Neutrosophic Sets and Systems, {Special Issue: Artificial Intelligence, Neutrosophy, and Latin American Worldviews: Toward a Sustainable Future (Workshop – March 18–21, 2025, Universidad Tecnológica de El Salvador, San Salvador, El Salvador)}, Vol. 84, 2025

University of New Mexico

Neutrosophic computational model for identifying trends in scientific articles using Natural Language Processing

Omar Mar Cornelio [1]*, and Barbara Bron Fonseca [2]

[1] University of Computer Sciences. Havana, Cuba. omarmar@uci.cu
[2] Soluciones MarBro SRL. Artemisa. Cuba. barbara.bron.fonseca@gmail.com

Abstract: This study presents an innovative neutrosophic computational model for identifying trends in scientific articles through Natural Language Processing (NLP). The primary objective was to develop a system that could effectively analyze a large body of academic literature and extract relevant summaries that capture the most significant themes and findings. Advanced NLP techniques were employed to extract, process, and synthesize key information, enabling a deeper understanding of emerging trends. Furthermore, the model was designed to evaluate the relevance and validity of the generated summaries, ensuring that the information presented is accurate and useful for researchers and practitioners. The results demonstrate that the system is not only efficient in generating coherent summaries but also facilitates the identification of critical themes in scientific research, potentially guiding future research and applications in the field. Thus, this model is established as a valuable tool for the academic community, promoting more agile and effective access to relevant information in a constantly evolving scientific context.

Keywords: Natural Language Processing; Neutrosophic Logic; Trend Analysis; Information Extraction; Summary Validity; Neutrosophic Computational Model.

1. Introduction

The advent of the Internet has generated a massive flow of information, making its retrieval more difficult. Most scientific information is found in scientific articles, and with the expansion of research fields, it can be quite difficult for scholars to find documents relevant to their interests [1]. Even query-based searches for some specific fields return a large number of relevant articles that far exceed human processing capabilities [2]. Automatic summarization of these articles would be useful to reduce the time needed to review them in their entirety and obtain the essence of the information contained in them [3]. Mainly, summaries could be generated in two ways: single-document summaries where the task is to generate a summary from a single source and multi-document summaries where different but related documents are summarized, comprising only the essential materials or main ideas of a document in less space [4].

The scientific research process generally begins with the examination of academic scientific journals based on an infographic methodological analysis. Infometric methodological analysis refers to the evaluation of the quality and relevance of scientific research through bibliometric indicators and impact metrics [5]. By examining academic scientific articles, researchers can identify the most influential publications in a specific field, as well as evaluate the quality of research based on the number of citations received, the journal's impact factor, the rejection rate, among other indicators.

Automatically summarizing scientific articles would help researchers in their research by speeding up the research process. Automatic summarization of scientific articles differs from summarization of generic texts by its specific structure and the inclusion of citation phrases. Most of the valuable information in scientific articles is presented in tables, figures, and pseudocode of algorithms. These elements, however, do not typically appear in generic text. One of the technologies available for this purpose is information extraction (IE), which has been widely applied in different domains to date [6].

Information extraction (IE) and natural language processing (NLP) are interrelated fields that transform unstructured data into useful, structured information. IE involves a range of tasks, such as tokenization, syntactic parsing, entity identification, and relationship extraction, which require natural language understanding. This is where NLP provides the necessary tools and techniques, utilizing algorithms and statistical models [7].

In this context, we have analyzed the integration of Neutrosophic Logic, proposed by [8, 9], which expands the concept of truth to include indeterminacy, thereby addressing the inherent ambiguities of human language. For example, in NLP, many words can have multiple meanings, and neutrosophic logic helps capture these ambiguities by considering different degrees of truth. This also applies to sentiment analysis, where it allows for a more nuanced assessment than categorizing a sentiment as simply positive or negative.

