Analysis of Neutrosophic Elements in the Determination of Bankruptcies in SMEs Using Machine Learning

J. Ramón R. de Vega, A. G. Ruiz Conejo, Carlos C. Carranza, Vladimir R. Cairo

National University of San Marcos, Lima, Peru

Emails: jramonr@unmsm.edu.pe; agallegosr@unmsm.edu.pe; ccabrerac@unmsm.edu.pe, vrodriguezc@unmsm.edu.pe

Abstract

Nowadays, Machine Learning techniques stand out, especially in the business sector, in predicting bankruptcies in small and medium-sized enterprises (SMEs). This reduces the probability of making bad investments when creating SMEs. Therefore, a systematic review of Machine Learning for predicting bankruptcies in SMEs was conducted to identify ideal articles. The search was conducted on Taylor & Francis Online, IEEE Xplore, ARDI, ScienceDirect, ACM Digital Library, Google Scholar, and ProQuest. As a result, information was collected from 84 definitive studies on determining bankruptcies in SMEs using Machine Learning. Therefore, this study aims to determine the state-of-the-art regarding determining bankruptcies in SMEs using Machine Learning. To obtain the results, the Saaty Neutrosophic AHP method was used to identify the most applied business sector and predict possible bankruptcy due to its broad nature of indeterminacy in that subset. The systematic review results have allowed for determining essential details regarding the state-of-the-art of determining bankruptcies in SMEs using Machine Learning.

Keywords: Keywork one; Keywork two; Keywork three; Keyword four; ….

1. Introduction

There is a lack of knowledge of the Determination of Bankruptcy in SMEs when using Machine Learning. The bankruptcy of SMEs in the market has been generated more frequently in recent years, constituting a consistent factor in a country's financial and labor crisis. Recent studies confirm a 30% increase in the desertion of SMEs. The continuity of SMEs in the market is significant, as it indicates progress and advancement, whether at the country or global level [1]. Therefore, it is evident and necessary to predict bankruptcy in SMEs since it maintains the development of an active market[2, 3]. The bankruptcy models provided by Machine Learning offer crucial information on the financial state of SMEs to investors, representing a significant aid in making investment decisions, identifying internal problems, and assessing viability. This article focuses on identifying relevant trends in bankruptcy prediction techniques with the support of Machine Learning [4, 5].

An exhaustive review of the models was developed for the study's modeling to identify the prediction of bankruptcies in SMEs [6]. However, it should be noted that subsets of indeterminate nature with neutrosophic nature prevail to analyze the set of variables that influence state of the art regarding the Determination of Bankruptcies in SMEs using Machine Learning. Among these subsets with the highest incidence in the study result are the business sectors that use the Neutrosophic AHP Saaty modeling to detect the point of highest incidence in the set[7].

Therefore, this review aims to determine *state of the art regarding the determination of Bankruptcy in SMEs using Machine Learning*. For this purpose, a systematic review of the literature is carried out, considering the

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

research questions, which play a very significant role in the search, extraction, and analysis of data. In addition to formulating the research questions, their respective objectives are also formulated and shown in Table 1.

Table 1: Research questions and objectives.

Research Question	Objectives
RQ1: In which business sectors are Bankruptcies in SMEs being determined?	Determine the business sectors in which bankruptcies are being determined in SMEs.
RQ2: What are the types of investigation most used for the Determination of Bankruptcy in SMEs using Machine Learning?	Determine what are the types of investigation most used for the Determination of Bankruptcy in SMEs using Machine Learning.
RQ3: What are the most used concepts in research on the Determination of Bankruptcy in SMEs using Machine Learning?	Identify the most used concepts in research on the Determination of Bankruptcy in SMEs using Machine Learning.
RQ4: What are the most used and relevant key phrases in the investigations on the Determination of Bankruptcy in SMEs using Machine Learning?	Determine the most used and relevant key phrases in the investigations on the Determination of Bankruptcy in SMEs using Machine Learning.
RQ5: What are the keywords that present co-occurrence in the investigations on the Determination of Bankruptcy in SMEs using Machine Learning?	Determine the keywords that present co-occurrence in the investigations on the Determination of Bankruptcy in SMEs using Machine Learning.

2. Materials and methods

2.1 Theoretical background

Currently, there are not many Systematic Literature Reviews that focus on the prediction of SME bankruptcies using Machine Learning techniques[8]. But some studies analyze documents on bankruptcy prediction, whether at the business or micro-enterprise level, and it is considered essential to define it, since it allows guiding the review during the process of exclusion and analysis of the studies. In addition, a notorious application of Machine Learning has been visualized in this business sector to predict bankruptcy in SMEs, to maintain the continuity of activity in the financial and labor sectors. This highlights the great usefulness of Neural Networks, as they have a prediction effectiveness according to AUC (area under the curve), which ranges from 81% to 88%, representing an effective prediction of the use of Neural Networks[9]. This also confirms that there is a high interest in predicting bankruptcies in SMEs using Machine Learning [10, 11]. Therefore, the review method has been carried out considering the stages of the suggested methodology for the systematic literature review[12]. The method is made up of the steps for the systematic review process:

- Research problems
- Sources and search strategies
- Selection criteria
- Selection of studies
- Quality assessment
- Data evaluation
- Data synthesis

