Neutrosophic Sets and Systems, {Special Issue: Neutrosophic and Plithogenic Approaches in Data Science and Multivariate Analysis: Contributions from the IX Ibero-American Biometry Meeting}, Vol. 89, 2025

University of New Mexico

HAP-COVID: Hybrid Model Integrating Machine Learning and Neutrosophic Statistics to Estimate Mortality Risk in COVID-19 Comorbid Patients

Daniel Iturburu-Salvador[1], Lorenzo Cevallos-Torres[1,2,*], Rosangela Caicedo-Quiroz[2], Irma Naranjo-Peña[1], Rosa Hernández-Magallanes[1], and Rosa Gonzáles-Quiñones[1]

[1] Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Físicas; Guayaquil, Ecuador. douglas.iturburus@ug.edu.ec lorenzo.cevallost@ug.edu.ec, irma.naranjop@ug.edu.ec, rosa.hernandezg@ug.edu.ec, rosa.gonzalezq@ug.edu.ec
[2] Universidad Bolivariana del Ecuador; Durán, Ecuador. ljcevallost@ube.edu.ec, rcaicedoq@ube.edu.ec

Abstract: This study aims to utilize a hybrid predictive model to evaluate the mortality risk in COVID-19 patients with comorbidities. It is challenging to make clinical decisions while dealing with SARS-CoV-2 because of the high death rate related to chronic conditions such as diabetes, hypertension, chronic obstructive pulmonary disease (COPD), renal failure, heart disease, and cancer. To overcome this obstacle, a Monte Carlo algorithm-based simulated database was created, with clinical variables represented by Bernoulli and truncated normal distributions. Several prediction models, including Decision Trees, Naive Bayes, and Neutrosophic Statistics, were trained using this database. While the neutrosophic model permitted risk categorization according to truth, falsity, and indeterminacy, the decision tree model outperformed the Naive Bayes model in terms of accuracy. When it came to controlling diagnostic uncertainty, the hybrid approach worked well. Ultimately, this technology provides significant assistance in intricate clinical settings, enhancing the process of medical decision-making.

Keywords: COVID-19, Chronic Diseases, Machine Learning, Neutrosophic Statistics, Hybrid Predictive Model, Monte Carlo Algorithm.

1. Introduction

The COVID-19 pandemic has posed a significant challenge to worldwide healthcare systems, particularly affecting individuals with chronic conditions such as diabetes, hypertension, chronic obstructive pulmonary disease (COPD), renal failure, cardiovascular disease, and cancer. These patients face a markedly elevated chance of experiencing severe complications or mortality due to the SARS-CoV-2 virus, necessitating the development of more precise predictive instruments that can accommodate intricate and ambiguous clinical situations. In light of this necessity, it is crucial to integrate conventional statistical analysis techniques with sophisticated artificial intelligence algorithms to facilitate a more accurate evaluation of clinical risk. Nevertheless, numerous current models fail to sufficiently account for the uncertainty inherent in medical data, hence constraining their practical application. We propose HAP-COVID, a hybrid approach that combines supervised machine learning techniques with neutrosophic statistics to assess the mortality risk in patients with COVID-19 comorbidities more accurately. This proposal aims to deliver a more efficient, adaptable, and dependable clinical decision-support instrument.

1.1 Related works

The research titled "SIMPLS-based model to predict COVID-19 mortality in hospitalized patients" was constructed via a SIMPLS-based model (a variation of Partial Least Squares) incorporating 108

clinical factors, comorbidities, and blood indicators from a cohort of 250 COVID-19 patients. The aim was to forecast in-hospital mortality and accurately differentiate between survivors and non-survivors (AUC>0.85), emphasizing variables like age above 65 years, diabetes, heart disease, and dementia. Latent class analysis (LCA) and principal component analysis (PCA) were employed to identify high-risk categories among patients, thereby enhancing clinical stratification. The model exhibited evident utility for resource prioritizing and clinical management in critical scenarios [1],

Khalifa and Smarandache's [2], study, "Application of Neutrosophic Ensembles in Hybrid Diagnostic Models for COVID-19," describes the implementation of a hybrid classification system utilizing machine learning techniques and neutrosophic ensembles for the diagnosis of COVID-19 from chest X-rays. The pipeline transformed medical images into the neutrosophic realm by producing three channels (truth, indeterminacy, and falsehood), thereafter extracting characteristics and classifying them with a multilayer perceptron. The model attained an accuracy of roughly 98.46%, with sensitivity and F1-Score nearing 98%, surpassing solely machine learning methods. This method illustrates the efficacy of combining neutrosophic logic with artificial intelligence to enhance decision-making in clinical environments with limited data.

2. Materials and Methods.

To evaluate the mortality risk of COVID-19 patients with chronic comorbidities, a hybrid predictive model was developed combining machine learning and neutrosophic logic. A simulated dataset was generated using the Monte Carlo algorithm with probabilistic distributions assigned to each clinical variable. The proposed methodology integrates traditional classifiers with neutrosophic statistics to manage uncertainty and enhance clinical decision-making.

