NSS
Neutrosophic Sets and Systems, Vol. 82, 2025
University of New Mexico

Early Parkinson's Disease Detection and Classification using Machine Learning and Neutrosophic Sets

Hasan H. Oudah[1], Ahmed A. Metwaly[2], Mohamed eassa[3,4], Ahmed Abdelhafeez[3,4], Ahmad M. Nagm[5], Ahmed S. Salama[5]

[1]College of Arts / Department of Journalism, Ahl Al Bayt University, Karbala, Iraq, hasan.hadi@abu.edu.iq

[2]Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Egypt, a.metwaly23@fci.zu.edu.eg

[3]Computer Science Department, Faculty of Information System and Computer Science, October 6 University, Giza, 12585, Egypt mohamed.eassa.cs@o6u.edu.eg; aahafeez.scis@o6u.edu.eg

[4]Applied Science Research Center. Applied Science Private University, Amman, Jordan

[5]Department of Computer Engineering and Electronics, Cairo Higher Institute for Engineering, Computer Science and Management, New Cairo, Egypt ahmadnagm@alazhar.edu.eg, A.salama@chi.edu.eg

Abstract:

Optimizing therapy and rehabilitation for Parkinson's disease (PD) requires early identification and precise evaluation of the illness's course. However, there is disagreement about the best way to use gait analysis to categorize the severity of motor symptoms and identify early-stage Parkinson's disease. The precision of machine learning (ML) models in identifying early and intermediate stages of Parkinson's disease was assessed in this study. Six ML models are used in this study for the prediction of PD. Different metrics are used to evaluate ML models such as accuracy, precision, recall, f1-score, and AUC score. Then we propose a multi-criteria decision-making methodology (MCDM) to evaluate the ML models and select the best one. Two MCDM methods are used such as CRITIC method to compute the criteria weights and the TOPSIS method to rank the alternatives. These methods are used under the bipolar neutrosophic sets (BNSs) to deal with uncertainty information. The results show the support vector machine is the best ML model for the prediction of PD.

Keywords: Parkinson's Disease; Machine Learning; Neutrosophic Numbers; Early Detection; Uncertainty.

---

1. Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that primarily affects motor functions due to the gradual loss of dopamine-producing neurons in the brain. Early diagnosis is crucial for managing symptoms and slowing disease progression, yet detecting PD in its initial stages remains a significant challenge. Traditional diagnostic methods rely on clinical assessments and neurological examinations, which may not always be accurate or feasible for

detecting the disease before visible motor symptoms appear. This limitation has driven the need for advanced computational techniques, particularly machine learning (ML) and artificial intelligence (AI), to aid in the early detection and classification of Parkinson's Disease[1], [2].

Emerging research has explored various biomarkers for PD detection, including voice abnormalities, handwriting patterns, gait disturbances, and neuroimaging data. These biomarkers provide valuable information that, when analyzed using computational models, can significantly improve diagnostic accuracy. Speech-based assessments, for example, have gained considerable attention due to their non-invasive nature and ease of implementation. Individuals with early-stage PD often exhibit subtle voice changes, such as reduced vocal intensity and increased speech irregularities, which can be quantified using ML algorithms[3], [4].

Machine learning models such as Support Vector Machines (SVM), Random Forest, and deep learning techniques like Convolutional Neural Networks (CNNs) have shown promise in detecting early-stage PD by analyzing structured patient data. These models are trained on datasets containing clinical features, speech patterns, or movement data, allowing them to identify hidden patterns indicative of the disease. Additionally, ensemble learning and hybrid approaches have further enhanced classification accuracy by combining multiple models' strengths to improve predictive performance[5], [6].

The integration of wearable sensors and mobile health technologies has also revolutionized Parkinson's detection by enabling continuous and real-time monitoring of patients. Smart devices equipped with accelerometers and gyroscopes can capture fine motor impairments, such as tremors and bradykinesia, which may go unnoticed in clinical settings. When combined with cloud-based AI analytics, these technologies provide a scalable and accessible approach for early PD detection, particularly in remote or underserved populations[7], [8].

Despite the promising advancements in AI-driven PD detection, several challenges remain. Computational complexity, data privacy concerns, and the need for high-quality labeled datasets are significant barriers to widespread clinical adoption. Moreover, ensuring model interpretability and reliability is critical, as clinicians must trust and understand AI-driven decisions. Future research must focus on refining existing models, incorporating multi-modal data fusion techniques, and validating AI-driven approaches through extensive clinical trials to ensure robust and accurate detection[9].

Early detection and classification of Parkinson's Disease using computational intelligence hold immense potential in transforming neurodegenerative disease diagnosis. By leveraging ML algorithms, wearable technology, and multi-modal data analysis, researchers can develop more effective diagnostic tools to identify PD at its earliest stages. While challenges persist, continued innovation in AI-driven healthcare solutions offers hope for improved patient outcomes and a better quality of life for those at risk of developing Parkinson's Disease[10], [11].