2.2 Neutrosophic AHP

The Analytic Hierarchy Process (AHP) was proposed by Thomas Saaty in 1980. It is one of the most widespread methods for solving multicriteria decision-making problems[13, 14]. This technique models the problem, forming a representative hierarchy of the associated decision-making schema. This hierarchy presents the objective pursued in solving the problem at the upper level and at the lower level. The different alternatives are included from which a decision must be made [15]. The intermediate levels detail the set of criteria and attributes considered. For the description of the method, it is necessary to expose the following definitions[16, 17]:

Definition 1: The Neutrosophic set N is characterized by three membership functions, which are the truth-membership function TA, indeterminacy-membership function IA, and falsehood-membership function FA, where U is the Universe of Discourse and $\forall x \in U, TA(x), IA(x), FA(x) \subseteq]-0, 1+[$, and $-0 \leq inf\ TA(x) + inf\ IA\ (x) + inf\ FA\ (x) \leq sup\ TA(x) + sup\ IA\ (x) + sup\ FA\ (x) \leq 3+$. Notice that, according to the

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

definition, $TA(x)$, $IA(x)$ and $FA(x)$ are real standard or non-standard subsets of $]-0,1+[$ and hence, $TA(x)$, $IA(x)$ and $FA(x)$ can be subintervals of.

Definition 2: The *Single-Valued Neutrosophic Set* (SVNS) N over U is $A = \{< x; TA(x), IA(x), FA(x) > : x \in U\}$, where $TA: U \rightarrow [0,1], IA: U \rightarrow [0,1], and\ FA: U \rightarrow [0,1], 0 \leq TA(x) + IA(x) + FA(x) \leq 3$. The *Single-Valued Neutrosophic Number* (SVNN) is represented by $N = (t, I, f)$, such that $0 \leq t, I, f \leq 1\ and\ 0 \leq t + I + f \leq 3$.

Definition 3: the *single-valued trapezoidal neutrosophic number*, $\tilde{a} = \langle(a_1, a_2.a_3, a_4); \alpha_{\tilde{a}}, \beta_{\tilde{a}}, \gamma_{\tilde{a}}\rangle$, is a neutrosophic set on $\mathbb{R}$, whose truth, indeterminacy and falsehood membership functions are defined as follows, respectively:

$$T_{\tilde{a}}(x) = \begin{cases} \alpha_{\tilde{a}\left(\frac{x-a_1}{a_2-a_1}\right)}, & a_1 \leq x \leq a_2 \\ \alpha_{\tilde{a}}, & a_2 \leq x \leq a_3 \\ \alpha_{\tilde{a}\left(\frac{a_3-x}{a_3-a_2}\right)}, & a_3 \leq x \leq a_4 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$I_{\tilde{a}}(x) = \begin{cases} \frac{(a_2 - x + \beta_{\tilde{a}}(x - a_1))}{a_2 - a_1}, & a_1 \leq x \leq a_2 \\ \beta_{\tilde{a}}, & a_2 \leq x \leq a_3 \\ \frac{(x - a_2 + \beta_{\tilde{a}}(a_3 - x))}{a_3 - a_2}, & a_3 \leq x \leq a_4 \\ 1, & \text{otherwise} \end{cases} \tag{2}$$

$$F_{\tilde{a}}(x) = \begin{cases} \frac{(a_2 - x + \gamma_{\tilde{a}}(x - a_1))}{a_2 - a_1}, & a_1 \leq x \leq a_2 \\ \gamma_{\tilde{a}}, & a_2 \leq x \leq a_3 \\ \frac{(x - a_2 + \gamma_{\tilde{a}}(a_3 - x))}{a_3 - a_2}, & a_3 \leq x \leq a_4 \\ 1, & \text{otherwise} \end{cases} \tag{3}$$

Where $\alpha_{\tilde{a}}, \beta_{\tilde{a}}, \gamma_{\tilde{a}} \in [0,1]$, $a_1, a_2, a_3, a_4 \in \mathbb{R}$ and $a_1 \leq a_2 \leq a_3 \leq a_4$.

Definition 4: given $\tilde{a} = \langle(a_1, a_2, a_3, a_4); \alpha_{\tilde{a}}, \beta_{\tilde{a}}, \gamma_{\tilde{a}}\rangle$ and $\tilde{b} = \langle(b_1, b_2, b_3, b_4); \alpha_{\tilde{b}}, \beta_{\tilde{b}}, \gamma_{\tilde{b}}\rangle$ two single-valued trapezoidal neutrosophic numbers and λ any non-null number in the real line. Then, the following operations are defined:

Addition: $\tilde{a} + \tilde{b} = \langle(a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4); \alpha_{\tilde{a}} \wedge \alpha_{\tilde{b}}, \beta_{\tilde{a}} \vee \beta_{\tilde{b}}, \gamma_{\tilde{a}} \vee \gamma_{\tilde{b}}\rangle$

Subtraction: $\tilde{a} - \tilde{b} = \langle(a_1 - b_4, a_2 - b_3, a_3 - b_2, a_4 - b_1); \alpha_{\tilde{a}} \wedge \alpha_{\tilde{b}}, \beta_{\tilde{a}} \vee \beta_{\tilde{b}}, \gamma_{\tilde{a}} \vee \gamma_{\tilde{b}}\rangle$ (4)

Inversion: $\tilde{a}^{-1} = \langle(a_4{}^{-1}, a_3{}^{-1}, a_2{}^{-1}, a_1{}^{-1}); \alpha_{\tilde{a}}, \beta_{\tilde{a}}, \gamma_{\tilde{a}}\rangle$, where $a_1, a_2, a_3, a_4 \neq 0$.