In IE, context often affects the meaning of information, which is why Neutrosophic Logic has been considered to offer flexibility in classifying data relevance, recognizing that information may not be completely relevant or may be subject to interpretation. Therefore, integrating Neutrosophic Logic into information extraction and natural language processing not only improves human language understanding but also enables greater accuracy in information interpretation, which is key in a world where ambiguity and subjectivity are the norm in communication.

Integrating information extraction and natural language processing within a neutrosophic framework can be highly beneficial in a variety of applications, such as information retrieval, text mining, semantic search, business intelligence, knowledge extraction, and others. By combining these disciplines, the extraction of useful knowledge from large amounts of unstructured data can be automated, leading to more informed and efficient decision-making in various fields, especially in trend analysis, which is the focus of this research. Based on the potential of the techniques and tools described above, this research aims to develop a neutrosophic computational model based on Natural Language Processing techniques for trend analysis in scientific articles.

2. Materials and Methods

The Neutrosophic Computational Model for identifying trends in scientific articles using Natural Language Processing is designed to address the extraction and generation of summaries of scientific articles, optimizing both textual coherence and information relevance. This model is structured in several fundamental stages: first, the text is divided into segments using coherence scores, ensuring that related sentences are appropriately grouped and minimizing interruptions at inappropriate points.

Next, in the summary generation phase, the model uses the Transformer architecture to produce a concise summary of each segment, conditional on the context of the previous fragments. This summary is generated sequentially, allowing each word to be based on previously written content and the context of its fragments. The model then expands these summaries into full texts, generating the constituent words of each segment in the same sequential manner.

Model training involves optimizing the negative likelihood loss to improve the quality of the generated summaries and texts. During the inference phase, a beam search strategy is employed, complemented by mutual information reclassification to improve the accuracy of the generated sentences. Overall, this model enables a more efficient and coherent understanding of trends in the scientific literature, facilitating the extraction of relevant knowledge from large volumes of unstructured data. Figure 1 shows the overall structure of the proposed model:

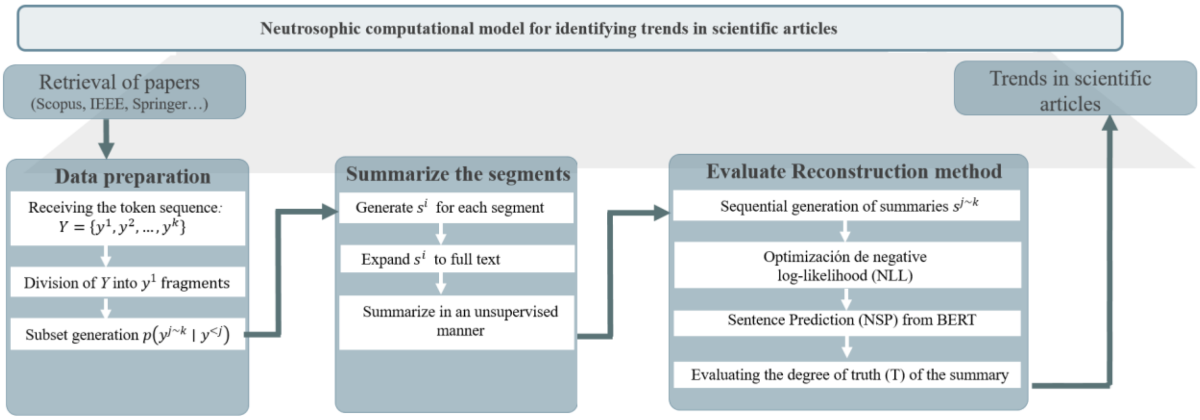

Figure 1. Structure of the Neutrosophic Computational Model for identifying trends in scientific articles using NLP.