1.1 Neutrosophic Sets

Using a multi-criteria decision-making (MCDM) strategy to rank machine learning models offers a more realistic answer to this issue. When faced with the challenge of identifying or articulating preferences, as well as when decisions must be made based on several conflicting signs or elements that vie for significance, MCDM techniques are very helpful. Then, using the chosen criteria or alternatives, MCDM techniques assist in resolving conflicting issues and selecting the optimal answer. Choosing options with competing criteria is a difficulty that decision-makers encounter frequently. One specific MCDM issue that requires careful consideration of several conflicting intangible factors is the ranking of ML models[12], [13].

To put it briefly, MCDM techniques have been applied more and more to address a variety of decision-making issues, including ranking ML models. Few studies, meanwhile, have been able to develop a definitive MCDM approach, especially for ML model selection[14], [15].

The neutrosophic set is a subfield of neutrosophy, a field of philosophy that examines the nature, origin, and extent of neutralities as well as how they interact with other ideational spectra. Truth membership (T), indeterminacy membership (I), and falsity membership (F) are the three components of neutrosophic[16]. Every member value in the neutrosophic set is a real standard or non-standard. The single-valued neutrosophic set (SVNS), an extension of the neutrosophic set that takes the value from the subset of [0, 1], was then suggested by Wang et al. Numerous studies have focused specifically on the T, I, and F components, which ultimately led to the definition of a specific example of neutrosophic sets[17], [18].

2. Neutrosophic Numbers and Machine Learning

This section shows the steps of the neutrosophic approach to selecting the best ML models to be selected to predict Parkinson's disease. We use two MCDM methods such as CRITIC method to compute the criteria weights and the TOPSIS method to rank the ML models. These methods are used under the bipolar neutrosophic sets (BNSs). We show some definitions of BNSs such as:

Definition 1.

The bipolar neutrosophic sets (BNSs) can be defined as[19], [20]:

$$X = \left\{z, \left(T_X^+(z), I_X^+(z), F_X^+(z), T_X^-(z), I_X^-(z), F_X^-(z)\right) z \in Z\right\} \quad (1)$$

$$T_X^+(z), I_X^+(z), F_X^+(z): Z \to [0,1] \quad (2)$$

$$T_X^-(z), I_X^-(z), F_X^-(z): Z \to [-1,0] \quad (3)$$

Definition 2.

Let two bipolar neutrosophic numbers (BNNs) such as:

$X_1 = \{T_1^+(z), I_1^+(z), F_1^+(z), T_1^-(z), I_1^-(z), F_1^-(z)\}$ and $X_2 = \{T_2^+(z), I_2^+(z), F_2^+(z), T_2^-(z), I_2^-(z), F_2^-(z)\}$

$$X_1 \cup X_2 = \begin{pmatrix} \max\left(T_1^+(z), T_2^+(z)\right) \\ , \frac{I_1^+(z)+I_2^+(z)}{2}, \\ \min\left(F_1^+(z), F_2^+(z)\right), \min \left(T_1^-(z), T_2^-(z)\right) \\ , \frac{I_1^-(z)+I_2^-(z)}{2} \\ \max\left(F_1^-(z), F_2^-(z)\right) \end{pmatrix} \quad (4)$$

Definition 3.

We show operations of BNNs such as:

$$X_1 + X_2 = \begin{pmatrix} T_1^+(z) + T_2^+(z) - T_1^+(z)T_2^+(z), \\ I_1^+(z)I_2^+(z), \\ F_1^+(z)F_2^+(z), \\ -T_1^-(z)T_2^-(z), \\ -\begin{pmatrix} -I_1^-(z) - I_2^-(z) - \\ I_1^-(z)I_2^-(z) \end{pmatrix}, \\ -\begin{pmatrix} -F_1^-(z) - F_2^-(z) \\ -F_1^-(z)F_2^-(z) \end{pmatrix} \end{pmatrix} \quad (5)$$

$$X_1 X_2 = \begin{pmatrix} T_1^+(z)T_2^+(z), \\ I_1^+(z) + I_2^+(z) \\ -I_1^+(z)I_2^+(z) + \\ F_1^+(z) + F_2^+(z) \\ -F_1^+(z)F_2^+(z), \\ -\begin{pmatrix} -T_1^-(z) - T_2^-(z) \\ -T_1^-(z)T_2^-(z) \end{pmatrix}, \\ -I_1^-(z)I_2^-(z), \\ -F_1^-(z)F_2^-(z) \end{pmatrix} \quad (6)$$