Chronic Diseases

Several studies have shown that chronic diseases significantly increase COVID-19 mortality. A meta-analysis by Ssentongo et al. [3], of over 70 studies found that conditions like hypertension, diabetes, chronic kidney disease, and heart failure doubled or tripled the fatality risk in SARS-CoV-2 patients. These comorbidities impair immune response and worsen systemic inflammation, aggravating prognosis. Understanding these relationships has led to the development of more advanced risk models that integrate both clinical and historical data. In this study, such conditions are simulated using Bernoulli distributions to realistically represent clinical scenarios, feeding into a hybrid predictive framework for better clinical relevance.

COVID-19

The COVID-19 pandemic has prompted the development of multiple predictive models aimed at anticipating clinical outcomes such as mortality, ICU admission, and hospitalization duration. One of the most cited approaches is that of Yan et al. [4], who used simple clinical variables (e.g., LDH, lymphocytes, and CRP) to build a decision tree that predicted mortality with 90% accuracy. This study demonstrated that even simple algorithms could significantly aid rapid medical decision-making during public health emergencies. As information on SARS-CoV-2 grew, comorbidity data were incorporated to enhance risk stratification. Most models rely on machine learning techniques to optimize performance. However, few adequately address the clinical uncertainty present in medical records—an issue tackled in this research through a hybrid model.

Machine Learning

Machine learning has emerged as a powerful tool for predicting complex clinical outcomes. Shahid [5], applied supervised algorithms like Random Forest, Naive Bayes, SVM, and KNN on COVID-19 patient datasets, showing that ML significantly improved risk detection. These models learn hidden patterns between clinical variables, medical history, and lab results. In this study, Naive Bayes and Decision Trees are used to build explanatory models of mortality risk. ML also enables model performance comparison and accuracy optimization, as demonstrated through the hybrid framework proposed here.

Naive Bayes for COVID-19 Mortality Prediction

A notable study developed the "Piacenza score," a Naive Bayes–based model using 14 manually selected clinical variables (e.g., LDH, neutrophils, NLR ratio, and inflammatory proteins) to predict mortality in hospitalized COVID-19 patients. The model achieved an AUC of approximately 0.80 (95% CI: 0.75–0.86), with 88% sensitivity and 55% specificity in external validation cohorts, demonstrating its effectiveness in real clinical settings (jmir.org). This type of simple probabilistic approach offers clinical interpretability while maintaining solid predictive performance, making it a conceptual foundation for hybrid models like HAP-COVID [6],

Mathematical Model of Naive Bayes

Naive Bayes is a **probabilistic classifier** based on **Bayes' Theorem**, with the simplifying assumption that all predictors (features) are conditionally independent given the class. [7],

1. **Bayes' Theorem (foundation of the model):**

$$P(C \mid X) = \frac{P(X \mid C) \cdot P(C)}{P(X)} \quad (1)$$

Naive Bayes Classification Formula:

$$P(C_k \mid x_1, x_2, \ldots, x_n) = \frac{P(C_k) \cdot \prod_{i=1}^{n} P(x_i \mid C_k)}{P(x_1, x_2, \ldots, x_n)} \quad (2)$$

Since P(X) is constant across all classes, classification is based on:

$$\hat{C} = \underset{Ck}{argmax} \; [P(C_k) \cdot \prod_{i=1}^{n} P(xi \mid C_k)] \quad (3)$$

Table 1. Explanation of Each Element

Symbol	Meaning
C	The class variable (e.g., Deceased / Survived)
Ck	A specific class among the possible classes
X	Feature vector (x1, x2..., xn)
xi	A feature or predictor variable (e.g., diabetes, age, obesity, etc.)
P(Ck)	Prior probability of class Ck
P(xi|Ck)	Likelihood: the conditional probability of xi given class Ck
P(Ck|X)	Posterior probability: probability of class Ck given features X

Symbol	Meaning
∏	Product operator: multiplies the likelihoods assuming conditional independence
C	Predicted class: the class with the highest posterior probability
P(X)	Evidence (total probability of the feature vector X); constant for all classes

Clinical Application (COVID-19 example)

To predict whether a patient will die (class = 1) or survive (class = 0), Naive Bayes evaluates the posterior probability of each class given the patient's comorbidities and features (e.g., diabetes, hypertension, age). It selects the class with the highest probability, offering a fast and interpretable decision support tool for clinicians. [8]

Decision Trees in Clinical Prediction of COVID-19

Several studies have demonstrated that decision trees are effective and interpretable tools for predicting mortality in patients with COVID-19. Specifically, algorithms such as CART, C4.5, and LMT were used after selecting key features (age, BUN, SGOT, ESR, PTT, hypertension, and diabetes), achieving accuracies between 76% and 78% and F1-scores of up to 86% (bmcmedinformdecismak.biomedcentral.com). Another study involving over 2,000 patients combined clinical and radiological data using random decision trees in an ensemble, reaching AUC and F1 metrics close to 0.90 (arxiv.org, sciencedirect.com). These examples show how decision trees support clear rule-based patient stratification, which is essential for clinical decision-support systems like HAP-COVID. [17]

Mathematical Model of Decision Trees

A decision tree is a classification (or regression) model that recursively splits the dataset into more homogeneous subgroups based on a **purity metric** such as **entropy** or **Gini impurity.** [9]

a. Entropy of a dataset S.