Model Design

A long sequence of tokens $Y = \{y^1, y^2, \dots, y^k\}$ is divided into a series of fragments $y^i s$ where k denotes the number of constituent fragments. Bold font **y** was used to denote fragments and regular font y was used to denote tokens. The number of tokens N within each fragment is a hyperparameter. We also use the superscript$i$ to denote the index of a chunk and the subscript l to denote the index of a token. Each y^i consists of a sequence of tokens $y^i = \{y_1^i, \dots, y_{n_i}^i\}$, where n_i denotes the length of y_i. The goal is to generate a subset of Y, denoted as $y^{j\sim k} = \{y^j, y^{j+1}, \dots, y^k\}$ given its preceding chunks, denoted by $p(y^{j\sim k} \mid y^{<j})$. Each fragment y^i s associated with a summary short $s^i = \{s_1^i, s_2^i, \dots, s_{m_i}^i\}$, where s_l^i denotes the tokens and m_i is the number of tokens in s^i.

Instead of generating all the constituent words in Y one by one, a hierarchical strategy was adopted, as proposed by [10]. The generation process of $y^{j\sim k}$ is decoupled into the following two activities:

- 1: Summarizing the fragments: The summary s^i , is generated sequentially for each fragment, conditioning it on the summaries of the previous fragments. This mimics the catalog generation process when humans type.
- 2: Expanding summaries into texts: Each summary s^i is expanded to the entire segment by sequentially generating its constituent words.

In the training phase, the model learns to generate summaries. Unsupervised Extractive Summarization was used [11]. For each fragment y^i, its summarys^i is extracted in an unsupervised manner, and the extracted s^i is used as the optimal summary for training. The Reconstruction method [12] was also used to evaluate the importance of selecting summary sentences.

To measure the degree of sentence reconstruction skill, a seq2seq model was used to predict the original text given the summary sentence [13], whose probability is considered the reconstruction score. The reconstruction score for y^i is denoted by $Score(y^{i,j})$ calculated according to equation 1:

$$Score(y^{i,j}) = \frac{1}{|y^i|} \log p(y^i \mid y^{i,j}) \quad (1)$$

To obtain $p(y^i \mid y^{i,j})$, another seq2seq model is trained, where the input is $y^{i,j}$ for each j and the output is y^i sequentially predicting the tokens in y^i. Given the trained model, all sentences in y^i are classified, and the one with the highest score is used as the optimal summary s^i.

In the summary generation stage $s^{j\sim k}$ is sequentially generated, given $y^{<j}$:

$$p(s^{j\sim k} \mid y^{<j}) = \prod_{i\in[j,k]} p(s^i \mid y^{<i}, s^{<i}) \tag{2}$$

The generation of thes^i summary can be factored into the sequential generation of the constituent words within the summary itself:

$$p(s^i \mid y^{<i}, s^{<i}) = \prod_{l\in[1,m_i]} p\left(s_l^i \mid s_{<l}^i, y^{<i}, s^{<i}\right) \tag{3}$$

This process ends when a Special End-of-Sequence (EOS) token is generated or a specified digest length m is reached.

Each digest s^i is then expanded to the full text of each segment by sequentially generating its constituent words. The expansion process has the same termination conditions as in digest generation.

$$p(y^i \mid y^{<i}, s^i) = \prod_{l\in[1,n_i]} p\left(y_l^i \mid y_{<l}^i, s^i, s^{<i}\right) \tag{4}$$

For summary generation, the Transformer model [14] takes $[y^{<i}, s^{<i}]$ as input and is optimized by minimizing the Negative Likelihood Loss (NLL) [15]:

$$loss - \log\ p(s^i \mid y^{<i}, s^{<i}) \tag{5}$$

The tokens in $y^{<i}$ mostly come from the segment just before it, i.e., y^{i-1} , while $s^{<i}$ comes from multiple previous segments since the summary is more concise. For the summary expansion stage, the Transformer model takes $[y^{<i}, s^i]$ s input and is optimized by minimizing the NLL as shown in equation (6).

$$loss - \log\ p(\widehat{y^i} \mid y^{<i}, s^{<i}) \tag{6}$$

The two models, i.e., summary generation and summary expansion, share parameters, with a task-specific token added at the start to inform the model about what to generate: summaries or segments. In the case of segments, the simplest way is to split the full text equally. However, this is suboptimal, as the cut point could be between two closely related sentences, and a segment could contain multiple aspects.

