International Journal of Computers and Informatics

Journal Homepage: https://www.ijci.zu.edu.eg

Int. j. Comp. Info. Vol. 4 (2024) 31–45

Paper Type: Review Article

A Thorough Survey on the Fusion of Machine Learning Algorithms and Neutrosophic Theory

Samia Mandour [1,*]

[1] Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, Zagazig, Sharqiyah, 44519, Egypt; samia.rmdan@fci.zu.edu.eg.

Received: 20 May 2024 **Revised**: 10 Jul 2024 **Accepted**: 13 Aug 2024 **Published**: 15 Aug 2024

Abstract

Massive, complex, and varied collections of data that are difficult to store, process, and visualize to use in later processes or results are referred to as "big data". Data mining is the process of examining and evaluating enormous volumes of data to look for important patterns and principles. Because data mining can reveal useful patterns that were previously unknown, it is essential to many human efforts. There are a lot of machine learning algorithms that are widely used in this regard. Currently, one of the most effective approaches for addressing the aforementioned problems when combined with machine learning tools is recognized as being neutrosophic set theory (NS). Generalizations of interval fuzzy sets, or neutrosophic set theory (NS), have gained significant attention in the data mining and machine learning sectors during the past ten years due to their multitude of applications. Since natural ambiguity is dealt with by neutrophilic group theory (NS), academics have been motivated to incorporate NS into machine learning algorithms in order to eliminate ambiguity from data, add more precise data values, and improve the accuracy and efficiency of mining methods. A large number of research papers have been published on the hybridization of neutrosophic set theory (NS) and machine learning techniques. This has motivated us to present a review of the literature on the application of NS with machine learning approaches to address data mining challenges in the years 2020–2024.

Keywords: Big Data; Data Mining; Machine Leaning; Interval Fuzzy Sets; Neutrosophic Set.

1 | Introduction

Our lives are digitally gathered, and we are living in the era of data when every detail is linked to a database[1]. In the modern electronic world, for example, there is an abundance of diverse types of data, including data from the Internet of Things (IoT), cybersecurity, smart cities, enterprises, cellphones, social networking sites, medical care, COVID-19, and numerous additional sources. Varieties of machine learning methods and real-world data are growing daily. By concluding these data, numerous smart applications in the pertinent fields can be constructed. For example, appropriate cybersecurity data can be utilized to develop an intelligent and autonomous cybersecurity platform powered by data [2]; similarly, relevant smartphone information may be utilized to develop context-aware, tailored smart smartphone apps; and so on [3]. Therefore, applications in real life that rely on managing data technologies and methodologies capable of intelligently and quickly extracting knowledge or valuable information from data are desperately needed.

Corresponding Author: samia.rmdan@fci.zu.edu.eg

In the last few years, AI and machine learning (ML) have evolved significantly in the realms of analysis of data and computation, which usually enables the applications to operate intelligently. ML, often known as the most well-liked newest technologies in the so-called fourth industrial revolution, often gives systems the capacity to learn and improve from experience autonomously without being specifically designed. "Industry 4.0" generally refers to the continuous automation of traditional industrial processes and manufacturing, involving exploring data processing, through the use of new intelligent technologies like automation for machine learning. Therefore, machine learning techniques are essential for both intelligently analyzing these data and creating the related applications that exist in reality [4]. Four primary categories of learning algorithms may be distinguished in the field: supervised, unsupervised, semi-supervised, and reinforcement learning [5], this categorization can be displayed in Figure 1.

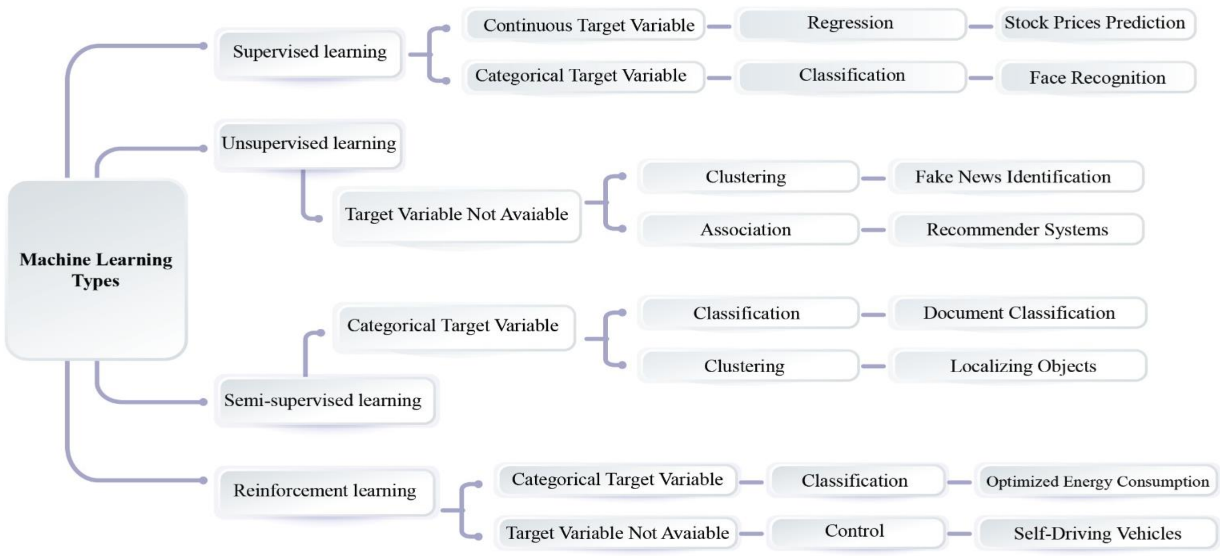

Figure 1. Machine learning algorithms.