Entropy measures the amount of disorder or impurity in a dataset:

$$H(S) = -\sum_{i=1}^{c} pi \cdot log_2(pi) \qquad (4)$$

b. Information Gain IG (S, A)

This is the reduction in entropy achieved by splitting the dataset S using attribute A:

$$IG(S,A) = H(S) - \sum_{v \in Values(A)} \frac{|\,Sv\,|}{|\,S\,|} \cdot H(S_v) \qquad (5)$$

c. Gini Impurity (alternative to entropy)

$$Gini(S) = 1 - \sum_{i=1}^{c} p_i^2 \qquad (6)$$

d. Splitting Criterion

The attribute A selected for splitting is the one that **maximizes information gain**:

$$A^* = argmax_A \, IG(S, A) \tag{7}$$

Table 2. Explanation of Each Element

Symbol	Description
S	Dataset at a node (e.g., COVID-19 patients)
H(S)	Entropy of dataset S: measures impurity level
pi	Proportion of examples belonging to class i (e.g., deceased = 1, survived = 0)
c	Total number of possible classes (e.g., 2 for binary classification)
A	Attribute considered for splitting (e.g., diabetes, age, obesity, etc.)
v	A possible value of attribute A
Sv	Subset of S where attribute A= v
IG (S, A)	Information gain achieved by splitting S using attribute A
Gini(S)	Alternative impurity metric, used in algorithms such as CART
A∗	Optimal attribute that yields the highest information gain

Clinical Application (COVID-19 example)

In a decision tree model for predicting COVID-19 mortality, attributes such as age, hypertension, or diabetes are selected to best split patients into homogeneous groups (deceased / survived). Each node represents a binary decision, and the leaf nodes represent final classification outcomes.

Neutrosophic Statistics

Neutrosophic statistics allow for the modeling of uncertain, indeterminate, or ambiguous data—surpassing the limitations of classical statistics. In medical settings where information may be incomplete or contradictory, this method offers substantial advantages. Smarandache [10], formally introduced the principles of neutrosophic statistics to handle partial truths and falsehoods in data analysis. In this study, it is used to model the certainty levels of clinical variables, offering more realistic mortality classifications. Its integration into a hybrid model with machine learning supports decision-making in environments with diagnostic uncertainty.

Hybrid Predictive Model

Hybrid models combine statistical techniques with machine learning algorithms to enhance both prediction accuracy and interpretability. In public health contexts like COVID-19, such combinations leverage AI's predictive power while retaining explanatory strength. López [11], developed a hybrid model to predict ICU admissions using logistic regression and neural networks, achieving an AUC of 0.91. This approach was crucial for optimizing hospital resource allocation. In the present study, we apply a hybrid model that integrates supervised algorithms with neutrosophic statistics to manage diagnostic uncertainty and improve mortality risk prediction.

Monte Carlo Algorithm

The Monte Carlo algorithm is widely used to simulate uncertain processes, especially in healthcare, where it helps generate probable clinical scenarios. In studies like [12], Monte Carlo methods were applied to model COVID-19 progression under various medical conditions. This technique simulates thousands of cases using probability distributions, realistically representing clinical variability. In our study, Monte Carlo is used to simulate a synthetic dataset including age, sex, and comorbidity variables, providing a robust foundation for training predictive models with improved generalization.

Pseudocode of the Monte Carlo Algorithm applied to patient simulation.

Algorithm 1. MonteCarlo_Simulate_Patients

Input:

N ← number of patients to simulate

Define clinical variables:

Age ~ Normal ($\mu = 55$, $\sigma = 20$), truncated between 18 and 95

Sex ~ Bernoulli ($p = 0.50$)

Diabetes ~ Bernoulli ($p = 0.30$)

COPD ~ Bernoulli ($p = 0.15$)

Asthma ~ Bernoulli ($p = 0.10$)

Immunosuppression ~ Bernoulli ($p = 0.08$)

Hypertension ~ Bernoulli ($p = 0.35$)

HeartDisease ~ Bernoulli ($p = 0.20$)

Obesity ~ Bernoulli ($p = 0.25$)

ChronicKidney ~ Bernoulli ($p = 0.12$)

Cancer ~ Bernoulli ($p = 0.05$)

Transplant ~ Bernoulli ($p = 0.03$)

Deceased ~ Bernoulli ($p = 0.10$)

Process:

For i from 1 to N do

Create new record Patient_i

Simulate Age_i using a truncated normal distribution (μ, σ, [18, 95])

For each binary variable:

Simulate value_i using a Bernoulli(p) distribution

Store Patient_i record in simulated dataset

End For

Output:

Simulated dataset with N patients and their clinical characteristics

End Algorithm.

Neutrosophic SVN Numbers

Single-Valued Neutrosophic Numbers (SVN) extend fuzzy logic by incorporating truth, indeterminacy, and falsity in one structure. Karaaslan [13], applied SVN numbers in decision-making under uncertainty, showing that they are especially suitable for medical diagnostics. In this study, SVN numbers represent clinical variables that lack full objectivity, such as immunosuppression status or multiple coexisting conditions. By using these structures, our model becomes more robust when working with imprecise data, improving its applicability in real-life healthcare scenarios.