Multiplication by a scalar number:

$$\lambda\tilde{a} = \begin{Bmatrix} \langle(\lambda a_1, \lambda a_2, \lambda a_3, \lambda a_4); \alpha_{\tilde{a}}, \beta_{\tilde{a}}, \gamma_{\tilde{a}}\rangle, & \lambda > 0 \\ \langle(\lambda a_4, \lambda a_3, \lambda a_2, \lambda a_1); \alpha_{\tilde{a}}, \beta_{\tilde{a}}, \gamma_{\tilde{a}}\rangle, & \lambda < 0 \end{Bmatrix}$$

Definitions 3 and 4 refer to *single-valued triangular neutrosophic number* when the condition $a_2 = a_3$. For simplicity, the linguistic scale of triangular neutrosophic numbers is used, see Table 2.

Formulating the decision-making problem in a hierarchical structure is the first and main stage. In this stage, the decision maker must break down the problem into its relevant components. The hierarchy is constructed so that the elements are of the same order of magnitude and can be related to some of the next level. In a typical hierarchy the highest level locates the problem of decision making[18]. The elements that affect decision making are represented at the intermediate level, the criteria occupying the intermediate levels. At the lowest level the decision options are understood. The criteria' importance or weighting levels are estimated through paired comparisons between them. This comparison uses a scale, as expressed in equation (6).

$$S = \left\{\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1,3,5,7,9\right\} \tag{5}$$

In a neutrosophic framework, AHP technique can be used to model the indeterminacy of decision-making by applying neutrosophic AHP or NAHP for short. Equation 7 contains a generic neutrosophic pair-wise comparison matrix for NAHP[19].

$$\tilde{A} = \begin{bmatrix} \tilde{1} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \vdots & & \ddots & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots & \tilde{1} \end{bmatrix} \quad (6)$$

Matrix $\tilde{A}$ must satisfy condition $\tilde{a}_{ji} = \tilde{a}_{ij}^{-1}$, based on the inversion operator of Definition 4.

To convert neutrosophic triangular numbers into crisp numbers, there are two indexes defined in, they are the so-called score and accuracy indexes, respectively, see Equations 8 and 9[20, 21]:

$$S(\tilde{a}) = \frac{1}{8}[a_1 + a_2 + a_3](2 + \alpha_{\tilde{a}} - \beta_{\tilde{a}} - \gamma_{\tilde{a}}) \quad (7)$$

$$A(\tilde{a}) = \frac{1}{8}[a_1 + a_2 + a_3](2 + \alpha_{\tilde{a}} - \beta_{\tilde{a}} + \gamma_{\tilde{a}}) \quad (8)$$

Step 1 Select a group of experts.
Step 2 Structure the neutrosophic pair-wise comparison matrix of factors, sub-factors, and strategies, through the linguistic terms shown in Table 2.

Table 2: Saaty's scale translated to a neutrosophic triangular scale.

Saaty's scale	Definition	Neutrosophic Triangular Scale
1	Equally influential	$\tilde{1} = \langle(1,1,1); 0.50, 0.50, 0.50\rangle$
3	Slightly influential	$\tilde{3} = \langle(2,3,4); 0.30, 0.75, 0.70\rangle$
5	Strongly influential	$\tilde{5} = \langle(4,5,6); 0.80, 0.15, 0.20\rangle$
7	Very strongly influential	$\tilde{7} = \langle(6,7,8); 0.90, 0.10, 0.10\rangle$
9	Absolutely influential	$\tilde{9} = \langle(9,9,9); 1.00, 1.00, 1.00\rangle$
2, 4, 6, 8	Sporadic values between two close scales	$\tilde{2} = \langle(1,2,3); 0.40, 0.65, 0.60\rangle$ $\tilde{4} = \langle(3,4,5); 0.60, 0.35, 0.40\rangle$ $\tilde{6} = \langle(5,6,7); 0.70, 0.25, 0.30\rangle$ $\tilde{8} = \langle(7,8,9); 0.85, 0.10, 0.15\rangle$

The neutrosophic scale is attained according to expert opinions. The neutrosophic pair-wise comparison matrix of factors, sub-factors, and strategies are as described in Equation 6.
Step 3 Check the consistency of experts' assessments.
If the pair-wise comparison matrix has a transitive relation, i.e., $a_{ik} = a_{ij}a_{jk}$ for all i,j and k, then the comparison matrix is consistent, focusing only on the lower, median and upper values of the triangular neutrosophic number of the comparison matrix.
Step 4 Calculate the weight of the factors from the neutrosophic pair-wise comparison matrix, by transforming it to a deterministic matrix using Equations 9 and 10. To get the score and the accuracy degree of $\tilde{a}_{ji}$ the following equations are used:

$$S(\tilde{a}_{ji}) = 1/S(\tilde{a}_{ij}) \quad (9)$$

$$A(\tilde{a}_{ji}) = 1/A(\tilde{a}_{ij}) \quad (10)$$

With compensation by accuracy degree of each triangular neutrosophic number in the neutrosophic pair-wise comparison matrix, the following deterministic matrix is derived:

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix} \quad (11)$$

Determine the ranking of priorities, namely the Eigen Vector X, from the previous matrix:
1. Normalize the column entries by dividing each entry by the sum of the column.
2. Take the total of the row averages.