These methods of learning are becoming more and more popular every day. Generally speaking, the efficacy and efficiency of a solution that uses machine learning rely on the type and properties of the data as well as the functionality of the learning methods. To effectively create data-driven structures, algorithms for machine learning can be used in classification analysis, regression, data grouping, feature engineering and reduction of dimensionality, rule-based association learning, or learning by reinforcement [6]. Additionally, deep learning is derived from artificially generated neural networks, which belong to a larger branch of machine learning techniques and are capable of doing intelligent data analysis [7]. It is therefore difficult to choose an appropriate learning algorithm that fits the intended application in a given domain. The explanation for this is that different learning algorithms have different goals, and even within the same category, different learning algorithms may produce different results based on the properties of the data [8]. To utilize algorithms for machine learning in a variety of real-world domains of application, including cybersecurity services, IoT systems, business and recommendation systems, cities that are smart, medical care and COVID-19, context-aware structures, agricultural sustainability, and many more, it is crucial to comprehend the underlying principles of these algorithms. There are a lot of ML algorithms and techniques that are used to solve data mining duties. So, there are a lot of ML algorithms that can be used for handling the problems of data mining, besides the well-known data mining algorithms, which can be classified as shown in Figure 2.

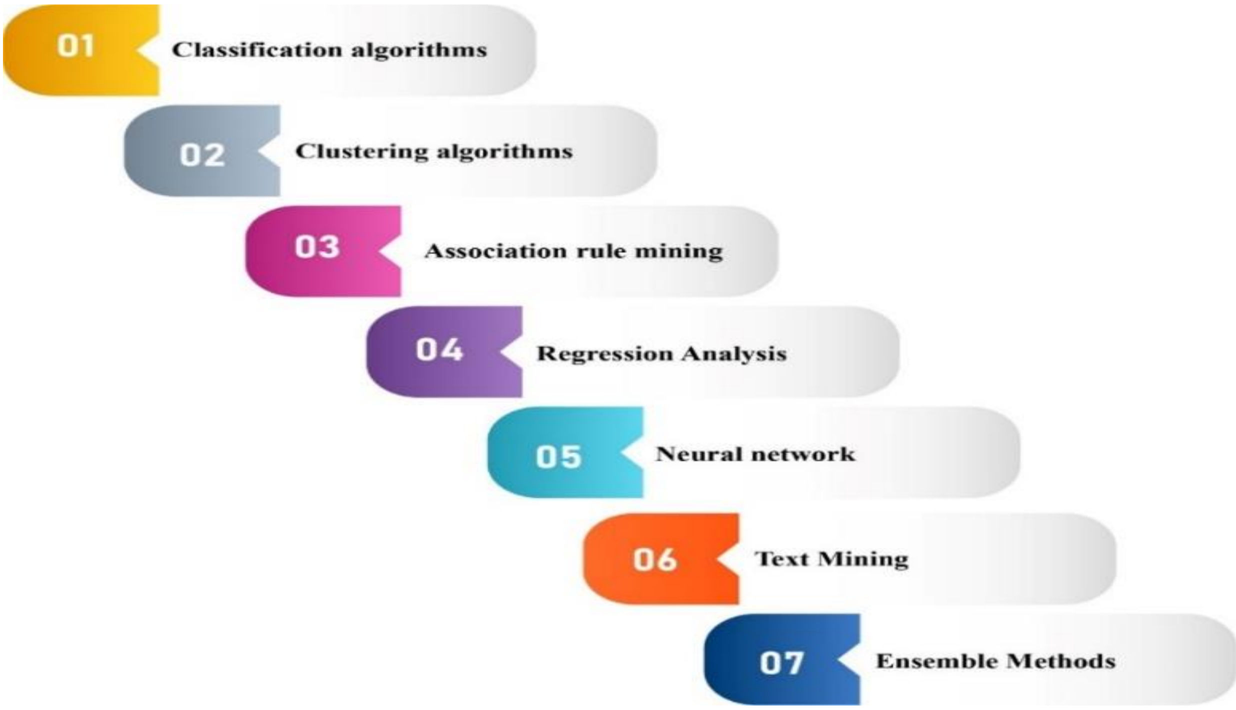

Figure 2. Data mining algorithms.

The ML algorithms and data mining algorithms have been spread across a wide range of fields such as decision-making [9, 10], medical sector [11-16], agriculture [17-20], financial [21-23], engineering [24-26], and other sectors. Although ML algorithms have made great strides in resolving real-world problems and data mining issues, the ambiguity surrounding the data in data mining problems and many other problems continues to stand in the way of finding the optimal answer, leading academics to turn to neutrosophic science as a workaround.

To represent the various types of uncertainty, Smarandache offers a domain area composed of three distinct subsets, which are known as neutrosophic sets. The definition of neutrosophic sets is those sets in which each element of the cosmos has a score of certainty, uncertainty, and indeterminacy that ranges from 0 to 1 [27]. The indeterminacy concept distinguishes between absoluteness and relativeness and expresses degrees of belonging and non-belonging to the neutrosophic sets, whereas impreciseness is expressed by truth and falsity values. This notation enables neutrosophic sets to handle the unpredictable nature of the system and reduce the paralysis caused by contradicting information. Consequently, compared to the numerous other types of fuzzy extensions, one could contend that this capability represents the most important advantage provided by neutrosophic sets [28].