In this case, splitting based on sentence-level coherence scores is employed. Using the Next Sentence Prediction (NSP) of the BERT model [16], the coherence score between two consecutive sentences with indices i and $i+1$ can be measured, denoted as$Score(i, i+1)$. Given a full text $y = y(1), y(2), \ldots, y(B)$, where B denotes the total number of sentences in y and $y(i)$, denotes the i-th sentence. Given a fixed value k for the number of segments to be divided, y will be divided into k segments, that is, $y^1, y^2, \ldots, y^K$, where each y^k consists of a group of consecutive sentences of y.

Let G_k be the list of sentence indices in the original tex y, where $G_k[1]$ denotes the index of the first sentence in $G_k, G_k[2]$ denotes the index of the second sentence, and so on. Where $R_k = |G_k|$ e is the number of sentences in G_k. The objective is to maximize the coherence scores between two consecutive sentences within the same segment and minimize the score between two consecutive sentences belonging to different segments. This gives rise to the following objective to be optimized:

$$L = \sum_{k=1}^{K} \sum_{i\in[1,R_k-1]} Score(G(k)[i], G(k)[i+1]) - \sum_{k=1}^{k-1} Score\ (G[k][R_k], G[k][1]) \tag{7}$$

Where $Score(G[k][R_k], G[k][1])$ is the coherence score between the last sentence of a segment and the first sentence of the next segment. Given $Score(i,j)$, equation (7) can be solved using linear programming.

Summary Truth (T)

Truth (T), the summary's degree of truth (validity). It is a number in the interval [0,1] that evaluates the summary's truth. Generally, only summaries with a high T value effectively support decision-making. This research proposes using Neutrosophic Logic to evaluate the validity of the generated summary. The proposal of [17], is adopted as an extension of the degree of truth using Neutrosophic Logic:

Neutrosophic set: Let X be a universe of discourse, a neutrosophic set A over X is an object of the form: $A = \{\langle \mu_A(x), \tau_A(x), \sigma_A(x)\rangle : x \in X\}$, and it is true that:

- $\mu_A(x) \in [0,1]$ is a membership function that represents the degree of certainty of the membership of the value x to the set A, see equation 8.
- $\tau_A(x) \in [0,1]$ membership function that represents the degree of indeterminacy of the value x to the set A, see equation 9.
- $\sigma_A(x) \in [0,1]$ membership function that represents the degree of non-membership (or falsity) of the value x to the set A, see equation 10.
- $0 \le \mu A(x) + \tau A(x) + \sigma A(x) \le 3 \;\; \forall x \in X$

$$\mu_A(x) = \begin{cases} (x-a)u_A/(b-a) & (a \le x < b) \\ u_A & (x = b) \\ (c-x)u_A/(c-b) & (b < x \le c) \\ 0 & In\ another\ case \end{cases} \tag{8}$$

$$\tau_A(x) = \begin{cases} \big(b-x+v_A(x-a)\big)/(b-a) & (a \le x < b) \\ r_A & (x = b) \\ \big(x-b+v_A(c-x)\big)/(c-b) & (b < x \le c) \\ 1 & In\ another\ case \end{cases} \tag{9}$$

$$\sigma_A(x) = \begin{cases} \big(b-x+f_A(x-a)\big)/(b-a) & (a \le x < b) \\ f_A & (x = b) \\ \big(x-b+f_A(c-x)\big)/(c-b) & (b < x \le c) \\ 1 & In\ another\ case \end{cases} \tag{10}$$

Let $y_i \in Y$ be such that, Y represents the database of n objects $|Y| = n$ and each object y_i is made up of k attributes. For each attribute z of an object y_i there exists a linguistic variable $V_z = \{A_1, A_2, \ldots, A_g\}$ made up of "g" fuzzy sets.

Let V_z be a linguistic variable associated with the z-th attribute of the objects in the database Y; the membership of an object $y_i \in Y$, in the fuzzy set $A \in V_z$, is denoted by $\mu_A(y_i) \in [0,1]$ and evaluates the z-th attribute of object y_i in the fuzzy set A.