Note that Step 3 refers to consider the use of the calculus of the *Consistency Index* (CI) when applying this technique, which is a function depending on λ_{max}, the maximum eigenvalue of the matrix. Saaty establishes that consistency of the evaluations can be determined by equation:

$$CI = \frac{\lambda_{max} - n}{n-1}, \quad (12)$$

where *n* is the order of the matrix. In addition, the *Consistency Ratio* (CR) is defined by equation:

$$CR = \frac{CI}{RI} \quad (13)$$

RI is given in Table 3.

Table 3: RI associated to each order.

Order (n)	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

If $CR \leq 0.1$ it can be considered that experts' evaluation is sufficiently consistent and hence NAHP can be used. This procedure is applied to matrix "*A*" in Equation 12.

3.Results

3.1 Information Sources and Search Strategies

The following libraries were selected as sources of information: Taylor & Francis Online, ARDI, ScienceDirect, ACM Digital Library, Google Scholar, IEEE Xplore and ProQuest.

On the other hand, the search strategy included keywords that were relevant to the study (see Table 4).

Table 4: Search descriptors and their synonyms..

Descriptor
Bankruptcy prediction/ Failure Prediction
Small companies/ Small Businesses/ SMEs
Machine Learning / ML

The search procedure has been carried out considering the search equations for the study, as shown in Table 5.

Table 5: Search equations.

Source	Search query
ACM Digital Library	[[All: "machine learning"] OR [All: ml]] AND [[All: "bankruptcy prediction"] OR [All: "failure prediction"]] AND [[All: "small companies"] OR [All: "small businesses"] OR [All: sme]]
Taylor & Francis Online	[[All: "machine learning"] OR [All:ml]] AND [[All: "bankruptcy prediction"] OR [All: "failure prediction"]] AND [[All: "small companies"] OR [All: "small businesses"] OR [All: sme]]
ScienceDirect	("machine learning" OR ML) AND (("bankruptcy prediction" OR "failure prediction") AND ("small companies" OR "small businesses" OR SMEs))

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

ProQuest	("machine learning" OR ML) AND (("bankruptcy prediction" OR "failure prediction") AND ("small companies" OR "small businesses" OR SMEs))
ARDI	(("machine learning") OR (ml)) AND (("bankruptcy prediction") OR ("failure prediction")) AND (("small companies") OR ("small businesses") OR (SME))
Google Scholar	("machine learning" OR ML) AND (("bankruptcy prediction" OR "failure prediction") AND ("small companies" OR "small businesses" OR SMEs))
IEEE Xplore	("All Metadata":"Machine Learning") OR ("All Metadata":ML) AND ("All Metadata":"bankruptcy prediction") OR ("All Metadata":"failure prediction") AND ("All Metadata":"small companies") OR ("All Metadata":"small businesses") OR ("All Metadata":SMEs)

3.2 Search Results

After carrying out the article search, several articles are obtained as a result, which are shown in Figure 1.

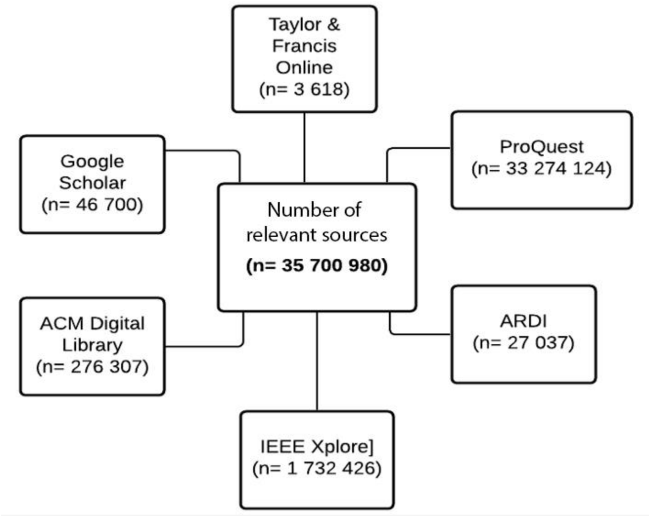

Figure 1: Articles found grouped by Source

3.3 Exclusion criteria

The exclusion criteria were defined to evaluate all the articles found carefully. Each article was validated considering the following criteria:

EC1. The articles are older than 5 years.
EC2. The articles are not written in English.
EC3. The articles were not published in Peer-reviewed Conferences or Journals
EC4. The articles do not belong to the appropriate specialty.
EC5. The titles and keywords of the articles are not very suitable.
EC6. Items are not unique.

This helps to have greater clarity in the selection of studies. And it allows obtaining an adequate number of articles to be reviewed regarding the topic under study.

3.4 Study Selection

Initially, 35700980 articles were obtained after a rigorous search using the keywords relevant to the study. The stages shown in Figure 2 were used for the selections and filters. The result was 84 articles.

3.5 Quality Assessment and Data Extraction Strategies

The application of a set of quality assessment (QA) standards was the next step to determine the final list of articles to review. Seven QAs were used to estimate the quality of the selected articles:

QA1. Is the document well organized?
QA2. Is there enough background information in the document?
QA3. Is the specific topic area used clearly defined?

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

QA4. Are the methods used to analyze the results appropriate?
QA5. Are the results of the experiments performed clearly identified and reported?
QA6. Are the research objectives clearly identified in the document?
QA7. Is the document considered useful?

At this stage, the final list of articles was used to extract the information necessary to answer the research questions.

The information extracted from each article includes the following characteristics: article ID, article title, URL, source, year, country, number of pages, language, publication type, publication name, authors, quartile, H-index, affiliation, number of citations, abstract, keywords, sample size.