Through the utilization of truth, falseness, and ambiguity values, domain areas can be established using neutrosophic sets. This field enables the independent execution of different types of computational tasks in the face of uncertainty. Researchers have demonstrated a great deal of interest in NSs. NSs have garnered significant interest from researchers due to their capacity to manage uncertainty [29-32]. Neutrosophic logic has proven to be an essential technique for eliminating ambiguity, according to the preceding. It has thus been applied and modified by many academics to address real-world problems and data mining challenges. So, in this article, we present a thorough view of several forms of algorithms for machine learning incorporated with neutrosophic set theory that may be implemented to improve the intelligence and capacities of an application, based on the significance and potential of "Machine Learning " to evaluate the data indicated above in the period from 2019 to 2024.

The following is a summary of the main contributions made by this paper:

- Define the area of our research by considering the nature and properties of different real-world data kinds and the capacities of different learning techniques.
- To offer a thorough understanding of machine learning techniques that can be used to improve a fueled by data application's intellect and capabilities.

This is how the remainder of the paper is structured. Neutrosophic set theory is discussed in the second section. The categories of data and machine learning methods are defined and presented in a more generalized manner in the third section. In the fourth part that follows, we review and explain the integration between machine learning techniques and neutrosophic logic in addressing real-world problems. We outline several research questions and possible future paths during the penultimate part, and this study is concluded in the last section.

2 | Neutrosophic Theory

Numerous pieces of information generated by applications that mimic the actual world are inadequate, erroneous, unclear, and inconsistent. There's a potential that the uncertainty results from things like incomplete data, errors made when gathering data, or chance. To deal with such ambiguous data, many contemporary theories and methods have been proposed, including intuitionistic fuzzy set (IFS), fuzzy set (FS) theory, statistical and probability theories, and para-consistent logic theory. Nevertheless, instead of addressing the complete problem under a unified framework, these speculations are limited to addressing a single imprecise component of a problem. For example, the FS cannot handle any gaps or irregularities in the data it is processing; it can only handle vague and hazy data. To resolve such issues within a single structure, it is thus advised to apply the neutrosophy technique, a school of philosophy that integrates the concepts of logic, statistical methods, probability, and set hypotheses. Smarandache initially presented a neutrosophic concept at the close of the 1990s [33]. The neutrosophic set is considered the generalization of the fuzzy set [34]. Neutrosophic logic, which represents indeterminacy by employing an innovative framework known as "Neut-A" to handle some problems that fuzzy logic is unable to resolve, is based on the principles of neutrosophy. Fuzzy logic is typically best described as a form of two-valued logical reasoning, where statements can possess a truth degree ranging from 0 to 1 rather than having to be true or false. IF logic and neutrosophic logic frequently exhibit higher levels of "uncertainty" than any other logic [35].

Despite this, the neutrosophic reasoning allows each element (T, I, and F) to be equally over flooded (simmering) throughout 1, or "1+," or under dried under 0 (freezing), or "0," to distinguish between relative reality and real truth as well as between relative falsehood and false. This is due to unintentional variables that may be concealed through specific unknowingness or suggestions [36]. Smarandache states that "neutrosophy is a new field of philosophy that explores the genesis, nature, and extent of neutralities, in addition to their connections with other ideational spectra." As an illustration of varied value logic, the model of neutrosophy takes into account fuzzy logic, imperfect likelihood, and classical logic. Neutrosophy is more in line with human reason since it clarifies the ambiguity of the evidence or the linguistic inaccuracy that has been verified by multiple witnesses. In general, the term " neutrosophy " relates to the investigation of neutralities' nature, source, and extent as well as the way they interact. The neutrosophic set represents a single facet of neutrosophy. There is a degree of truth, a degree of untruth, and a degree of ambiguity associated with each occurrence that the neutrosophy theory describes; these three degrees ought to be assessed independently of one another. In recent years, the neutrosophic set has become recognized as a comprehensive, ideal, and suitable structure [37].

Every proposition is simulated in the neutrosophic theory to determine the degrees of truth. $T_n(x)$ ambiguity $A_n(x)$ and falseness $F_n(x)$. The theory of neutrosophy is an extension of rational thought to include intuitionistic fuzzy sets and fuzziness. As a response to several real-world issues involving ambiguity, inaccuracy, vague, inadequacy inconsistency, and indeterminacy, neutrosophic theory has gained traction. Information that is highly erratic and unclear is handled by neutrosophic reasoning. Consequently, the neutrosophic theory is applied to indeterminacy-related problems in many other domains. We want to havespecific notions for defining the neutrosophic variable in order to cope with ambiguity.

According to the mathematical neutrosophic set concept, each element x in the neutrosophic set x has three membership degrees as mentioned above, these memberships fall under the following constraints:

$$0^{-} \leq T_n(x), A_n(x), F_n(x) \leq 1^{+} \quad (1)$$

$$0^{-} \leq T_n(x) + A_n(x) + F_n(x) \leq 3^{+} \quad (2)$$

Equations limiting a type 1 of theneutrosophic fuzzy set are as follows:

$$0^{-} \leq T_n(x), A_n(x), F_n(x) \leq 1^{+} \quad (3)$$

$$T_n(x) \wedge A_n(x) \wedge F_n(x) \leq 0.5 \quad (4)$$

$$0^{-} \leq T_n(x) + A_n(x) + F_n(x) \leq 3^{+} \quad (5)$$

The following formula forces the third categorization, a neutrosophic intuitionistic set of type 2:

$$0.5 \leq T_n(x), A_n(x), F_n(x) \quad (6)$$

$$T_n(x) \wedge A_n(x) \leq 0.5,\ T_n(x) \wedge F_n(x) \leq 0.5,\ A_n(x) \wedge F_n(x) \leq 0.5 \quad (7)$$

$$0^{-} \leq T_n(x) + A_n(x) + F_n(x) \leq 2^{+} \quad (8)$$

3 | Machine Learning Algorithms Types

Computer algorithms' study that enables machines to automatically pick up new skills and get better with time is known as machine learning. Most people consider it to be a branch of artificial intelligence. Algorithms for machine learning enable the systems to make judgments on their own without assistance from outside sources. These choices are determined by sifting through complex data to identify important underlying patterns.