Using neutrosophic logic, we seek to propose a more focused indicator for assessing summary certainty, which is better able to discriminate summaries and objects with a higher degree of certainty from those that lack it.

$$T = \mu_Q\left(\frac{s^i_{\alpha*}(y)}{s^*(y)}\right), T \epsilon [0,1] \tag{11}$$

Where, μ_Q is the fuzzy set representing the summary quantifier.

Hardware resources

To implement the neutrosophic computational model, we chose a server with moderate specifications that, while not state-of-the-art, offered sufficient resources to carry out model development and validation. The operating system was Ubuntu 16.04, equipped with a quad-core Intel Core i7 processor and 16GB of RAM. An NVIDIA GeForce GTX 1050 graphics card, which, while not high-end, enabled the basic parallel processing required for model training tasks. This configuration was capable of handling small to medium-sized datasets, allowing for initial testing and fine-tuning of the model.

Although hardware limitations prevented work on large volumes of data and complex processing tasks, the approach adopted was sufficient to perform preliminary validation of the model's capabilities and initial experimentation in identifying trends in scientific articles. This experience underscores the importance of model optimization and preprocessing techniques, which made it possible to maximize the performance of the available hardware.

3. Results

To conduct a comprehensive case study demonstrating the applicability and implementation of the Neutrosophic Computational Model for identifying trends in scientific articles using Natural Language Processing, we will approach the process in several steps, including data preparation, abstract generation, expansion of those abstracts to full texts, and assessment of the validity of the generated content. This case study will focus on identifying trends in a set of scientific articles related to a specific topic, in this case: Artificial Intelligence.

Data Preparation

Corpus Selection: For this case study, we will select a set of scientific articles on Artificial Intelligence from sources such as Scopus, IEEE, or Google Scholar. We will assume we have extracted the text of 5 relevant articles. Example Articles: Let's assume the extracted texts are the following:

- Article 1: "Deep learning algorithms have revolutionized the way artificial intelligence systems are built."
- Article 2: "Reinforcement learning approaches are enabling significant advances in automated decision-making."
- Article 3: "Convolutional neural networks are especially effective for image processing."
- Article 4: "Ethics in artificial intelligence is becoming a topic of growing concern among researchers."
- Article 5: "Natural language processing has improved thanks to models such as BERT and GPT-3."

Token Format: From these texts, we will proceed to tokenize, that is, divide the text into sequences of tokens $Y = \{y^1, y^2, \dots, y^k\}$.

Division into segments

Fragmentation: We will use a value of $N = 2$ or the maximum token length per fragments. The 5 articles will be divided into 3 fragments:

- $y^1 = \{y_1^1, y_2^1\}$ = Article 1 + Article 2
- $y^2 = \{y_3^2, y_4^2\}$= Article 3 + Article 4
- $y^3 = \{y_5^3\}$ = Article 5

Here denotes the number of tokens in y^1.

Summary Generation

Summarize the snippets: For each snippet, we generate a summary sequentially:

- Fragment 1**:** Generate s^1 for y^1 (We will use extracted terms in the process):
  - Extracted summary (s^1) "Deep learning algorithms are revolutionizing artificial intelligence."
- Fragment 2: Generate s^2 for y^2.
  - Extracted summary (s^2): "Convolutional neural networks are effective in images, and ethics in AI are concerning."
- Fragment 3: Generate s^3 for y^3:
  - Extracted summary (s^3): "Models like BERT have improved Natural Language Processing."

Expansion of summaries to texts

Expand s^i to Text: We expand each s^i summary to its full text:

- y^1 : Expand s^1 → "Deep learning algorithms have revolutionized the way artificial intelligence systems are built. Reinforcement learning approaches are enabling significant advances in automated decision making."
- y^3 : Expand s^2 → "Convolutional neural networks are especially effective for image processing. Ethics in artificial intelligence is becoming a topic of growing concern among researchers."
- y^5: Expand s^3 → "Natural language processing has improved thanks to models like BERT and GPT-3."