It is also important to note that all the articles responded positively to all the QAs.

Mendeley was used to perform the data extraction as shown in Figure 3.

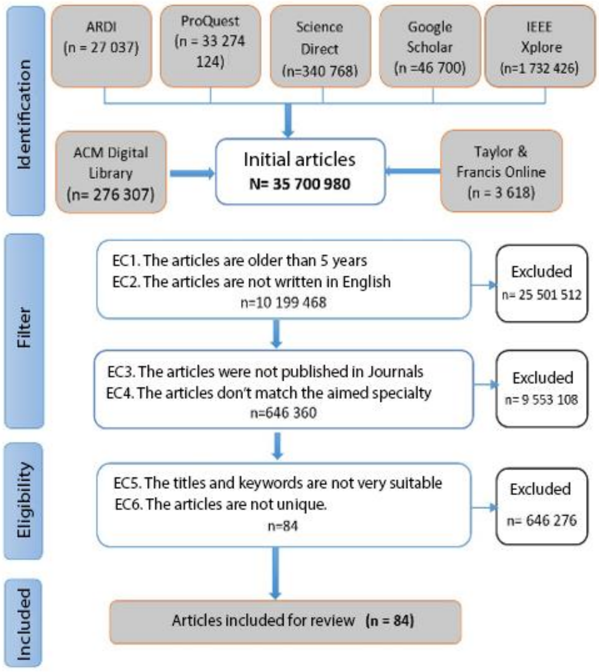

Figure 2: Quality assessment.

3.6 Synthesis of Findings or Synthesis of the Data

The information extracted for research questions RQ1-RQ5 was tabulated and presented as quantitative data that was used to develop a statistical comparison between the different findings for each research question. These data have been used to respond to each of the five RQs previously stated.

4.Discussion

4.1 Overview of studies

In the development of the selection process, 84 studies were obtained as a result for data extraction and analysis. Figure 3 presents the distribution of the studies published from 2016 to 2021. In 2020, there was a greater record of studies, which can be deduced that recently there has been a greater emphasis in studies on the prediction of bankruptcy of SMEs in relation to Machine Learning. In addition, Figure 4 shows the number of articles by information sources, where there is a higher percentage in ARDI (34.5%) with 29 published articles.

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

Figure 3: Articles published by Year

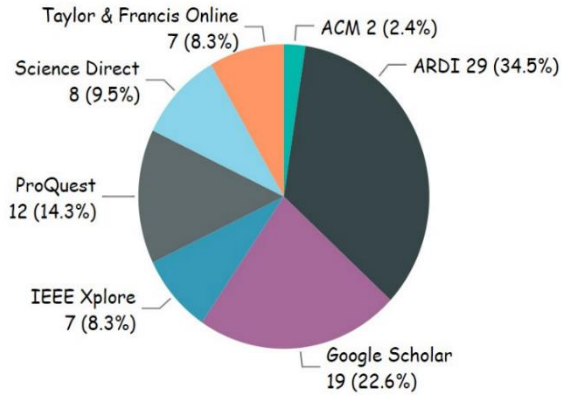

Figure 4: Articles by Source.

4.2 Answers to Research Questions

RQ1. In which business sectors is the Bankruptcy of SMEs being determined?
According to the results of the systematic review of the literature, eight areas of application of machine learning for the prediction of bankruptcies in companies were identified. Of the eight areas, it was decided to leave 6 areas, for this the *Marketing & Entrepreneurship/startup* sector was combined within the same element to the neutrosophic subset on the same percentage level of the analyzed reference. It is necessary to clarify that the element Sales is ruled out due to the low representativeness of 2.5% of the references analyzed. Table 6 shows that the Finance area with 36.9% use and a weight of 0.51 over the other elements analyzed, so that it constitutes the most applied to predict possible bankruptcy. The Neutrosophic AHP method was used to obtain the exposed data, where the weights of the principles on which the procedure for evaluating social projects will be determined (see tables 6 and 7).

Table 6: Neutrosophic AHP paired matrix.

Areas	**Finance**	**Economy**	**banking**	**business**	**logistics**	**Marketing**	**Weight**
Finance	**1**	〈(6,7,8);0.90,0.10,0.10〉	〈(6,7,8);0.90,0.10,0.10〉	〈(6,7,8);0.90,0.10,0.10〉	〈(7,8,9);0.85,0.10,0.15〉	〈(5,6,7);0.70,0.25,0.30〉	**0.51**
Economy	〈(6,7,8);0.90,0.10,0.10〉	**1**	〈(2,3,4);0.30,0.75,0.70〉	〈(4,5,6);0.80,0.15,0.20〉	〈(3,4,5);0.60,0.35,0.40〉	〈(6,7,8);0.90,0.10,0.10〉	**0.21**
Banking	〈(6,7,8);0.90,0.10,0.10〉	〈(2,3,4);0.30,0.75,0.70〉	**1**	〈(2,3,4);0.30,0.75,0.70〉	〈(2,3,4);0.30,0.75,0.70〉	〈(2,3,4);0.30,0.75,0.70〉	**0.11**
Business	〈(6,7,8);0.90,0.10,0.10〉	〈(4,5,6);0.80,0.15,0.20〉	〈(2,3,4);0.30,0.75,0.70〉	**1**	〈(1,1,1);0.50,0.50,0.50〉	〈(2,3,4);0.30,0.75,0.70〉	**0.06**
Logistics	〈(7,8,9);0.85,0.10,0.15〉	〈(3,4,5);0.60,0.35,0.40〉	〈(2,3,4);0.30,0.75,0.70〉	〈(1,1,1);0.50,0.50,0.50〉	**1**	〈(1,1,1);0.50,0.50,0.50〉	**0.05**
Marketing & Entrepreneurship/ startup	〈(5,6,7);0.70,0.25,0.30〉	〈(6,7,8);0.90,0.10,0.10〉	〈(2,3,4);0.30,0.75,0.70〉	〈(2,3,4);0.30,0.75,0.70〉	〈(1,1,1);0.50,0.50,0.50〉	**1**	**0.04**
Total	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**-**

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

Table 7: Machine Learning application areas.