There are four main types of machine learning algorithmic methods: supervised, unsupervised, reinforcement learning, and a hybrid algorithmic method called semi-supervised learning. These types of algorithms are determined by the learning procedure, the kind of data that serves as input and output, along with the kind of issue that is solved. Some hybrid strategies and other widely used techniques provide a natural extension of machine learning issue formats [38].

In cases where the data are output goal values and variables are inputs, supervised learning can be used. Through the input to the output, the function that maps it is learned by the algorithm. For applications when data is sparse, it is an expensive strategy due to the accessibility of massively labeled data samples [39]. When there are no related output variables and the data is solely provided as an input, unsupervised learning can be used [40]. To understand further the properties of the data, these methods represent the underlying patterns in it. When making a series of decisions that lead to a final incentive, reinforcement learning is used [41]. An intelligent agent receives incentives or punishments for the tasks it completes throughout the learning cycle. Its objective is to optimize the overall gain. The approach of semi-supervised learning lies in the middle between supervised and unsupervised learning methods. Labeled and unlabeled data are used in conjunction to train these types of algorithms. A typical scenario has a very high proportion of unidentified data and just a small amount of data that is labeled. Essentially, an unsupervised learning technique is employed for clustering comparable data first, after which the remaining unlabeled data is classified using the data with labels that have already been collected [42].

4 | Neutrosophic Logic-Based Machine Learning Algorithms for Tackling the Problems of the Real World

The digital world is rife with data in this Fourth Industrial Revolution (4IR) or Industry 4.0 era. Examples of this data include cybersecurity, handheld devices, social networking sites, company operations, Internet of Things (IoT), and healthcare data. The key to developing intelligent analyses of these data and creating correspondingly clever and automated applications is understanding artificial intelligence (AI), and specifically machine learning (ML). There are many different kinds of machine learning techniques in the field, including supervised, unsupervised, semi-supervised, and reinforcement learning as shown previously in Figure 1. Furthermore, deep learning, a subset of a larger class of machine learning techniques, is capable of large-scale, intelligent data analysis. Therefore, the fundamental contribution of this study is to shed light on the concepts of different methods of the integration between machine learning and neutrosophic logic and how they may be applied in a variety of real-world domains of application, including healthcare, e-commerce, the agricultural sector, cybersecurity structures, intelligent cities, and many more. This is because, It is acknowledged that consistency, ambiguity, and inconsistency pose complicated issues in the data processing phase when it comes to veracity in the analysis of large amounts of data. A strong ability to model this complicated information has been demonstrated by Neutrosophic logic. In the phase of preprocessing, numerous ML and data mining methods have been suggested to cope with this type of dirty data. Nevertheless, very few research addresses the inaccurate and inconsistent data that is present during the modeling phase. This study, however, provides an overview of all the research done on the translation of machine learning methods from crisp numerical space to the neutrosophic environment. Additionally, we go over the incorporation and contributions made by machine learning methods with Neutrosophic logic to the modeling of imperfect information, as well as the effects of these models on solving data mining problems and real-world situations.

Scholars hurriedly compiled a large number of research publications, including neutrosophic logic-based machine learning algorithms in the field of data science. Thus, it inspired us to begin this survey paper. Some of the research publications on neutrosophic-based machine-learning algorithms from 2019 to 2024 are mentioned in Table 1.

Table 1. An overview of some neutrosophic-based machine learning algorithms and data mining algorithms for tackling real-world problems.

No	Ref	Methodology	Dataset	Advantages	shortcomings
1	[43]	This work introduces a new combination of localization-based decision (DLBD) alongside fuzzy weighting using zero inconsistency (FWZIC) for evaluating hybrid multi-deep transfer and machine learning (HMDTML) frameworks of chest X-ray for COVID-19 case diagnosis in a probabilistic single-valued neutrosophic hesitant fuzzy set (PSVNHFS) environment.	Two datasets are used: the first dataset has 660 CXR scans, among which 468 show patients who are COVID-19 positive, while the rest show patients who are COVID-19 negative (including those with MERS, SARS, and ARDS). A total of 5679 CXR scans, comprising 1406 instances for normal images and 4273 samples for pneumonia images, make up the second dataset.	The model is sturdy and has high reliability.	First, only one fuzzy method (PSVNHS) can be applied to convert the EDM and dynamic localization decision matrix into a PSVNHS-EDM. Secondly, it was decided that the use of FS with DBLD was required because the data was ambiguous and imprecise. The problem of lacking data in a decision matrix about target values has not been considered.