$$\Xi X_1 = \begin{pmatrix} \left(1 - \begin{pmatrix} 1 - \\ T_1^+(z) \end{pmatrix}\right)^{\Xi}, \\ \left(I_1^+(z)\right)^{\Xi}, \\ \left(F_1^+(z)\right)^{\Xi}, \\ -\left(-\left(T_1^-(z)\right)^{\Xi}\right), \\ -\left(-\left(I_1^-(z)\right)^{\Xi}\right), \\ -\left(1 - \begin{pmatrix} 1 - \\ F_1^-(z) \end{pmatrix}\right)^{\Xi} \end{pmatrix} \quad (7)$$

$$a_1^{\Xi} = \begin{pmatrix} \left(T_1^+(z)\right)^{\Xi}, \\ \left(1-\left(1-I_1^+(z)\right)\right)^{\Xi}, \\ \left(1-\left(1-F_1^+(z)\right)\right)^{\Xi}, \\ -\left(1-\left(1-T_1^-(z)\right)\right)^{\Xi}, \\ -\left(-\left(I_1^-(z)\right)^{\Xi}\right) \\ -\left(-\left(F_1^-(z)\right)^{\Xi}\right) \end{pmatrix} \tag{8}$$

Definition 4.

The score function of BNSs is:

$$S(X_1) = \frac{\begin{pmatrix} T_1^+(z)+1-I_1^+(z) \\ +1-F_1^+(z)+ \\ 1+T_1^-(z) \\ -I_1^-(z)-F_1^-(z) \end{pmatrix}}{6} \tag{9}$$

We show the steps of the CRITIC method to compute the criteria weights.

Create the decision matrix.

The decision matrix is created as follows:

$$X = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & \ddots & \vdots \\ z_{m1} & \cdots & z_{mn} \end{bmatrix}_{m\times n} ; i = 1, \dots, m; j = 1, \dots, n \tag{10}$$

Normalize the decision matrix

We normalize the decision matrix for beneficial and non-beneficial criteria such as:

$$y_{ij} = \frac{x_{ij}-\min x_{ij}}{\max x_{ij}-\min x_{ij}} \tag{11}$$

$$y_{ij} = \frac{x_{ij}-\max x_{ij}}{\min x_{ij}-\max x_{ij}} \tag{12}$$

Determine the correlation between the criteria F_{jk}

Determine the C index

$$C_j = \xi_j \sum_{k=1}^{n}\left(1-F_{jk}\right) \tag{13}$$

Where ξ_j refers to the standard deviation

Compute the criteria weights

$$W_j = \frac{C_j}{\sum_{j=1}^{n} C_j} \tag{14}$$

Then we show the steps of the TOPSIS method to rank the alternatives. The TOPSIS method starts with the decision matrix.

Normalization decision matrix

$$a_{ij} = \frac{z_{ij}}{\sum_{i=1}^{i} z_{ij}^2} \tag{15}$$

Determine the weighted normalized matrix

$$d_{ij} = w_j a_{ij} \tag{16}$$

Determine ideal and non-ideal solution

$$L^+ = \left\{ \begin{matrix} \max d_{ij} \\ \left(\min d_{ij}\ \right) \end{matrix} \right\} \tag{17}$$

$$L^- = \left\{ \begin{matrix} \min d_{ij} \\ \left(\max d_{ij}\ \right) \end{matrix} \right\} \tag{18}$$

Determine the separation measures

$$B_i^+ = \sum_{j=1}^{n} \left(d_{ij} - d_i^+\right)^2 \tag{19}$$

$$B_i^- = \sum_{j=1}^{n} \left(d_{ij} - d_i^-\right)^2 \tag{20}$$

Determine the relative closeness values

$$S_i = \frac{B_i^-}{B_i^- + B_i^+} \tag{21}$$

3. Implementation

This section shows the implementation of the proposed approach to show the best ML models on the Parkinson dataset.

3.1 Description of the Dataset

The Parkinson dataset is gathered from the Kaggle website to be analyzed in this study. Table 1 shows some information on the dataset. This dataset has 24 features and 195 rows. This dataset has no missing values. We show some statistical analysis of this dataset as shown in Table 2.

Table 1. Some information from the dataset.

	phon_R01_S01_1	phon_R01_S01_2	phon_R01_S01_3	phon_R01_S01_4	phon_R01_S01_5
MDVP:Fo(Hz)	119.992	122.4	116.682	116.676	116.014
MDVP:Fhi(Hz)	157.302	148.65	131.111	137.871	141.781
MDVP:Flo(Hz)	74.997	113.819	111.555	111.366	110.655
MDVP:Jitter(%)	0.00784	0.00968	0.0105	0.00997	0.01284
MDVP:Jitter(Abs)	0.00007	0.00008	0.00009	0.00009	0.00011
MDVP:RAP	0.0037	0.00465	0.00544	0.00502	0.00655
MDVP:PPQ	0.00554	0.00696	0.00781	0.00698	0.00908
Jitter:DDP	0.01109	0.01394	0.01633	0.01505	0.01966