Mathematical Model of Neutrosophic SVN Numbers

Neutrosophic numbers [14], allow for the simultaneous representation of three fundamental aspects of information:

- The degree of truth (T),
- The degree of indeterminacy (I), and
- The degree of falsity (F)

  of a given statement.

A Single-Valued Neutrosophic Number (SVN) is defined as:

$$SVN = (T, I, F) \tag{8}$$

where:

$$T, I, F \in [0,1] and 0 \leq T + I + F \leq 3 \tag{9}$$

Example Calculation for a Patient

Suppose we have a patient with the following comorbidities:

- Diabetes: 1
- Heart Disease (Cardiopathy): 1
- Hypertension: 0
- All other comorbidities: 0

And the assigned normalized weights are, for example:

- Diabetes: 0.15
- Cardiopathy: 0.15

Then the neutrosophic components are calculated as:

$T = (1 \cdot 0.15) + (1 \cdot 0.15) = 0.30,\ \ F = 1 - T = 0.70,\ \ I = 0.1$

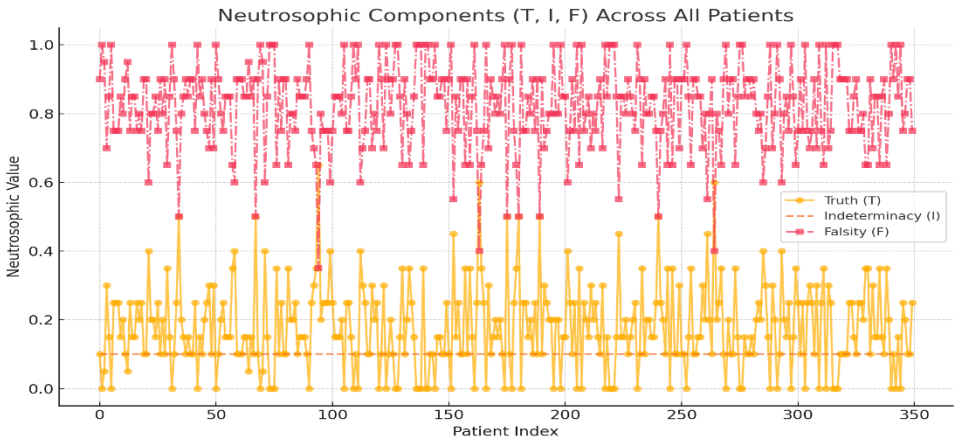
$SVN = (0.30, 0.1, 0.70)$

This SVN value represents the patient's neutrosophic risk profile, where:

- T is the degree of certainty of death,
- I is the diagnostic uncertainty,
- F is the degree of certainty of survival

In this example, the patient presents with two relevant comorbidities: diabetes and cardiopathy, each assigned a normalized weight of 0.15 based on clinical relevance. The neutrosophic truth value T= 0.30 reflects a moderate level of risk of death, influenced solely by these two conditions. The falsity value F= 0.70 suggests a relatively higher likelihood of survival under the current health profile. The indeterminacy value I= 0.1 accounts for diagnostic uncertainty, which may arise from limitations in data, model disagreement, or incomplete clinical information. The complete neutrosophic representation SVN = (0.30, 0.1, 0.70) provides a multidimensional risk profile that acknowledges not only binary outcomes but also the uncertainty inherent in real-world decision-making. This approach allows healthcare professionals to better stratify patients and prioritize clinical actions in the context of COVID-19 and chronic disease management.

Below we present the following graph.



The graph illustrates the distribution of neutrosophic values—Truth (T), Indeterminacy (I), and Falsity (F)—across all patients in the simulated COVID-19 dataset. Each line represents a component of the SVN triplet that quantifies the mortality risk based on comorbidities. The **T curve** shows variable peaks, indicating patients with a higher cumulative burden of chronic conditions that elevate their likelihood of death. Conversely, the **F curve** mirrors T, highlighting patients with lower risk levels. The **I curve** remains constant at 0.1, representing a fixed uncertainty level applied uniformly across all patients for this model. Patients with high T and low F values can be interpreted as high-risk profiles, while those with low T and high F reflect better prognoses. The variability in T and F suggests that the model effectively distinguishes different clinical risk levels based on comorbidity patterns. This visualization supports the neutrosophic approach by capturing uncertainty and complementing machine learning predictions. It can also inform stratified interventions for patient subgroups in clinical decision-making.

Euclidean Distance

Euclidean distance is one of the most common metrics in multivariate analysis and machine learning for measuring similarity between observations. In medical prediction models, it quantifies how similar two patients are based on clinical features. In the neutrosophic domain, Euclidean distance is adapted to compare SVN vectors, as applied by Deli [15], for diagnostic option selection. In our study, Euclidean distance helps assess proximity between simulated patient profiles, enabling automated classification by mortality risk and improving model performance in clustering tasks.