2	[44]	To overcome the shortcomings of decision-making techniques when dealing with uncertain information, this study presents a unique decision-making structure that combines Neutrosophic Triplets (NTs) with the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). The benefits of the suggested approach are illustrated by using it in the crucial field of sustainable supply chain management—green supplier selection.	A large dataset comprising more than 400 check values.	The outcomes of the comparison analysis show how the suggested smart TOPSIS approach has a high degree of accuracy and a low level of computing complexity. The research also sheds light on how long various approaches take to execute, highlighting the Smart TOPSIS method being particularly noteworthy for having a much lower time level of complexity.	The presumption that the suggested intelligent technique will always work well in a variety of decision-making scenarios is the source of this study's weaknesses. This calls for careful examination and validation throughout many diverse domains of application.
3	[45]	The multi-polar interval-valued neutrosophic set and the hypersoft set can be combined, as this article demonstrates. You can use this to address DM issues with several attributes. Furthermore, for multipolar interval-valued neutrosophic hypersoft sets (mPIVNHSs), we construct metrics of similarity. When choosing locations for new stores, K-Nearest Neighbor (KNN), a machine learning technique, is also used to determine ranking. They are useful in everyday life because they leverage the prospect of each data point resemblance.	This model is tested on a set of shopping sites.	This method performs well with the hybrid models that incorporate several decision-making processes, adapt the framework to accommodate for shifting data types and uncertainties, and increase computing efficiency.	Not mentioned
4	[46]	In this research, the authors provide a model of trust that evaluates the confidence score of devices connected to the industrial Internet of Things (IIoT) using the neutrosophic weighted product methodology (WPM), which is employed by the industrial Internet of Things IIoT apps. Depending on the behavioral patterns,	The authors do simulations of various IIoT device properties using five correlated sensors. A dataset for 100 industrial Internet of Things devices has been created.	The findings of the simulation indicate that the suggested trust model is more accurate at identifying IIoT device misbehavior. The accuracy of the suggested neutrosophic	Addressing the scalability issues in the industrial trust computation is not handled.

		temporal experiences, and spatial expertise gleaned from IIoT devices, the proposed model evaluates the reliability of devices. Additionally, the framework suggests using neutrosophic support vector machines (SVM) along with K-NN clustering to classify the retrieved attributes. The final trust score is generated by the suggested neutrosophic SVM algorithm, which can also accurately detect the trust borders.		SVM algorithm is 100%.	
5	[47]	This work presents a novel approach to histogram equalization for picture enhancement, and it suggests a computer-aided diagnosis system for the identification of a malignant cancer skin lesion. From this point on, the pre-processed picture proceeds through the segmentation stage, when the Neutrosophic approach is used for dividing the suspected lesion. To create a segmentation masking of the suspicious skin lesion, the process of segmentation combines a pentagonal neutrosophic framework with a thresholding-based technique. The deep neural network proposed in this research is based on Inception and residual blocks, which includes a soft-max block placed after each residual blocking to widen the layer and speed up the process of learning the important features.	The effectiveness of the suggested technique for digital artifact removal for the PH2, ISIC 2017, ISIC 2018, and ISIC 2019 datasets, respectively.	The suggested strategies appeared to yield good sensitivity and specificity ratings for the model.	There is a lack of attention to a wide range of categories for different skin conditions, including actinic keratoses, basal cell carcinoma, squamous cell carcinoma, vitiligo, psoriasis, atopic dermatitis, and lamellar ichthyosis.
6	[48]	This research produces a trustworthy placing of hospital constructing assets depending upon their variables significance levels as well as deficiencies by combining Neutrosophic reasoning, the Analytic Network Process (ANP), and Multi-Attribute Utility Theory (MAUT). This reduces the subjectivity associated with expert-	The suggested model was used on medical institutions in Canada. Using all pertinent datasets related to the 394 distinct asset categories found in five hospital structures in the state of Alberta, the approach was utilized for each and every data point in order to determine	A consistent, objective, and automated priority system for hospital asset renewals is anticipated to be created with the help of the consolidated structure that has been built. This, in	When it comes to correctly recognizing the high priority cases, the suggested approach is no more accurate than models found in the literature.

		driven choices. This method is further combined with the cutting-edge application of machine learning techniques in this domain, including Decision Trees, K-Nearest Neighbors, and Naïve Bayes, to streamline and make repeatable the procedure of selecting priorities, hence reducing the requirement to seek additional expert opinions. The created concept was used with medical facilities in Canada.	each asset's unique CI and PDI.	consequently, will hopefully lead to an allocation of resources process that is effective, sound, and well-informed.	
7	[49]	This study employed three methods for predicting heart illness: machine learning structures, association rules, and the neutrosophic analytical hierarchy process (AHP) to be a feature selection mechanism. The process of determining feature weights and choosing the best features involves the usage of the neutrosophic AHP approach. All data sets' outcomes are governed by rules that are provided by the association rules. The optimal feature to enter into machine learning algorithms was then chosen using the neutrosophic AHP to be a feature selection method. For predicting heart attacks, we employed nine machine learning algorithms.	The dataset was gathered as a component of a continuing cardiovascular study involving residents of the Massachusetts city of Framingham. You can find the patient details in the dataset. There are roughly 4,000 rows and fifteen distinct properties in it.	The comparison analysis's results demonstrate how the innovative strategy that has been recommended has a high degree of accuracy and a low level of computational complexity.	-Not mentioned
8	[50]	A unique, reliable, automatic, and intelligent method for COVID-19 categorization utilizing chest X-rays has been suggested in this study. The suggested pipeline combines neuromorphic and machine learning methodologies in a hybrid fashion. The ML structures were trained to categorize the chest X-rays into two groups: normal participants and positive COVID-19 patients.	For assessment, a combination of two different databases is used. The first dataset: 319 X-rays of COVID-19, Middle East Respiratory Syndrome (MERS), Extreme Acute Metabolic Syndrome (SARS), and respiratory diseases were compiled from internet sources and published studies to create the COVID-19 Dataset, a public database on GitHub. Standard (or Healthy) Subjects Database is	The great accuracy of the suggested pipeline was further confirmed by the testing results.	The shortcoming of our suggested pipeline lies in the fact it only provides the doctor with a clear-cut positive or negative COVID-19 assessment.

