Socioeconomic Costs". 2023. 10.1016/j.eswa.2023.120275

[35] Ribeiro B., Silva C., Chen N., Vieira A., Carvalho Das Neves J. "Improved models of default risk with SVM". 2012. 10.1016/j.eswa.2012.02.142

[36] Sigrist F., Leuenberger N. "Machine learning for corporate default risk: Multi-period prediction, frailty corlation, loan portfolios, and tail probabilities." 2023. 10.1016/j.ejor.2022.06.035

[37] Siswoyo B., Abas Z.A., Pee A.N.C., Komalasari R., Suyatna N. "Optimization of the joint machine learning algorithm of bank failure prediction". 2022. 10.11591/ijai.v11.i2.pp679-686

[38] Song Y., Peng Y. A. "MCDM-based assessment approach for unbalanced classification methods in financial risk prediction". 2019. 10.1109/ACCESS.2019.2924923

[39] Yu, Q., Lendasse, A., Severin, E. "Ensemble KNNs for Bankruptcy Prediction. Helsinki University of Technology". 2018. Department of Information and Informatics. Espoo. Finland. University of Lille. Laboratoire Economie Management. Villeneuve d'Ascq cedex. France. https://doi.org/10.1007/s10902-010-9244-4

[40] Zhang, W., Yan, S., Li, J., Tian, X., and Yoshida, T. "Predicting SME Credit Risk in Supply Chain Finance by Merging Demographic and Behavioral Data." 2022. Transportation Research Part E: Logistics and Transportation Review, 158, 102611. https://doi.org/10.1016/j.tre.2022.102611

[41] Zou Y., Gao C., Gao H. "Business failure prediction based on a cost-sensitive extreme gradient increase machine." 2022. 10.1109/ACCESS.2022.3168857