MDVP:Shimmer	0.04374	0.06134	0.05233	0.05492	0.06425
MDVP:Shimmer(dB)	0.426	0.626	0.482	0.517	0.584
Shimmer:APQ3	0.02182	0.03134	0.02757	0.02924	0.0349
Shimmer:APQ5	0.0313	0.04518	0.03858	0.04005	0.04825
MDVP:APQ	0.02971	0.04368	0.0359	0.03772	0.04465
Shimmer:DDA	0.06545	0.09403	0.0827	0.08771	0.1047
NHR	0.02211	0.01929	0.01309	0.01353	0.01767
HNR	21.033	19.085	20.651	20.644	19.649
status	1	1	1	1	1
RPDE	0.414783	0.458359	0.429895	0.434969	0.417356
DFA	0.815285	0.819521	0.825288	0.819235	0.823484
spread1	-4.81303	-4.07519	-4.44318	-4.1175	-3.74779
spread2	0.266482	0.33559	0.311173	0.334147	0.234513
D2	2.301442	2.486855	2.342259	2.405554	2.33218
PPE	0.284654	0.368674	0.332634	0.368975	0.410335

Table 2. Some statistics analysis of the Parkison disease dataset.

	count	mean	std	min	25%	50%	75%	max
MDVP:Fo(Hz)	195	154.2286	41.39007	88.333	117.572	148.79	182.769	260.105
MDVP:Fhi(Hz)	195	197.1049	91.49155	102.145	134.8625	175.829	224.2055	592.03
MDVP:Flo(Hz)	195	116.3246	43.52141	65.476	84.291	104.315	140.0185	239.17
MDVP:Jitter(%)	195	0.00622	0.004848	0.00168	0.00346	0.00494	0.007365	0.03316
MDVP:Jitter(Abs)	195	0.000044	0.000035	0.000007	0.00002	0.00003	0.00006	0.00026
MDVP:RAP	195	0.003306	0.002968	0.00068	0.00166	0.0025	0.003835	0.02144
MDVP:PPQ	195	0.003446	0.002759	0.00092	0.00186	0.00269	0.003955	0.01958
Jitter:DDP	195	0.00992	0.008903	0.00204	0.004985	0.00749	0.011505	0.06433
MDVP:Shimmer	195	0.029709	0.018857	0.00954	0.016505	0.02297	0.037885	0.11908
MDVP:Shimmer(dB)	195	0.282251	0.194877	0.085	0.1485	0.221	0.35	1.302
Shimmer:APQ3	195	0.015664	0.010153	0.00455	0.008245	0.01279	0.020265	0.05647
Shimmer:APQ5	195	0.017878	0.012024	0.0057	0.00958	0.01347	0.02238	0.0794
MDVP:APQ	195	0.024081	0.016947	0.00719	0.01308	0.01826	0.0294	0.13778
Shimmer:DDA	195	0.046993	0.030459	0.01364	0.024735	0.03836	0.060795	0.16942
NHR	195	0.024847	0.040418	0.00065	0.005925	0.01166	0.02564	0.31482
HNR	195	21.88597	4.425764	8.441	19.198	22.085	25.0755	33.047
status	195	0.753846	0.431878	0	1	1	1	1
RPDE	195	0.498536	0.103942	0.25657	0.421306	0.495954	0.587562	0.685151
DFA	195	0.718099	0.055336	0.574282	0.674758	0.722254	0.761881	0.825288
spread1	195	-5.6844	1.090208	-7.96498	-6.4501	-5.72087	-5.04619	-2.43403
spread2	195	0.22651	0.083406	0.006274	0.174351	0.218885	0.279234	0.450493
D2	195	2.381826	0.382799	1.423287	2.099125	2.361532	2.636456	3.671155
PPE	195	0.206552	0.090119	0.044539	0.137451	0.194052	0.25298	0.527367

Fig 1 shows the number of rows in each class of the dataset. The first class has 147 samples and the second class has 48 samples.