			the second database: a very well-known database of 5247 standard, infectious agent, and microorganism pneumonic x-rays of the chest, with precision ranging from 400 to 2000 pips.		
9	[51]	An effective machine learning-enabled Neutrosophic C-Means Clustering Skin Lesion Segmenting and Classifying model (NCCOML-SKSC) is presented in this research. For segmenting the dermoscopic imagery, the suggested NCCOML-SKSC paradigm produces an NCC-based segmented method. Additionally, the AlexNet model is used to produce a feature vector. Using a whale optimization technique (WOA) to select the MLP factors, the optimal multilayer perceptron (MLP) structure is employed for the process of classification in the last stage.	A skin lesion benchmark dataset.	In comparison to other models, the results showed that the NCCOML-SKSC paradigm performed effectively with the fewest clusters under all images. The experimental findings also demonstrated the higher accuracy of the NCCOML-SKSC model than the other approaches.	Not mentioned
10	[52]	A backpropagation algorithm-based Interval Neutrosophic Rough Neural Network framework serves as the foundation for the suggested neutrosophic diagnostic method. It makes use of the benefits of neutrosophic set theory to outperform the competing algorithms in addition to enhancing the efficiency of rough neural network structures.	The IN-RNN framework is trained and tested on the cardiotocography (CTG) dataset. The University of California, Irvine (UCI) machine learning archive website has the CTG dataset available for download. To simultaneously detect the condition of fetal heart rate and uterine contraction, CTG uses 2126 instances and 21 input attributes.	The suggested framework is an effective and workable classifier. Furthermore, the suggested paradigm can be used for a number of real-world issues, including social media, satellite images, and coronavirus categorization.	-Not mentioned
11	[53]	The Dynamic Neutrosophic Cognitive Map with Improved Cuckoo Search Algorithm (DNCM-ICSA) with a group classifier is presented in this technological study in order to acquire a profile of gene expression that differentiates between likely	E-GEOD-64707 dataset	-The suggested model operates with acceptable classification and precision.	-Time complexity

		subjects to be in control and those affected by RA. This work consists of four primary phases: feature selection, forecasting, categorization, and data preparation.			
12	[32]	This paper proposes a novel classification procedure. The researchers first reduced the characteristics of the chosen database using the Kernel Principle Component Analysis (KPCA) model. The rationale for using that dimensional technique is to effectively handle computing time. An ideal rule was created in the Neutrosophic Set (NS) to improve the classification procedure, and the Oppositional-based method is integrated with the Elephant Herd Algorithm and this integration was employed to optimize this rule. Following feature selection, the Deep Learning (DL) method was implemented to categorize the data.	Three datasets were selected: Isolet dataset, Musk dataset, and Hill Valley database.	-High Performance	-Not mentioned
13	[54]	In this study, blended feature selection with ensemble machine learning techniques are presented for the successful forecasting of rice illnesses. First, preliminary processes centered around SMOTE are suggested to perform data normalization. Subsequently for effective segmentation, the Adapted Feature weighted Fuzzy Clustering (MFWFC) based segmentation is suggested. EICAs, or enhanced independent component analyses, are subsequently employed for obtaining features in order to improve the accuracy of classifiers. Using HNCMs (Hybrid Neutrosophic Cognitive Maps), the features are chosen.	A dataset of sick rice plants and five standard datasets (UCI, 2010) were included for assessment.	The suggested approach offers a low-cost, high-performance, straightforward, and simple means of identifying rice diseases.	Long computation time

14	[55]	In the suggested approach, the optimal subsets combining GLCM and statistical data that can effectively describe the leaf diseases are chosen using Neutrosophic Cognitive Maps (NCM). The suggested method is tested against eight cutting-edge feature selection algorithms currently in use in order to demonstrate its efficacy on publically accessible photos taken from the plant village archive.	Plant village repository (apple leaf images dataset)	High classification precision	-Computational resources
15	[56]	The suggested method offers three main advantages. First, the classical set is made more inclusive by permitting interpretations of ambiguity in data and expanding the concepts of the truth, ambiguity, and falsehood. The second method goes one step more by using these features in train and verifying the models used for classification throughout a range of distinct domains. It combines a strong technique for picking features that utilizes neutrosophic theory of sets which is considered optimum by evolutionary algorithms.	Three datasets are used for classification:Cleveland Heart Disease, A financial dataset (Customer Satisfaction dataset), and Social media (Twitter Sentiment Analysis).	The suggested neutrosophic technique improves the effectiveness of classification while also enhancing general robustness and dependability in big data analytics.	-Not mentioned

5 | Conclusion

Neutronosophic logic is gaining popularity lately since a lot of real-world issues involve ambiguity, inconsistency, inadequate data, and indeterminacy. A mathematical formula of ambiguity is traditionally indicated by neutrosophic logic (NL). It is recognized as a framework for evaluating truth, falsity, and indeterminacy. Recently, NS and machine learning have been mixed to produce decision strategies for a range of applications that are used in several computer science domains, including forecasting, modeling neutrosophic variables, cancer detection, the Internet of Things, the medical sector, and many other fields. Given the wide application of this integration in many fields, we decided to write a review paper that addresses what has been published in this regard in the period between 2020 to 2024. In this research, we have investigated how the neutrophilic approach might improve algorithms for machine learning in general as well as how to simulate and take advantage of insufficient data, like uncertainty, as a source of knowledge rather than a kind of noise. We made an effort to discuss hybrid techniques and the various domains in which they are applied. Nevertheless, there are still several machine learning methods that can be mapped to a neutrosophic context and show how machine learning and neutrosophic can be used to address real-world problems.