Areas	W x Weight	Approximate eigenvalues
Finance	3.92	7.632215878
Economy	1.46	6.77524613
Banking	0.73	6.384170759
Business	0.40	6.151204195
Logistics	0.31	6.491604374
Marketing & Entrepreneurship/startup	0.27	6.115818517
8.830773916 Eigen value		

The analysis of the consistency of the method revealed that its eigenvalue is 8.83, IC=0.12 and RC=0.08, therefore it is confirmed that the modeling was satisfactory. The Finance area was the most prominent, with 31 publications, where machine learning was applied to predict bankruptcies in companies, implying that it is the area with the greatest focus on this problem to be detected.

RQ2. What are the types of investigation most used for the determination of Bankruptcy in SMEs when using Machine Learning?

According to the review of the studies, to obtain the most used types of research, a pie chart was developed, as shown in Figure 5, where it is shown that there are more studies of the Experimental type. 48.8% of the studies are of an experimental type, constituting the largest number of all the studies. It is recommended to consider the experimental articles, where the work with real data is visualized and the effectiveness of the determination of Bankruptcy in SMEs can be verified when using Machine Learning.

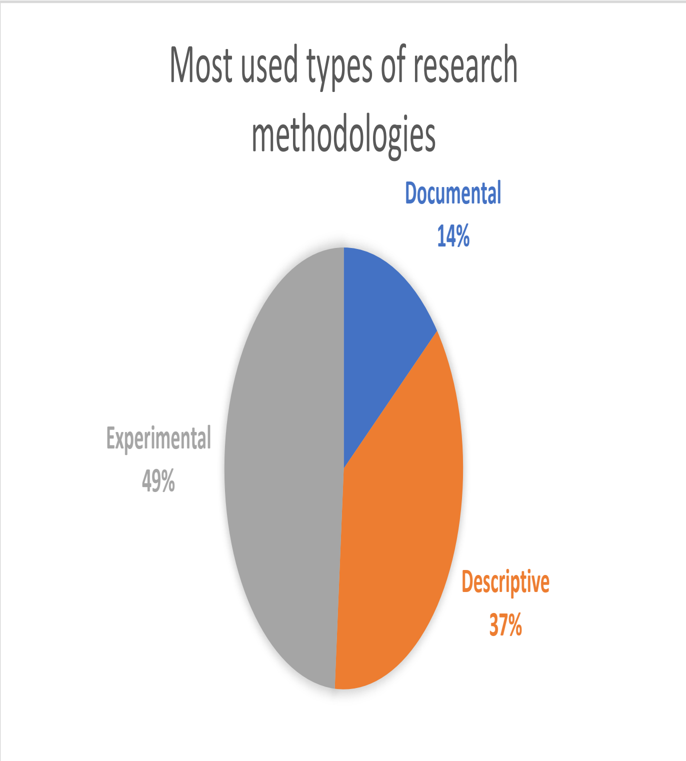

Figure 5: Types of research most used.

RQ3. What are the most used concepts in the investigations on determining Bankruptcy in SMEs when using Machine Learning?

The articles selected in this SRL were classified into 4 quadrants with the most used concepts according to a portfolio of trigrams, which is shown in Figure 6. In the CORE quadrant, the "support vector machines" trigram is present in many articles and many citations, in the EMERGING quadrant the trigram that appears the most in the last 3 years is "artificial neural network", in the DECLINING quadrant the trigram that appears the least in the last

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

3 years is "absolute percentage error" and finally, in The ESTABLISHED quadrant shows for the second time the concept "support vector machines" as the trigram that appears with many citations greater than 5 and less than 15.

DECLINING

Trigrama	2018	2019	2020	Total
absolute percentage error		1		1
accuracy precision sensitivity	1			1
accuracy true negative			1	1
accurate bankruptcy prediction		1		1
accurate decisionmaking model		1		1
achieve better performance	1			1
achieve better results		1		1
achieves much better		1		1
active learning skills			1	1
actlct cashat fat			1	1
actual price series	1			1
adaboost xgboost svc			1	1
adaptive elman network	1			1
adaptive elman tool	1			1
add certain degree			1	1
Total	511	481	641	1633

Figure 6: Declining classification.

According to the portfolio, the main approaches, or the most outstanding concepts of all the articles can be highlighted. Similarities are found in two of the machine learning methods as the most widely used for predicting bankruptcies in companies ("artificial neural networks" and "support vector machines"), by highlighting the years reviewed for each article and the number of citations of each trigram. To obtain these results, natural language processing (NLP) was used, as it is something innovative to obtain this type of response.

RQ4. What are the most used and relevant key phrases in the investigations on the Determination of Bankruptcy in SMEs using Machine Learning?