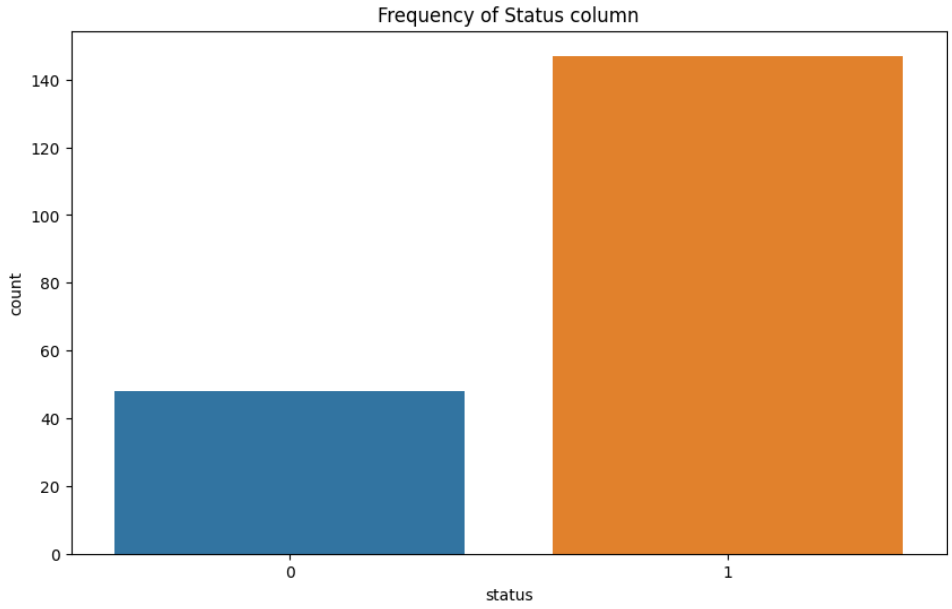

Fig 1. Samples of each class.

3.1 Metrics Measure

We can evaluate the ML models by different metrics evaluation such as accuracy, precision, recall, f1-score, and AUC score.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1-Score = 2 * \frac{Precision*Recall}{Precision+Recall}$$

TP refers to true positive, TN refers to true negative, FP refers to false positive, and FN refers to false negative.

4. Analysis

In this section, we apply six ML models on the Parkinson's disease dataset such as Logistic Regression (LR), k-nearest Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes (NB), Decision Tree (DT), and Random Forest. Table 3 shows the comparison analysis of the six ML models.

Table 3. Comparative analysis of ML models.

	Accuracy	Precision	Recall	F1-Score	AUC Score
LR	0.820512821	0.9	0.870968	0.885246	0.747984
KNN	0.769230769	0.923077	0.774194	0.842105	0.762097

SVM	0.8974359	0.88571	1	0.93939	0.75
NB	0.61538462	1	0.51613	0.68085	0.75806
DT	0.74358974	0.95652	0.70968	0.81481	0.79234
RF	0.82051282	0.9	0.87097	0.88525	0.74798

To show the best ML models we use the MCDM methodology under the BNSs to deal with uncertainty information. We use the CRITIC methodology to compute the criteria weights and select the best criterion. Then we use the TOPSIS methodology to select the best ML models based on the criteria weights. Fig 2 shows the criteria and alternatives of this study.

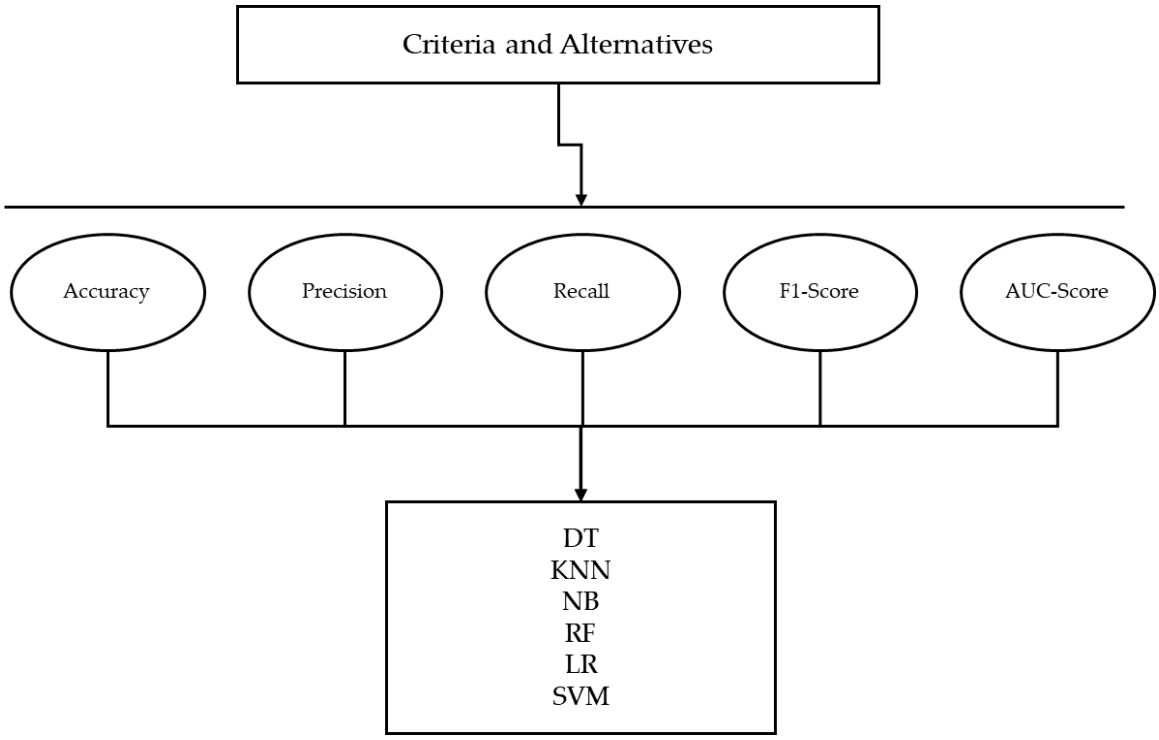

Fig 2. Set of criteria and alternatives.