Acknowledgments

The author is grateful to the editorial and reviewers, as well as the correspondent author, who offered assistance in the form of advice, assessment, and checking during the study period.

Funding

This research has no funding source.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the privacy-preserving nature of the data but are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Cao, L., Data science: a comprehensive overview. ACM Computing Surveys (CSUR), 2017. 50(3): p. 1-42.
[2] Sarker, I.H., et al., Cybersecurity data science: an overview from machine learning perspective. Journal of Big data, 2020. 7: p. 1-29.
[3] Sarker, I.H., et al., Mobile data science and intelligent apps: concepts, AI-based modeling and research directions. Mobile Networks and Applications, 2021. 26(1): p. 285-303.
[4] Sarker, I.H., Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2021. 2(3): p. 160.
[5] Mohammed, M., M.B. Khan, and E.B.M. Bashier, Machine learning: algorithms and applications. 2016: Crc Press.
[6] Han, J., J. Pei, and H. Tong, Data mining: concepts and techniques. 2022: Morgan kaufmann.
[7] Janiesch, C., P. Zschech, and K. Heinrich, Machine learning and deep learning. Electronic Markets, 2021. 31(3): p. 685-695.
[8] Zhu, L. and P. Spachos, Towards image classification with machine learning methodologies for smartphones. Machine Learning and Knowledge Extraction, 2019. 1(4): p. 1039-1057.
[9] Kusiak, A., et al., Autonomous decision-making: A data mining approach. IEEE Transactions on Information technology in Biomedicine, 2000. 4(4): p. 274-284.
[10] Jayatilake, S.M.D.A.C. and G.U. Ganegoda, Involvement of machine learning tools in healthcare decision making. Journal of healthcare engineering, 2021. 2021(1): p. 6679512.
[11] Albahri, A.S., et al., Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. Journal of medical systems, 2020. 44: p. 1-11.
[12] Babichev, S., L. Yasinska-Damri, and I. Liakh, A hybrid model of cancer diseases diagnosis based on gene expression data with joint use of data mining methods and machine learning techniques. Applied Sciences, 2023. 13(10): p. 6022.
[13] Islam, M., et al., Likelihood prediction of diabetes at early stage using data mining techniques, in Computer vision and machine intelligence in medical image analysis. 2020, Springer. p. 113-125.
[14] Naresh, V.S. and M. Thamarai, Privacy-preserving data mining and machine learning in healthcare: Applications, challenges, and solutions. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2023. 13(2): p. e1490.
[15] Khafaga, D.S., et al., Data Mining Techniques in Predictive Medicine: An Application in hemodynamic prediction for abdominal aortic aneurysm disease. Journal of Artificial Intelligence and Metaheuristics, 2023. 5(1): p. 29-37.
[16] Shehab, M., et al., Machine learning in medical applications: A review of state-of-the-art methods. Computers in Biology and Medicine, 2022. 145: p. 105458.
[17] Liu, W., Smart sensors, sensing mechanisms and platforms of sustainable smart agriculture realized through the big data analysis. Cluster Computing, 2023. 26(5): p. 2503-2517.
[18] Badapanda, K., D.P. Mishra, and S.R. Salkuti, Agriculture data visualization and analysis using data mining techniques: application of unsupervised machine learning. TELKOMNIKA (Telecommunication Computing Electronics and Control), 2022. 20(1): p. 98-108.