After processing the articles' content using natural language processing (NLP), a series of relevant phrases (topics) were identified with their corresponding score. Then the generated key phrases were processed, and a word cloud was created using the scores as shown in Table 8.

Table 8: Most used key phrases and their Scores.

Keyphrase	Score	Keyphrase	Score	Keyphrase	Score
Dow Jones industrial average	48.00	Assuming p holds	16.00	Effectiveness became widely known	16.00
Root mean square error	44.34	Avoid oscillations towards reducing	16.00	Eliminate short term fluctuation	16.00
Multicriteria decision aid methodology	32.00	Compelling driving force pertaining	16.00	Establishing early warning systems	16.00
Autoregressive integrated moving average	31.17	Consistent estimator named meanmap	16.00	Evaluate business loan applications	16.00
Order single knot spline	31.00	Daily ecological practices allow	16.00	Existing functional forms relating	16.00
Squared automatic interaction detector	30.67	Declaring huge annual dividends	16.00	Federal deposit insurance corporation	16.00
Root mean squared error	30.07	Drawing receiver operating characteristic	16.00	Fourth industrial revolution refers	16.00
Ad hoc prudential requirements	16.00	effectively split since june	16.00	gini coefficient gini	16.00

The relevance of these key phrases in the influence of Machine Learning in bankruptcy prediction can be visualized. The most used phrases "Down Jones industrial average" and "Root mean square error" stand out with scores of 48 and 44.4 respectively.

RQ5. What are the keywords that present co-occurrence in the investigations on the Determination of Bankruptcy in SMEs using Machine Learning?

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

Based on the systematic review of the articles, the keywords were extracted and organized by means of a bibliometric network to visualize the co-occurrence between all the keywords throughout the articles (See Figure 7 and Figure 8). Of the two main keywords ("machine learning" and "bankruptcy prediction"), it is shown that they correlate more with the keywords "artificial neural network" and "support vector machine".

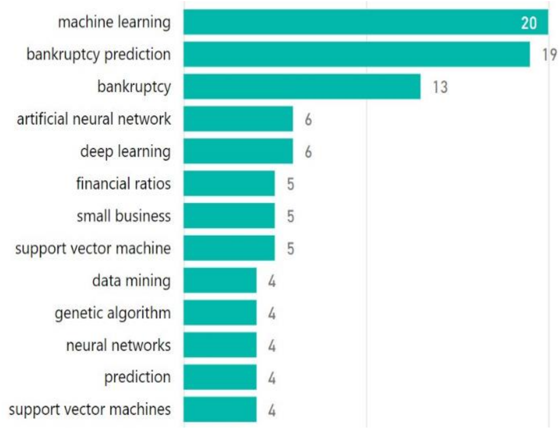

Figure 7: Keywords most used in the articles.

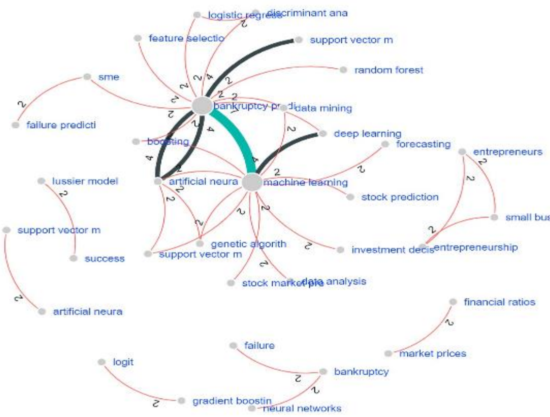

Figure 8: Keywords bibliometric network.

After investigating other systematic review articles, neither this type of analysis nor its results have been found, so it is not possible to make any comparison.

5. Conclusion

The development of this systematic review has allowed us to delve into the theories and applications of Machine Learning for the prediction of bankruptcies in SMEs, by extracting data from 84 studies published between 2016 and 2021. All to the data sources used for selecting bankruptcy states were written in English. Regarding the efficiency evaluation of the methods used, it was found that most of the articles used AUC (Area under the ROC curve) and ROC (Receiver operating characteristic curve). To minimize the indeterminacies existing in the subset of the analyzed variable (Machine Learning application areas), the modeling of the Saaty Neutrosophic AHP method is used. It defines 6 areas, where Finance stands out for presenting more publications with the use of machine learning in the prediction of bankruptcies in companies, implying that it is the area with the greatest focus to be analyzed. Although in the indeterminacies there are neutrosophic links in Economy and Banking. The study in most of the articles reviewed uses support vector machines (SVM) to extract characteristics of bankrupt SMEs. Most of the studies highlighted Machine Learning techniques as the evaluation criteria. The articles focused more on the methods than the areas for predicting bankruptcies. Although in the consulted systematic reviews it was not possible to find approaches on the area or the use of any tool to show bibliometric networks of the keywords or the key concepts to make the comparisons. Therefore, for future research, a more significant number of studies should be considered, which consider Machine Learning and its use to determine bankruptcies in SMEs and each neutrosophic subset that influences the results.

References

[1] Sampedro Guaman, C.R., D.P. Palma Rivera, S.A. Machuca Vivar, and E.V. Arrobo Lapo, *Digital Transformation of Marketing in Small and Medium Enterprises Through Social Networks: Plitogenic Decision-Making.* Neutrosophic Sets & Systems, 2021. **44**.