Firstly, we apply the steps of the CRITIC methodology to compute the criteria weights, then we apply the steps of the TOPSIS methodology to show the best ML models.

Four experts and decision-makers are evaluating the criteria and alternatives to create the decision matrix using BNNs as shown in Table 4. Then we use Eq. (9) to obtain the crisp values. Then we combine the decision matrix.

Eq. (11) is used to normalize the decision matrix using equations. (11 and 12) as shown in table 5.

Eq. (13) is used to determine the correlation between the criteria F_{jk} as shown in table 6.

Then we determine the C index using eq. (13).

Then we compute the criteria weights using eq. (14) as shown in Fig 3.

Table 4. The decision matrix.

	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNA_1	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)
BNA_2	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)
BNA_3	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)
BNA_4	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)
BNA_5	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)
BNA_6	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)
	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNA_1	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)
BNA_2	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)
BNA_3	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)
BNA_4	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)
BNA_5	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)
BNA_6	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)
	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNA_1	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)
BNA_2	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)
BNA_3	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)
BNA_4	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)
BNA_5	(0.1,0.4,0.3,-0.1,-0.2,-0.3)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)
BNA_6	(0.1,0.4,0.4,-0.1,-0.3,-0.5)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.7,0.3,0.2,-0.4,-0.2,-0.1)
	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNA_1	(0.7,0.3,0.2,-0.4,-0.2,-0.1)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)
BNA_2	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.1,0.4,0.3,-0.1,-0.2,-0.3)
BNA_3	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)
BNA_4	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.6,0.4,0.4,-0.3,-0.2,-0.3)	(0.1,0.4,0.4,-0.1,-0.3,-0.5)
BNA_5	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)
BNA_6	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.4,0.1,0.4,-0.1,-0.2,-0.5)	(0.5,0.4,0.3,-0.4,-0.3,-0.3)	(0.4,0.3,0.3,-0.1,-0.2,-0.3)

Table 5. Normalized decision matrix.

	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNA_1	0.25	0.333333	0.125	0	0.444444
BNA_2	1	1	0.75	0.090909	0
BNA_3	0.5	0.166667	0	0.090909	0.444444
BNA_4	0	0.333333	0.375	0.454545	0.833333
BNA_5	0.416667	0	1	0.636364	0.833333
BNA_6	0.25	0.222222	0.375	1	1

Table 6. The correlation matrix.

	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNC_1	1	0.67387	0.3665	-0.39723	-0.80372
BNC_2	0.67387	1	0.1548	-0.44222	-0.78124
BNC_3	0.3665	0.1548	1	0.346198	0.030031
BNC_4	-0.39723	-0.44222	0.346198	1	0.836291
BNC_5	-0.80372	-0.78124	0.030031	0.836291	1

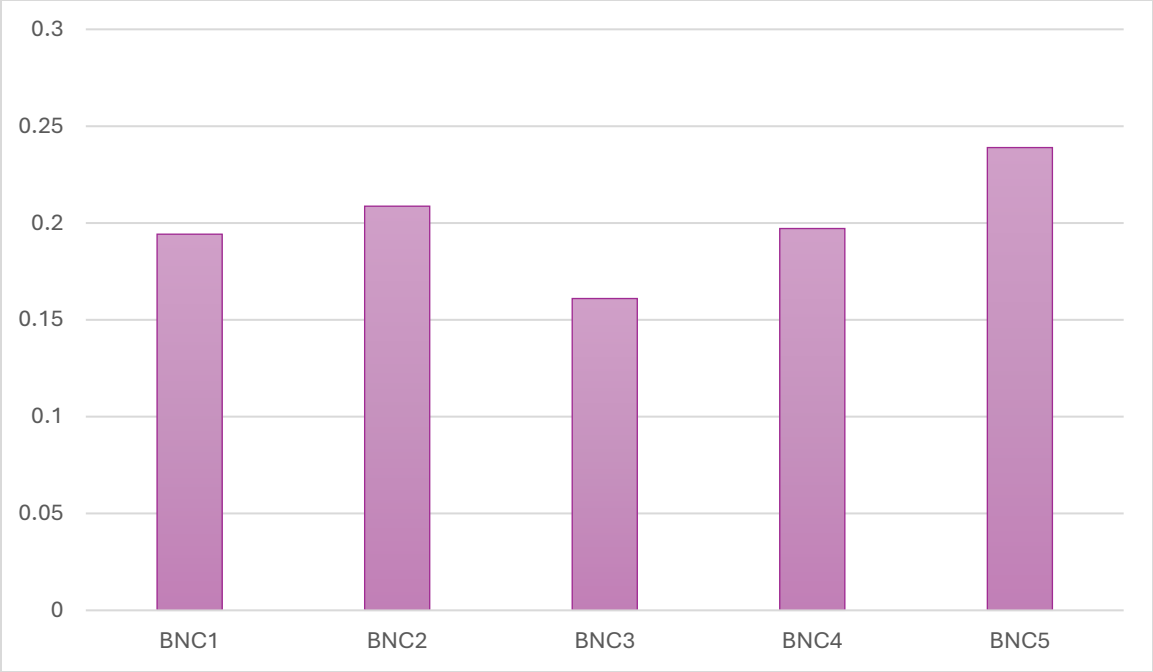

Fig 3. The importance of criteria.