[19] Sattari, M.T., et al., Soil temperature estimation with meteorological parameters by using tree-based hybrid data mining models. Mathematics, 2020. 8(9): p. 1407.
[20] Benos, L., et al., Machine learning in agriculture: A comprehensive updated review. Sensors, 2021. 21(11): p. 3758.
[21] Al-Hashedi, K.G. and P. Magalingam, Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 2021. 40: p. 100402.
[22] Kumar, T.S., Data mining based marketing decision support system using hybrid machine learning algorithm. Journal of Artificial Intelligence, 2020. 2(03): p. 185-193.
[23] Dixon, M.F., I. Halperin, and P. Bilokon, Machine learning in finance. Vol. 1170. 2020: Springer.
[24] Sáiz-Manzanares, M.C., et al., Systematic review on inclusive education, sustainability in engineering: An analysis with mixed methods and data mining techniques. Sustainability, 2020. 12(17): p. 6861.
[25] Le, Y., Research on data resource management of biomass energy engineering based on data mining. Energy Reports, 2022. 8: p. 1482-1492.
[26] Su, M., H. Peng, and S. Li, A visualized bibliometric analysis of mapping research trends of machine learning in engineering (MLE). Expert Systems with Applications, 2021. 186: p. 115728.
[27] Wang, H., et al., Single valued neutrosophic sets. 2010: Infinite study.
[28] Mandour, S., An Exhaustive Review of Neutrosophic Logic in Addressing Image Processing Issues. Neutrosophic Systems with Applications, 2023. 12: p. 36-55.
[29] Mondal, K., S. Pramanik, and B.C. Giri, Role of neutrosophic logic in data mining. New Trends in Neutrosophic Theory and Application. Pons Editions, Brussels, 2016: p. 15-23.
[30] Abdel-Basset, M., et al., Neutrosophic association rule mining algorithm for big data analysis. Symmetry, 2018. 10(4): p. 106.
[31] Awajan, I., M. Mohamad, and A. Al-Quran, Sentiment analysis technique and neutrosophic set theory for mining and ranking big data from online reviews. IEEE Access, 2021. 9: p. 47338-47353.
[32] Gomathy, V., et al., Optimal neutrosophic rules based feature extraction for data classification using deep learning model, in Soft Computing for Data Analytics, Classification Model, and Control. 2022, Springer. p. 57-79.
[33] Smarandache, F., A unifying field in Logics: Neutrosophic Logic, in Philosophy. 1999, American Research Press. p. 1-141.
[34] Zadeh, L.A., G.J. Klir, and B. Yuan, Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers. Vol. 6. 1996: World scientific.
[35] Chai, J.S., et al., New similarity measures for single-valued neutrosophic sets with applications in pattern recognition and medical diagnosis problems. Complex & Intelligent Systems, 2021. 7: p. 703-723.
[36] Abdel-Basset, M., et al., A novel group decision making model based on neutrosophic sets for heart disease diagnosis. Multimedia Tools and Applications, 2020. 79: p. 9977-10002.
[37] Edalatpanah, S. and F. Smarandache, Data envelopment analysis for simplified neutrosophic sets. 2019: Infinite Study.
[38] Sah, S., Machine learning: a review of learning types. 2020.
[39] Jiang, T., J.L. Gradus, and A.J. Rosellini, Supervised machine learning: a brief primer. Behavior therapy, 2020. 51(5): p. 675-687.
[40] Alloghani, M., et al., A systematic review on supervised and unsupervised machine learning algorithms for data science. Supervised and unsupervised learning for data science, 2020: p. 3-21.
[41] Matsuo, Y., et al., Deep learning, reinforcement learning, and world models. Neural Networks, 2022. 152: p. 267-275.
[42] Yang, X., et al., A survey on deep semi-supervised learning. IEEE Transactions on Knowledge and Data Engineering, 2022. 35(9): p. 8934-8954.
[43] Alsattar, H.A., et al., Developing deep transfer and machine learning models of chest X-ray for diagnosing COVID-19 cases using probabilistic single-valued neutrosophic hesitant fuzzy. Expert Systems with Applications, 2024. 236: p. 121300.
[44] Nafei, A., et al., Smart TOPSIS: A Neural Network-Driven TOPSIS with Neutrosophic Triplets for Green Supplier Selection in Sustainable Manufacturing. Expert Systems with Applications, 2024. 255: p. 124744.
[45] Saqlain, M., et al., Uncertainty and decision-making with multi-polar interval-valued neutrosophic hypersoft set: A distance, similarity measure and machine learning approach. Alexandria Engineering Journal, 2023. 84: p. 323-332.
[46] Khan, M.A. and N.S. Alghamdi, A neutrosophic WPM-based machine learning model for device trust in industrial internet of things. Journal of Ambient Intelligence and Humanized Computing, 2023. 14(4): p. 3003-3017.
[47] Singh, S.K., V. Abolghasemi, and M.H. Anisi, Skin cancer diagnosis based on neutrosophic features with a deep neural network. Sensors, 2022. 22(16): p. 6261.
[48] Ahmed, R., F. Nasiri, and T. Zayed, A novel Neutrosophic-based machine learning approach for maintenance prioritization in healthcare facilities. Journal of Building Engineering, 2021. 42: p. 102480.
[49] El-Douh, A.A., et al., Heart Disease Prediction under Machine Learning and Association Rules under Neutrosophic Environment. Neutrosophic Systems with Applications, 2023. 10: p. 35-52.
[50] Yasser, I., et al., A hybrid automated intelligent COVID-19 classification system based on neutrosophic logic and machine learning techniques using chest X-Ray images. Advances in Data Science and Intelligent Data Communication Technologies for COVID-19: Innovative Solutions Against COVID-19, 2022: p. 119-137.
[51] Taher, F. and A. Abdelaziz, Neutrosophic C-Means Clustering with Optimal Machine Learning Enabled Skin Lesion Segmentation and Classification. International Journal of Neutrosophic Science, 2022. 19(1): p. 177-187.
[52] Amin, B., et al., Intelligent neutrosophic diagnostic system for cardiotocography data. Computational Intelligence and Neuroscience, 2021. 2021(1): p. 6656770.

[53] Chithra, B. and R. Nedunchezhian, Dynamic neutrosophic cognitive map with improved cuckoo search algorithm (DNCM-ICSA) and ensemble classifier for rheumatoid arthritis (RA) disease. Journal of King Saud University-Computer and Information Sciences, 2022. 34(6): p. 3236-3246.

[54] Senthilkumar, T. and P. Prabhusundhar, ENSEMBLE MACHINE LEARNINGAND HYBRID NEUTROSOPHIC COGNITIVE MAPS BASED FEATURE SELECTION FOR RICE DISEASE PREDICTION.

[55] Shadrach, F.D. and G. Kandasamy, RETRACTED ARTICLE: Neutrosophic Cognitive Maps (NCM) based feature selection approach for early leaf disease diagnosis. Journal of Ambient Intelligence and Humanized Computing, 2021. 12(5): p. 5627-5638.

[56] Abdulbaqi, A.S., et al., Neutrosophic Sets in Big Data Analytics: A Novel Approach for Feature Selection and Classification. International Journal of Neutrosophic Science, 2024. 25(1): p. 428-28-438.