Summary Validity Assessment

Calculating the degree of truth (T): The validity of the summary will be assessed using Neutrosophic Logic. We define the degree of truth (T) of the generated summary as a number in the range [0, 1], where 0 means the summary is invalid and 1 means the summary is completely valid. To calculate the degree of truth, we use the reconstruction score defined in Equation 1.

$$Score(y^{i,j}) = \frac{1}{|y^i|}\log p(y^i \mid y^{i,j}) \tag{12}$$

Where $y^{i,j}$ is the input summary whose sentences we are evaluating.

Running the Evaluation:

1. For each fragment y^i, we will extract the sentences from the summary and use them to make predictions about the original text using a pre-trained seq2seq model to obtain $p(y^i \mid y^{i,j})$, which refers to the probability of reconstructing the original text given the summary.
2. The following probability values were obtained for our examples:
   - $p(y^1 \mid s^1)$=0.8
   - $p(y^2 \mid s^2)$=0.7
   - $p(y^3 \mid s^3)$=0.9
3. We used the values to calculate the reconstruction score:

For the first fragment:

$$Score(y^{1,j}) = \frac{1}{|y^i|}\log(0.8) \approx -0.223$$

For the second fragment:

$$Score(y^{2,j}) = \frac{1}{|y^2|}\log(0.7) \approx -0.357$$

For the third fragment:

$$Score(y^{3,j}) = \frac{1}{|y^3|}\log(0.9) \approx -0.105$$

With these scores, we can normalize them and convert them to a [0,1] interval, where higher values are more indicative of the summary's validity.

Trend Identification

Trend Analysis: From summaries s^1, s^2, y s^3, we identify common trends. We will analyze keywords that appear in more than one summary and interrelated concepts.

- Key Terms:
  - "Artificial Intelligence"
  - "Deep Learning"
  - "Neural Networks"
  - "Ethics"
  - "Natural Language Processing"

Trend Consolidation

We can group the emerging trends as follows:

1. Trends in Learning Techniques**:**
   - Increased use of neural networks and deep learning.
   - Implications of reinforcement learning.
2. Ethical Concerns in AI:
   - Ethics in artificial intelligence is becoming an increasingly critical issue in the research community.
3. Progress in Natural Language Processing**:**
   - Significant improvements thanks to models such as BERT and GPT-3**.**

The application of the neutrosophic computational model has allowed for the extraction of effective summaries of the selected articles and has facilitated the identification of relevant trends in the field of artificial intelligence. The observed trends suggest an increasing focus on technical and ethical approaches to the development of artificial intelligence, reflecting a commitment to continuous technical improvement and respect for the social impacts of these technologies.

4. Metrics

Scoring Accuracy: To measure the proximity of predictions to actual ratings, we propose using the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics, defined according to equations 12 and 13, respectively. For these metrics, the lower the values, the greater the accuracy of the predictions.

$$MAE = \frac{1}{|E^P|}\sum_{(i,\alpha)\in E^P}|r_{ia} - \hat{r}_{ia}| \quad (13)$$

$$RMSE = \sqrt{\frac{1}{|E^P|}\sum_{(i,\alpha)\in E^P}(r_{ia} - \hat{r}_{ia})^2} \tag{14}$$

Where:

$r_{i\alpha}$: is the actual rating value of user i on item α,

$\hat{r}_{ia}$: is the value predicted by the system for that rating,

E^P: subset that will be compared with the corresponding predictions to evaluate the effectiveness of the summaries.

To identify the most relevant summaries for the end user of the neutrosophic computational model proposed in this research, assuming that they are interested in the indicators in the top L positions, it is considered appropriate to use the Precision (P) and Recall (R) metrics. For a user i, equations 14 and 15 are defined respectively.

$$P_i(L) = \frac{d_i(L)}{L} \tag{15}$$

$$R_i(L) = \frac{d_i(L)}{D_i} \tag{16}$$

Where:

$d_i(L)$: indicates the number of relevant indicators among the first L in the summary list,

: is the total number of relevant indicators for user i. Thus, if all the Precision and Recall values for all users are averaged, the average values P(L) and R(L) are obtained.