Then we apply the steps of the TOPSIS methodology.

Eq. (15) is used for normalization decision matrix as shown in Table 7.

Eq. (16) is used to determine the weighted normalized matrix as shown in Table 8.

Then we determine the ideal and non-ideal solution

Then we determine the separation measures

Then we determine the relative closeness values using Eq. (21). Then we rank the alternatives as shown in Fig 4.

Table 7. Normalization matrix.

	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5
BNA_1	0.052573651	0.053082501	0.053284	0.052243	0.053038
BNA_2	0.056420504	0.058178421	0.055504	0.052682	0.049502
BNA_3	0.053855935	0.051808521	0.05284	0.052682	0.053038
BNA_4	0.051291367	0.053082501	0.054172	0.054438	0.056131
BNA_5	0.053428507	0.050534541	0.056392	0.055316	0.056131
BNA_6	0.052573651	0.052233181	0.054172	0.057072	0.057457

Table 8. Weighted decision matrix.

	BNC_1	BNC_2	BNC_3	BNC_4	BNC_5

BNA_1	0.010210779	0.011077008	0.008579	0.0103	0.012673
BNA_2	0.010957909	0.012140401	0.008936	0.010386	0.011829
BNA_3	0.010459822	0.01081116	0.008508	0.010386	0.012673
BNA_4	0.009961735	0.011077008	0.008722	0.010732	0.013413
BNA_5	0.010376808	0.010545311	0.009079	0.010905	0.013413
BNA_6	0.010210779	0.010899776	0.008722	0.011252	0.01373

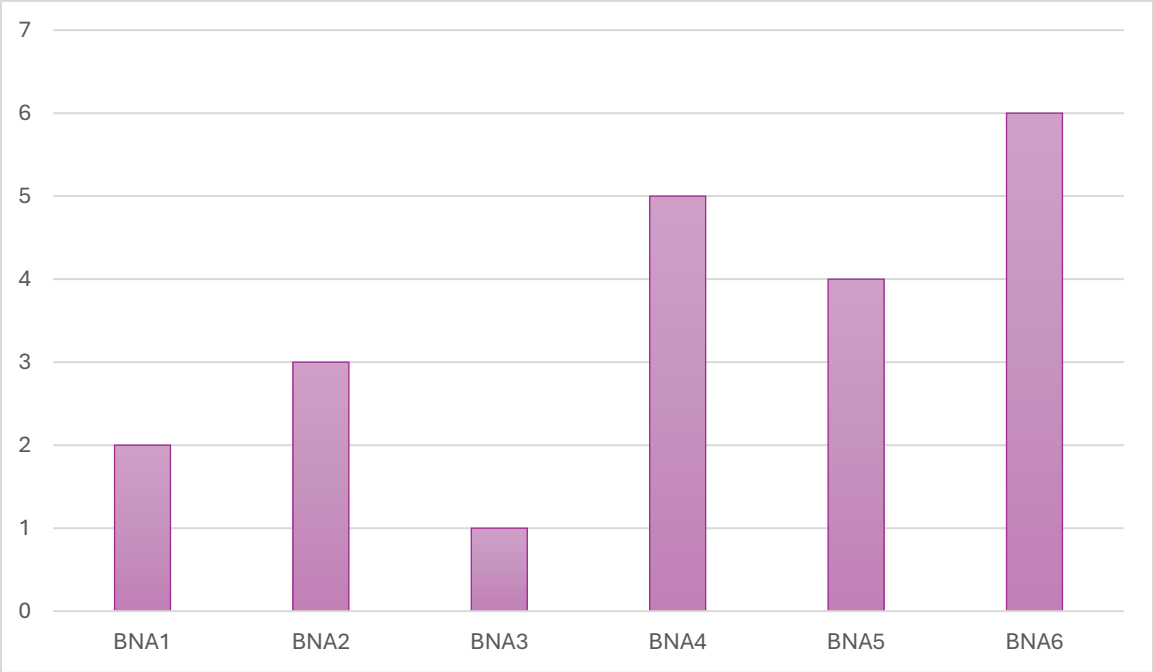

Fig 4. The rank of the alternatives.