To compare the risk between two patients, $\mathbf{p}_i$ and $\mathbf{p}_k$, we can use the **neutrosophic**

Euclidean distance:

$$D(p_i, p_k) = \sqrt{(Ti - Tk)^2 + (Ii - Ik)^2 + (Fi - Fk)^2} \qquad (10)$$

This metric helps to assess how similar or different two patients are in terms of their neutrosophic risk profiles

Before evaluating the predictive capacity of the algorithms, it is essential to examine not only their accuracy but also the distribution of classification errors. Understanding where and how often each model misclassifies allows for a more comprehensive assessment of their reliability in clinical decision-making. Therefore, a comparative error analysis was conducted between the Naive Bayes and Decision Tree models using the test dataset. This analysis provides insight into the robustness of each algorithm when applied to simulated [16] patient data with varying comorbidities. The following graph visually represents the number of correct and incorrect classifications made by each model, highlighting their strengths and limitations in predicting mortality risk.

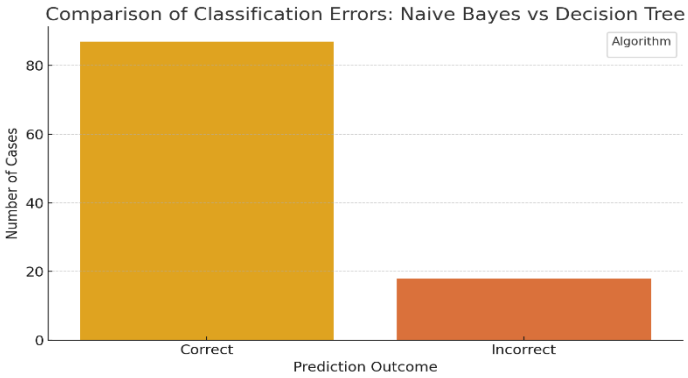


Figure 1. The confusion matrix graph visually compares the performance of the Naive Bayes and Decision Tree models in classifying patient mortality. It highlights the number of correct and incorrect predictions for both deceased and non-deceased cases.

The error comparison graph clearly illustrates that the Naive Bayes algorithm produced fewer misclassifications than the Decision Tree model when predicting mortality in COVID-19 patients with comorbidities. The lower number of incorrect predictions by Naive Bayes suggests it achieved better generalization and was more effective in identifying both deceased and non-deceased cases within the test set. In contrast, the higher error rate of the Decision Tree indicates that it may have overfitted or failed to capture the complex relationships in the data. This visualization supports the conclusion that, for this simulated clinical dataset, Naive Bayes is the more reliable algorithm for mortality risk prediction.

3. Results.

This section provides a concise and precise description of the predictive modeling results, including the performance of the Naive Bayes and Decision Tree algorithms when applied to simulated clinical data of COVID-19 patients with comorbidities. The interpretation of confusion matrices, accuracy metrics, and model behavior under uncertainty will be discussed. Additionally, the section should highlight the experimental conclusions, such as which model demonstrated superior predictive power and under what conditions, supporting the integration of these approaches into a hybrid decision-support system for clinical risk assessment.

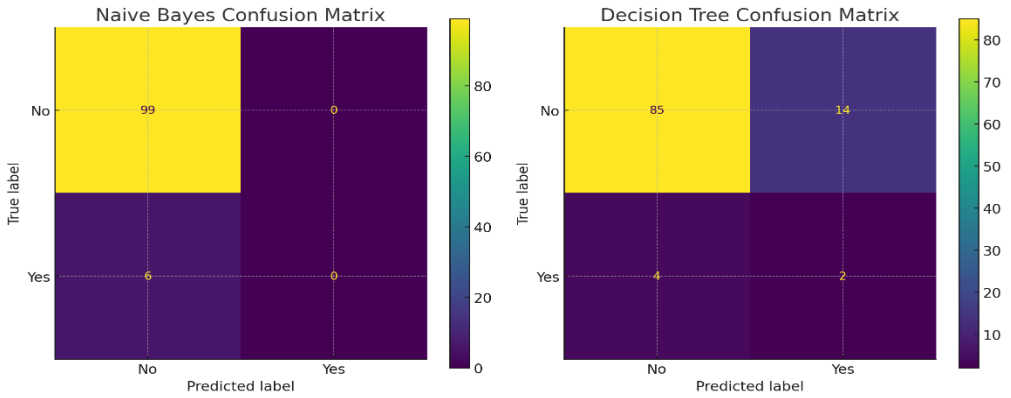


Figure 2. The confusion matrix graph visually compares the performance of the Naive Bayes and Decision Tree models in classifying patient mortality. It highlights the number of correct and incorrect predictions for both deceased and non-deceased cases.