Satisfaction Metric: In order to evaluate the usefulness of trend analysis for a user, the Half-life Utility (HL(L)) metric is used. This metric is based on the assumption that the probability of a user examining a constructed summary decreases exponentially with the classification of the indicators. The expected utility for the summaries, given to a user i, is defined according to equation (16).

$$HL_i = \sum_{\alpha=1}^{N}\frac{max(r_{ia} - d, 0)}{2^{(0_{i,\alpha}-1)/(h-1)}} \tag{17}$$

Where:

$0_{i,\alpha}$: represents the predicted ranking position for indicator α in the list of summaries for user i,

d: s the default rating (for example, it can be placed at the middle of the possible ratings),

h: is the indicator's position in the list with a 50% probability that the user will ultimately examine it.

In this metric, objects are sorted by their prediction $\hat{r}_{ia}$ in descending order. When HLi is averaged across all users, a utility for the entire system is obtained.

Table 1. Results of the Computational Model Implementation.

User i	*T*	MAE	RMSE	Precisión	Recall	Half-life Utility
1	0.85	0.15	0.20	0.75	0.60	0.65
2	0.90	0.10	0.15	0.80	0.70	0.70
3	0.78	0.20	0.25	0.70	0.65	0.50
4	0.82	0.18	0.23	0.78	0.68	0.60
5	0.88	0.12	0.18	0.85	0.75	0.75
Average	0.83	0.17	0.20	0.78	0.66	0.64

The results presented in the table reveal a positive evaluation of the computational model implemented for identifying trends in scientific articles. The average validity value T is 0.83, indicating that, in general, the generated summaries are considered valid and effectively reflect the original content of the texts. The error metrics, specifically the Mean Absolute Error (MAE) with an average of 0.17 and the Root Mean Squared Error (RMSE) with 0.20, suggest that the predictions made by the model are close to the actual ratings within an acceptable margin, implying good accuracy in reconstructing the texts.

The average precision of 0.78 and the recall of 0.66 indicate that the system manages to identify a significant proportion of relevant recommendations among the highest in the generated lists, confirming its effectiveness in generating useful suggestions for users. Furthermore, the average Half-Life Utility of 0.64 suggests that the recommendations are sufficiently engaging and useful to users, potentially resulting in a high acceptance and utilization rate. Taken together, these results indicate that the model is not only capable of generating coherent and valuable summaries but is also efficient in offering relevant recommendations, demonstrating its ability to identify and reinforce trends in the scientific literature.

As a line of future research, it would be valuable to incorporate neutrosophic stance detection techniques to deepen the analysis of how users position themselves—explicitly or implicitly—regarding scientific trends or recommendations. The approach proposed by Vázquez and Smarandache [18] offers a promising framework to detect ambivalent, uncertain, or contradictory positions in discourse, by integrating truth, indeterminacy, and falsity into stance modeling. This methodology could enhance the interpretability and adaptability of the system, especially when evaluating heterogeneous user feedback or guiding personalized scientific discovery.

5. Conclusions

The results highlight the effectiveness of the neutrosophic computational model as a valuable tool for identifying trends in scientific articles using Natural Language Processing (NLP). First, the model demonstrated its ability to analyze and synthesize large volumes of information, generating accurate and coherent summaries that adequately reflect the original content of the texts. This not only improves the accessibility of scientific information but also facilitates the identification of emerging and relevant research topics.

Second, the model was able to integrate various evaluation techniques, allowing for the validation of the summaries' effectiveness and continuous improvement of the system based on the results obtained.

Additionally, the summaries' relevance and usefulness indicators suggest that the model is attractive and functional for users, contributing to its potential application in academic and professional settings.