[2] Alshikho, M., M. Jdid, and S. Broumi, *Artificial Intelligence and Neutrosophic Machine learning in the Diagnosis and Detection of COVID 19.* **Prospects for Applied Mathematics and Data Analysis**, 2023. **Vol. 1**(No. 2): p. 17-27.

[3] Zuñiga, V.C., A.M. León, D.M. Nogueira, D.A. Valencia, and J.M. Romero, *Validation of A Model for Knowledge Management in the Cocoa Producing Peasant Organizations of Vinces Using Neutrosophic Iadov Technique.* Neutrosophic Sets and Systems, 2019. **30**: p. 253-260.

[4] Arnaiz, N.V.Q., N.G. Arias, and L.C.C. Muñoz, *Neutrosophic K-means Based Method for Handling Unlabeled Data.* Neutrosophic Sets and Systems, 2020. **37**: p. 308-315.

[5] Sharma, M., I. Kandasamy, and W. Vasantha, *Comparison of neutrosophic approach to various deep learning models for sentiment analysis.* Knowledge-Based Systems, 2021. **223**: p. 107058.

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023

[6] A. B. Bonilla Rodríguez , M.F.C.M., F. B. Morocho Quinchuela, *Construction of Sanda Teaching risk assessment index system using neutrosophic AHP method.* International Journal of Neutrosophic Science, 2022. **18**(4): p. 334-343.

[7] Comas Rodríguez, R., J.M.D. Oca Sánchez, and V. Lucero Salcedo, *Evaluation of Social Projects Using Neutrosophic AHP.* International Journal of Neutrosophic Science, 2022. **19**(1): p. 280-288.

[8] Abdel Nasser H. Zaied , Shaimaa Mohmed, *ERP Implementation Road Map for Small and Medium Size Enterprises (SMEs)*, Journal of Intelligent Systems and Internet of Things, Vol. 2 , No. 1 , (2020) : 14-25 (Doi : https://doi.org/10.54216/JISIoT.020102).

[9] Rashno, E., A. Akbari, and B. Nasersharif. *A convolutional neural network model based on neutrosophy for noisy speech recognition*. in *2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA)*. 2019. IEEE.

[10] Abdulaziz Shehab , Mahmood Mahmood, *An Intelligent Bankruptcy Prediction Model based on an Enhanced Sparrow Search Algorithm*, Journal of Intelligent Systems and Internet of Things, Vol. 6 , No. 1 , (2022) : 09-19 (Doi : https://doi.org/10.54216/JISIoT.060101).

[11] Kandasamy, I., W. Vasantha, J.M. Obbineni, and F. Smarandache, *Sentiment analysis of tweets using refined neutrosophic sets.* Computers in Industry, 2020. **115**: p. 103180.

[12] Woodall, W.H., A.R. Driscoll, and D.C. Montgomery, *A review and perspective on neutrosophic statistical process monitoring methods.* IEEE Access, 2022.

[13] Smarandache, F., J.E. Ricardo, E.G. Caballero, M.Y.L. Vázquez, and N.B. Hernández, *Delphi method for evaluating scientific research proposals in a neutrosophic environment.* Infinite Study, 2020. **34**: p. 204-213.

[14] Aurelia Maria, C.B., I.Q. Janneth Ximena, and P.P. Alex Javier, *Multi-Criteria Analysis of Pollution rivers.* International Journal of Neutrosophic Science, 2022. **19**(1): p. 306-313.

[15] Becerra Arévalo, N.P., M.F. Calles Carrasco, J.L. Toasa Espinoza, and M. Velasteguí Córdova, *Neutrosophic AHP for the prioritization of requirements for a computerized facial recognition system.* Neutrosophic Sets and Systems, 2020. **34**(1): p. 21.

[16] Ortega, R.G., M.D.O. Rodríguez, M.L. Vázquez, J.E. Ricardo, J.A.S. Figueiredo, and F. Smarandache, *Pestel analysis based on neutrosophic cognitive maps and neutrosophic numbers for the sinos river basin management.* Neutrosophic Sets and Systems, 2019. **26**: p. 105-113.

[17] Noboa, M.F.O., O.E.P. Copa, and A.N.G. Eloísa, *Comparative analysis of multicriteria methods based on single-valued neutrosophic numbers for the evaluation of medical technologies.* International Journal of Neutrosophic Science, 2022. **18**(4): p. 72-82.

[18] Bonilla Rodríguez, A.B., M.F. Cueva Moncayo, and F.B. Morocho Quinchuela, *Construction of Sanda Teaching risk assessment index system using neutrosophic AHP method.* International Journal of Neutrosophic Science, 2022. **18**(4): p. 334-343.

[19] R. Comas Rodríguez , J.M.D.O.S., V. Lucero Salcedo, *Evaluation of Social Projects Using Neutrosophic AHP.* International Journal of Neutrosophic Science, 2022. **Vol. 19**(No. 1).

[20] Barragán, M.F.L., R.F.G. Montenegro, and L.L.C. María De, *Application of Neutrosophic Techniques for the Selection of the in-Hospital Triage System.* International Journal of Neutrosophic Science, 2022. **18**(4): p. 116-124.

[21] Quemac, R.E.C., M.E.G. Santos, L.A.C. Ramos, and C.P.C. Zúñiga, *Neutrosophic Analytic Hierarchy Process for the Analysis of Innovation in Latin America.* Neutrosophic Sets and Systems, 2021. **44**: p. 411-419.

Doi : https://doi.org/10.54216/IJNS.200412
Received: January 13, 2023 Accepted: March 29, 2023