5. Conclusions

This study proposed a ML model with the neutrosophic sets for PD prediction and classification. We used six ML models for the prediction of the PD such as LR, SVM, RF, DT, KNN, and NB. We applied these ML models to the PD dataset. The dataset has 24 features and 195 samples. These samples are divided into two classes. Then we used the MCDM methodology to select the best ML models based on the different metrics measures such as accuracy, precision, recall, f1-score, and AUC score. Two MCDM methods are used in this study, such as CRITIC method to compute the criteria weights and the TOPSIS method is used to rank the ML models. These methods are used under the bipolar neutrosophic sets to deal with uncertainty information. The results show the SVM is the best ML model in this study.

References

[1] A. Landolfi *et al.*, "Machine learning approaches in Parkinson's disease," *Curr. Med. Chem.*, vol. 28, no. 32, pp. 6548–6568, 2021.

[2] G. Pahuja and T. N. Nagabhushan, "A comparative study of existing machine learning approaches for Parkinson's disease detection," *IETE J. Res.*, vol. 67, no. 1, pp. 4–14, 2021.

[3] Z. K. Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Med. Hypotheses*, vol. 138, p. 109603, 2020.

[4] R. Prashanth, S. D. Roy, P. K. Mandal, and S. Ghosh, "High-accuracy detection of early Parkinson's disease through multimodal features and machine learning," *Int. J. Med. Inform.*, vol. 90, pp. 13–21, 2016.

[5] A. Govindu and S. Palwe, "Early detection of Parkinson's disease using machine learning," *Procedia Comput. Sci.*, vol. 218, pp. 249–261, 2023.

[6] K. N. R. Challa, V. S. Pagolu, G. Panda, and B. Majhi, "An improved approach for prediction of Parkinson's disease using machine learning techniques," in *2016 international conference on signal processing, communication, power and embedded system (SCOPES)*, IEEE, 2016, pp. 1446–1451.

[7] A. K. Tiwari, "Machine learning based approaches for prediction of Parkinson's disease," *Mach Learn Appl*, vol. 3, no. 2, pp. 33–39, 2016.

[8] J. M. Templeton, C. Poellabauer, and S. Schneider, "Classification of Parkinson's disease and its stages using machine learning," *Sci. Rep.*, vol. 12, no. 1, p. 14036, 2022.

[9] E. Celik and S. I. Omurca, "Improving Parkinson's disease diagnosis with machine learning methods," in *2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT)*, Ieee, 2019, pp. 1–4.

[10] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early detection of Parkinson's disease using deep learning and machine learning," *IEEE Access*, vol. 8, pp. 147635–147646, 2020.

[11] I. Mandal and N. Sairam, "New machine-learning algorithms for prediction of Parkinson's disease," *Int. J. Syst. Sci.*, vol. 45, no. 3, pp. 647–666, 2014.

[12] R. M. Zulqarnain, X. L. Xin, M. Saeed, F. Smarandache, and N. Ahmad, *Generalized neutrosophic TOPSIS to solve multi-criteria decision-making problems*, vol. 38. Infinite Study, 2020.

[13] J. Peng, J. Wang, and W.-E. Yang, "A multi-valued neutrosophic qualitative flexible approach based on likelihood for multi-criteria decision-making problems," *Int. J. Syst. Sci.*, vol. 48, no. 2, pp. 425–435, 2017.

[14] C. Kahraman and İ. Otay, *Fuzzy multi-criteria decision-making using neutrosophic sets*, vol. 16. Springer, 2019.

[15] J. Peng, J. Wang, H. Zhang, and X. Chen, "An outranking approach for multi-criteria decision-making problems with simplified neutrosophic sets," *Appl. Soft Comput.*, vol. 25, pp. 336–346, 2014.

[16] H. Zhang, J. Wang, and X. Chen, "An outranking approach for multi-criteria decision-

making problems with interval-valued neutrosophic sets," *Neural Comput. Appl.*, vol. 27, pp. 615–627, 2016.

[17] M. Mohamed and A. Elsayed, "A novel multi-criteria decision making approach based on bipolar neutrosophic set for evaluating financial markets in egypt," *Multicriteria Algorithms with Appl.*, vol. 5, pp. 1–17, 2024.

[18] V. Ulucay, I. Deli, and M. Şahin, "Similarity measures of bipolar neutrosophic sets and their application to multiple criteria decision making," *Neural Comput. Appl.*, vol. 29, pp. 739–748, 2018.

[19] I. Deli, M. Ali, and F. Smarandache, "Bipolar neutrosophic sets and their application based on multi-criteria decision making problems," in *2015 International conference on advanced mechatronic systems (ICAMechS)*, Ieee, 2015, pp. 249–254.

[20] V. Uluçay, A. Kilic, I. Yildiz, and M. Șahin, *A new approach for multi-attribute decision-making problems in bipolar neutrosophic sets*. Infinite Study, 2018.

Received: Oct 23, 2024. Accepted: March 24, 2025