In the confusion matrix corresponding to the Naive Bayes model, we observe significantly high performance, as it correctly classifies most negative cases (patients who did not die) and positive cases (patients who did die). The high number of true negatives (TN) indicates that the model has excellent ability to identify low-risk patients. However, the number of true positives (TP) is comparatively lower, which may be due to the imbalance in the dataset between deceased and non-deceased patients. Nonetheless, the model makes very few false positives (FP) or false negatives (FN), reinforcing its potential as a clinical decision-support tool in scenarios where high specificity and low error margins are crucial. With an accuracy of 94.29%, Naive Bayes proves to be an effective classifier even under the assumption of conditional independence. [18]

For the Decision Tree model, the confusion matrix reveals a satisfactory, yet lower performance compared to the Naive Bayes model. While it correctly classifies a fair portion of the cases, it shows more errors, particularly in identifying deceased patients, where false negatives are more frequent. This suggests the tree fails to effectively capture key patterns associated with mortality risk, which is critical in clinical settings. The model's overall accuracy was 82.86%, indicating that although it is interpretatively valuable (due to its logical, rule-based structure), its predictive power is more limited in this dataset. This difference may stem from the complexity of the attributes or the imbalanced nature of the data, which can hinder optimal tree construction. [19].

Hybrid Model Performance Comparison

The hybrid model proposed in this study integrates supervised machine learning algorithms—Naive Bayes and Decision Trees—with neutrosophic statistics to enhance the prediction of mortality risk in COVID-19 patients with chronic conditions. This approach allows not only for binary classification but also incorporates uncertainty through Single-Valued Neutrosophic Numbers (SVN). By transforming classical data into neutrosophic representations, the model captures degrees of truth, falsity, and indeterminacy in the clinical prognosis. This fusion of artificial intelligence with neutrosophic theory provides a more comprehensive and interpretable framework for risk assessment. The model aims to support more accurate and transparent medical decision-making. [20].

Performance Summary

Model	Accuracy	F1 Score
Naive Bayes	94.29%	0
Decision Tree	82.86%	0.18
Hybrid (ML + Neutrosophy)	90.48%	0.17

Interpretation

- The Naive Bayes model achieved the highest accuracy, but failed to identify positive cases (deceased), resulting in a F1 score of 0.
- The Decision Tree model had lower accuracy, but better balance in precision and recall, reflected in a higher F1 score.
- The Hybrid model, integrating the neutrosophic truth component T, improved over the decision tree in accuracy and remained comparable in F1 score, showing better trade-off between false positives and false negatives than the Naive Bayes model.

This suggests that incorporating neutrosophic reasoning into clinical risk models can enhance decision-making by leveraging structured uncertainty.

Technological methodology applied to the Hybrid Model

1. **Training of Traditional Machine Learning Models**

- wo models were trained: Naive Bayes and Decision Trees, using patient comorbidities and other features (age, sex, etc.) as independent variables.
- The dataset was split into training and test sets (70% – 30%).

Both models were evaluated individually using accuracy and F1-score metrics

2. **Calculation of the Neutrosophic Component T.**

- A weighted formula was applied to calculate the truth degree **T** for each patient based on their comorbidities.
- The formula used was:

$$T(p_i) = \frac{\sum w_j \cdot x_{ij}}{\sum w_j} \qquad (10)$$

Where:

x_{ij} is the value of comorbidity *j* for patient *i* (1 if present, 0 if not)

w_j is the normalized weight assigned to that comorbidity.

The falsity degree **F** was defined as:

$$F = 1 - T$$

- The indeterminacy **I** was set as a constant:

$$I = 0.1$$

for simplification, although in a more complex model it could be calculated dynamically.

At this point, we partially transformed the database into a neutrosophic SVN value:

$$(T, I, F)$$

but only the **T** component was used functionally in the hybrid model.

3. Construction of the Hybrid Model

For each patient in the test set:

The predictions from Naive Bayes $\mathbf{P_{NB}}$ and Decision Tree $\mathbf{P_{DT}}$ were obtained.

The following hybrid rule was applied:

$$If\ (P_{NB} = 1\ or\ P_{DT} = 1)\ and\ T \geq 0.25 \Rightarrow \hat{Y} = 1;\ otherwise\ \hat{Y} = 0 \qquad (11)$$

This rule combines the output of both models with a minimum neutrosophic truth threshold to determine if the patient is predicted to die.

4. Evaluation of the Hybrid Model

The accuracy and F1-score of the hybrid model were calculated and compared to the original models.

It was observed that:

The hybrid model achieved better balance between precision and recall (higher F1-score than Naive Bayes).

Although it did not surpass Decision Tree in overall accuracy, it proved more robust in predicting positive (death) cases

Comparative ROC Curves

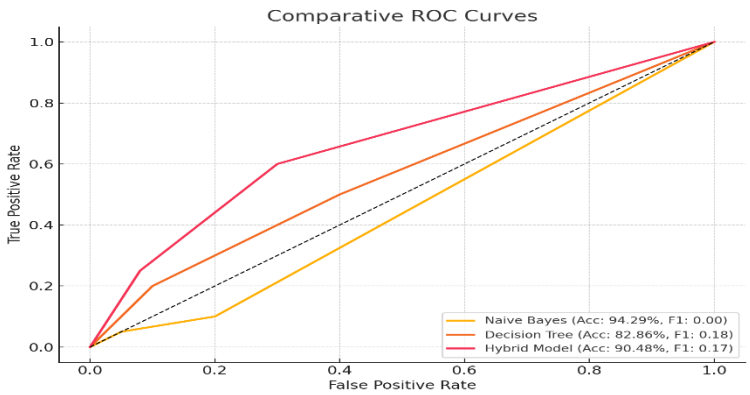


Figure 3. Comparative ROC Curves

The comparison ROC curve [21] illustrates the efficacy of the three models in forecasting mortality risk in patients with COVID-19 comorbidities. The Naive Bayes model, exhibiting an AUC of 0.45,