This research underscores the importance of computational models in the field of science, highlighting their role in transforming and optimizing access to knowledge, which could lead to more informed research decisions and more rapid progress in various areas of knowledge. The development of this model not only has significant implications for NLP, but also represents a notable advance in the way contemporary scientific literature can be explored and understood.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] G. Halevi, H. Moed, and J. Bar-Ilan, "Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the literature," *Journal of informetrics,* vol. 11, no. 3, pp. 823-834, 2017.

[2] S. Kapoor, and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns,* vol. 4, no. 9, 2023.

[3] B. Bron Fonseca, and O. Mar Cornelio, "Método para el análisis lingüístico de estadísticas médica," *Serie Científica de la Universidad de las Ciencias Informáticas,* vol. 18, no. 1, pp. 110-127, 2025.

[4] D. Dessì, F. Osborne, D. R. Recupero, D. Buscaldi, and E. Motta, "Generating knowledge graphs by employing natural language processing and machine learning techniques within the scholarly domain," *Future Generation Computer Systems,* vol. 116, pp. 253-264, 2021.

[5] P. M. Marcillo Sánchez, "Análisis del desarrollo de software con metodología ágil y la capacidad de la sostenibilidad implementada," ETSI_Sistemas_Infor, 2022.

[6] A. Sharma, and S. Kumar, "Machine learning and ontology-based novel semantic document indexing for information retrieval," *Computers & Industrial Engineering,* vol. 176, pp. 108940, 2023.

[7] M. H. A. Abdullah, N. Aziz, S. J. Abdulkadir, H. S. A. Alhussian, and N. Talpur, "Systematic literature review of information extraction from textual data: recent methods, applications, trends, and challenges," *IEEE Access,* vol. 11, pp. 10535-10562, 2023.

[8] F. Smarandache, "Significado Neutrosófico: Partes comunes de cosas poco comunes y partes poco comunes de cosas comunes," *Serie Científica de la Universidad de las Ciencias Informáticas,* vol. 18, no. 1, pp. 1-14, 2025.

[9] F. Smarandache, "Neutrosofía y Plitogenia: fundamentos y aplicaciones," *Serie Científica de la Universidad de las Ciencias Informáticas,* vol. 17, no. 8, pp. 164-168, 2024.

[10] X. Sun, Z. Sun, Y. Meng, J. Li, and C. Fan, "Summarize, outline, and elaborate: Long-text generation via hierarchical supervision from extractive summaries," *arXiv preprint arXiv:2010.07074*, 2020.

[11] M. Jang, and P. Kang, "Learning-free unsupervised extractive summarization model," *IEEE Access,* vol. 9, pp. 14358-14368, 2021.

[12] X. Mao, H. Yang, S. Huang, Y. Liu, and R. Li, “Extractive summarization using supervised and unsupervised learning,” *Expert systems with applications,* vol. 133, pp. 173-181, 2019.

[13] F. Carichon, F. Fettu, and G. Caporossi, “Unsupervised update summarization of news events,” *Pattern Recognition,* vol. 144, pp. 109839, 2023.

[14] D. Siva Krishna, P. Srikanth, R. Srinivas Rao, and R. HOLLA M, “Data2Summ: Review on Enhancing Summarization with Transformers and Attention Models for Speech and Text Compression,” *Physica Scripta*, 2025.

[15] H. Shakil, A. Farooq, and J. Kalita, “Abstractive text summarization: State of the art, challenges, and improvements,” *Neurocomputing*, pp. 128255, 2024.

[16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding." pp. 4171-4186.

[17] I. Pérez Pupo, P. López Gómez, E. Pérez Varona, P. Piñero Pérez, and R. García Vacacela, “Construcción de resúmenes lingüísticos a partir de rasgos de la personalidad y el desempeño en el desarrollo de software,” *Revista Cubana de Ciencias Informáticas,* vol. 12, pp. 135-150, 2018.

[18] Vázquez, M. Y. L., & Smarandache, F. (2025). Bridging Ancient Wisdom and Modern Logic: Neutrosophic Perspectives on Body, Mind, Soul and Spirit. Neutrosophic Sets and Systems, 82, 108-121.

Received: December 25, 2024. Accepted: April 6, 2025.