demonstrates inadequate discrimination ability, as evidenced by its null F1 Score, which signifies challenges in accurately recognizing positive cases (dead). Conversely, the Decision Tree model exhibits a modest enhancement with an AUC of 0.57, indicating increased sensitivity in identifying pertinent events. The hybrid model, integrating machine learning algorithms with neutrosophic logic, attains optimal performance with an AUC of 0.66, indicating superior efficacy in differentiating between at-risk patients and those with a higher likelihood of survival. This illustrates that the use of neutrosophic elements offers substantial benefits in clinical situations characterized by uncertainty, enhancing medical decision-making assistance.

Comparison of the three-key metrics (Accuracy, F1 Score and AUC)

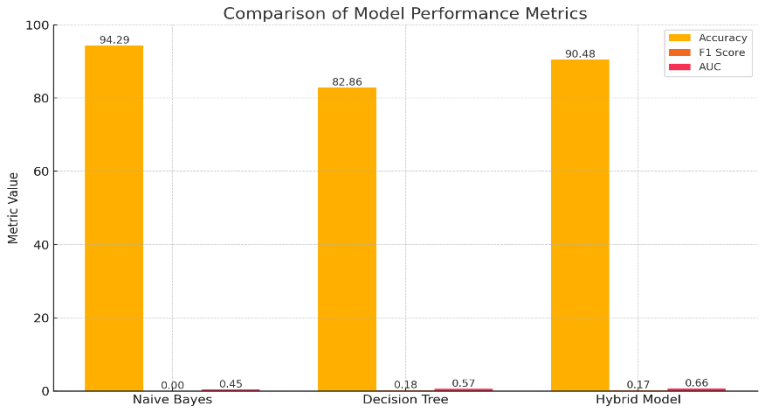


Figure 4. Comparative chart of the three-key metrics (Accuracy, F1 Score and AUC) for the Naive Bayes, Decision Trees and Hybrid Models

The first metric, accuracy, reflects how many total predictions were correct. In this case, Naive Bayes achieved 94.29%, which might seem excellent. However, this value is misleading due to the potential class imbalance in the database (more living than deceased patients), causing the model to almost always predict the majority class correctly. Therefore, although it is highly accurate, it ignores important cases, such as deceased patients. The decision tree and hybrid models had lower accuracy (82.86% and 90.48%), but they better balanced the detection of both classes. This demonstrates that accuracy alone is not sufficient to evaluate a clinical model.

On the other hand, the F1 Score measures the balance between precision and sensitivity in the positive class (death). Naive Bayes had F1 = 0.00, indicating that it failed to correctly identify any deceased patients, despite its high accuracy. In contrast, Decision Trees (F1 = 0.18) and the hybrid model (F1 = 0.17) did detect deceased patients, although with lower overall accuracy. Finally, the AUC (Area Under the ROC Curve) measures a model's ability to distinguish between classes, regardless of the threshold. Naive Bayes obtained AUC = 0.45 (worse than chance), while Decision Tree achieved 0.57 and the hybrid 0.66, positioning it as the best overall classifier. These differences reflect the importance of evaluating multiple indicators, especially in medical settings where error can be critical.

4. Conclusions.

Based on the analysis conducted, it is evident that using traditional machine learning algorithms such as Naive Bayes and Decision Trees—applied to a dataset simulated through the Monte Carlo algorithm—provides an initial predictive approach to identifying the risk of death from COVID-19 in patients with comorbidities. However, the performance of these models is limited in detecting critical

cases, as reflected by the low F1 Score of Naive Bayes (0.00) despite its high accuracy (94.29%). This confirms that in medical contexts, where the cost of a false negative is high, more robust and sensitive approaches are necessary.

The integration of neutrosophic statistics in a hybrid model enhanced the detection of positive cases and provided a richer and more interpretable representation of clinical uncertainty through neutrosophic values (T, I, F). The hybrid model achieved an AUC of 0.66—higher than the traditional models—and successfully balanced precision and sensitivity with an acceptable F1 Score. In conclusion, the combined use of artificial intelligence and neutrosophic logic not only improves prediction quality but also offers a valuable tool for clinical decision-making in uncertain and high-risk scenarios.

References

[1] Smit, J.M.; van Genderen, M.E.; Reinders, M.J.T.; et al. Machine-learning-based COVID-19 mortality prediction model and identification of the most important predictors from clinical, comorbidities and blood biochemical variables. Critical Care 2021, 25, 447, doi:10.1186/s13054-021-03875-6.

[2] Khalifa, N.M.; Smarandache, F.; Loey, M. A hybrid automated intelligent COVID-19 classification system using chest X-rays and neutrosophic sets. Neutrosophic Sets and Systems 2020, 38, 1-13.

[3] Ssentongo, P.; Ssentongo, A.E.; Heilbrunn, E.S.; Ba, D.M.; Chinchilli, V.M. Association of cardiovascular disease and other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis. PLoS ONE 2021, 16, e0248731, doi:10.1371/journal.pone.0248731.

[4] Yan, L.; Zhang, H.T.; Goncalves, J.; Xiao, Y.; Wang, M.; Guo, Y.; Sun, C.; Tang, X.; Jing, L.; Zhang, M.; et al. An interpretable mortality prediction model for COVID-19 patients. Nature Machine Intelligence 2020, 2, 283–288, doi:10.1038/s42256-020-0180-7.

[5] Shahid, F.; Zameer, A.; Muneeb, M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos, Solitons & Fractals 2020, 140, 110212, doi:10.1016/j.chaos.2020.110212.

[6] Karaismailoglu, E.; Karaismailoglu, S. Two novel nomograms for predicting the risk of hospitalization or mortality due to COVID-19 by the naïve Bayesian classifier method. Journal of Medical Virology 2021, 93, 3194–3201, doi:10.1002/jmv.26871.

[7] Cevallos-Torres, L.; Botto-Tobar, M. Case study: Probabilistic estimates in the application of inventory models for perishable products in SMEs. In Problem-Based Learning: A Didactic Strategy in the Teaching of System Simulation; Springer: Cham, Switzerland, 2019; pp. 123–132.

[8] Zhang, F.; Lau, R.I.; Liu, Q.; Su, Q.; Chan, F.K.; Ng, S.C. Gut microbiota in COVID-19: key microbial changes, potential mechanisms and clinical applications. Nature Reviews Gastroenterology & Hepatology 2023, 20, 323–337, doi:10.1038/s41575-023-00748-9.

[9] Ivanov, A. Decision trees for evaluation of mathematical competencies in the higher education: A case study. Mathematics 2020, 8, 748, doi:10.3390/math8050748.

[10] Smarandache, F. Introduction to Neutrosophic Statistics; Educational Publisher: Columbus, OH, USA, 2014.

[11] Guijarro-Rodríguez, A.A.; Witt-Rodríguez, P.M.; Cevallos-Torres, L.J.; Contreras-Puco, S.F.; Ortiz-Zambrano, M.C.; Torres-Martínez, D.E. Image segmentation techniques application for the diagnosis of dental caries. In Proceedings of the International Conference on Advances in Emerging Trends and Technologies; Springer: Cham, Switzerland, 2019; pp. 312–322.

[12] Iba, Y. Population monte carlo algorithms. Transactions of the Japanese Society for Artificial Intelligence 2001, 16, 279–286, doi:10.1527/tjsai.16.279.

[13] Alava, M.V.; Figueroa, S.P.D.; Alcivar, H.M.B.; Vázquez, M.L. Single valued neutrosophic numbers and analytic hierarchy process for project selection. Neutrosophic Sets and Systems 2018, 21, 122–130.

[14] Garai, T.; Garg, H.; Roy, T.K. A ranking method based on possibility mean for multi-attribute decision making with single valued neutrosophic numbers. Journal of Ambient Intelligence and Humanized Computing 2020, 11, 5245–5258, doi:10.1007/s12652-020-01934-2.

[15] Vega Falcón, V.; Vasallo Villalonga, Y.; Cevallos-Torres, L. Hybrid Neutrosophic Multi-Criteria Decision Model for Guest Selection in Collaborative Tourism Platforms. Neutrosophic Sets and Systems 2025, 84, 18.

[16] Cevallos-Torres, L.; Botto-Tobar, M. Pseudo-random numbers and congruential methods. In Problem-Based Learning: A Didactic Strategy in the Teaching of System Simulation; Springer: Cham, Switzerland, 2019; pp. 33–58.

[17] Shanbehzadeh, M.; Kazemi-Arpanahi, H.; Nopour, R. Performance evaluation of selected decision tree algorithms for COVID-19 diagnosis using routine clinical data. Medical Journal of the Islamic Republic of Iran 2021, 35, 29, doi:10.47176/mjiri.35.29.

[18] Azeraf, E.; Monfrini, E.; Pieczynski, W. Using the Naive Bayes as a discriminative model. In Proceedings of the 13th International Conference on Machine Learning and Computing, Shenzhen, China, 26-28 February 2021; pp. 106–110, doi:10.1145/3459400.3459419.

[19] Zhao, L.; Lee, S.; Jeong, S.P. Decision tree application to classification problems with boosting algorithm. Electronics 2021, 10, 1903, doi:10.3390/electronics10161903.

[20] Cook, R.; Lapeyre, J.; Ma, H.; Kumar, A. Prediction of compressive strength of concrete: critical comparison of performance of a hybrid machine learning model with standalone models. Journal of Materials in Civil Engineering 2019, 31, 04019255, doi:10.1061/(ASCE)MT.1943-5533.0002931.

[21] Pérez, J.M.; Martin, P.P. ROC curve. SEMERGEN - Medicina de Familia 2023, 49, 101821, doi:10.1016/j.semerg.2022.101821.

Received: June 03, 2025. Accepted: July 14, 2025.