






# Hybrid Model for Parkinson's Disease Prediction

Augusto Junio Guimarães<sup>1</sup> , Paulo Vitor de Campos Souza<sup>1,2</sup>  ,  
and Edwin Lughofer<sup>2</sup> 

<sup>1</sup> UNA University Center, Av. Gov. Valadares, 640 - Centro,  
Betim, MG 32510-010, Brazil  
augustojunioguimaraes@gmail.com

<sup>2</sup> Department Knowledge-Based Mathematical Systems,  
Johannes Kepler University Linz, Linz, Austria  
paulo.de\_campos\_souza@jku.at, Edwin.Lughofer@jku.at

**Abstract.** Parkinson's is a chronic, progressive neurological disease with no known cause that affects the central nervous system of older people and compromises their movement. This disorder can impair daily aspects of people and therefore identify their existence early, helps in choosing treatments that can reduce the impact of the disease on the patient's routine. This work aims to identify Parkinson's traces through a voice recording replications database applied to a fuzzy neural network to identify their patterns and enable the extraction of knowledge about situations present in the data collected in patients. The results obtained by the hybrid model were superior to state of the art for the theme, proving that it is possible to perform hybrid models in the extraction of knowledge and the classification of behavioral patterns of high accuracy Parkinson's.

**Keywords:** Parkinson's disease · Fuzzy neural network · Hybrid models

## 1 Introduction

Parkinson's disease - PD [12], progressively affects specific areas of the central nervous system composed of the brain and spinal cord, caused by the intense loss of nerve cells in parts of the basal ganglia known as the black substance, located mainly in a small region of the brain mass. These neurotransmitters are responsible for carrying out the voluntary movements of the body automatically, that is, all those in which there is no need to think to perform them, the muscles perform them to the presence of this substance in the brain. Dopamine is one of the main neurotransmitters in the basal ganglia, and its primary function is to intensify nerve impulses to the muscles [12]. In the absence of it, the individual's control is lost, causing characteristic signs and symptoms. The initial phase of the disease presents subtle symptoms such as variation in speech and other

notable ones, such as extensive and rhythmic tremors in the hands, stiffness in the muscles and joints making movement difficult, balance, and posture are also gradually compromised. PD is the second most common degenerative disease of the central nervous system, after Alzheimer's disease [6], so early diagnosis can provide the patient with a significant improvement in the quality of life and control of PD progression.

In order to carry out a diagnosis in an efficient and less invasive way to patients, methods that evaluate voice recordings are being used to perform an acoustic analysis on speech signals and tones to identify various diseases ([18, 21, 29, 33, 38]). Many works already consider this sound recording technique useful for the creation of intelligent systems or using smart approaches to diagnose the disease, and these approaches allow discriminating healthy people from those with PD ([2, 3, 7, 29, 30, 36, 41]). These structures facilitate the decision making of specialists in the correct diagnosis of patients, with a reduction in the risk of failures and a decrease in false-positive diagnoses, and taking into account the time provided to the patient in a medical follow-up. The classification is only possible by Naranjo et al. [28] that made available a database with 195 recordings consisting of people with PD and healthy people. This work proposes the use of a hybrid model, a Fuzzy Neural Network (FNN), in order to extract diffuse rules and maintain a high level of precision for the results obtained, this interpretability allows the extraction of knowledge from the database. The WEKA tool [14] presents the leading machine learning tools, and it will be possible to compare the results of the FNN network with the main machine learning techniques.

In addition to the introduction, this article presents the following sections: right after that, the related works (Sect. 2) and the main concepts that guide this research are present. In Sect. 3, concepts of fuzzy neural networks and the architecture used in this paper are presented to the reader. Also, in Sect. 4, Parkinson's detection tests are presented, and finally, in Sect. 5 it presents the conclusions obtained in the paper and future works that may expand what was accomplished in this paper.

## 2 Related Works

### 2.1 Parkinson's Disease

It can be said that the cause of PD is considered uncertain (first), with more than one factor involved in triggering the disease. These factors can be genetic or environmental. Studies indicate that due to the abnormal accumulation of synuclein (a protein in the brain that helps the communication of nerve cells). These neurotransmitters, called Lewy bodies, can accumulate in various regions of the brain, mainly in the black substance, and consequently interfere with brain function. When performing an early diagnosis, the impacts on quality of life are reduced by monitoring and treatment by specialists [39]. Medication and physical activities are the main assets for reducing the progression of PD.

## 2.2 Intelligent Approaches in Parkinson Disease Detection

The treatment and early identification of Parkinson's disease have stimulated several scientists to produce renowned academic works. In this paper, the database developed by Naranjo et al. [28] obtained excellent results using Bayesian models. Their studies generated other proposals and approaches with the same database. A summary of these works is listed in Table 1.

**Table 1.** Pattern Classification Models.

Models			
Applications	Accuracy	Sens	Spec.
Variable selection and classification approach for Parkinson's disease [29]	0.779 (0.080)	0.765 (0.135)	0.792 (0.150)
Bayesian approach [28]	0.752 (0.086)	0.786 (0.135)	0.718 (0.132)

## 3 Fuzzy Neural Networks

Fuzzy neural networks are an example of hybrid models that act in the synergy between interpretability (provided by fuzzy systems) and the ability to generalize training (provided by artificial neural networks) [24]. These models are seen as the union of a fuzzy inference system and a neural aggregation network responsible for carrying out actions of different natures, such as solving problems in the software area [8], astronomy [10], and the time series prediction [9,34]. It should be noted that these models have been working efficiently to become a reference in solving problems in the area of health and human behavior, such as in solving problems related to immunotherapy [20], breast cancer [32], ECG [23], autism detection [13], in addition to helping in the detection of cognitive and motor problems in children and adolescents [35]. The FNN model presented in this paper acts with three main layers, a fuzzification technique based on data density [17], training following the concepts of Extreme Learning Machine [16], and the classification of patterns performed by a singleton neuron that uses the ReLU [27] approach as a function of activation. Its architecture can be seen in Fig. 1, and its layers and training methodology are explained below.

### 3.1 First Layer- Data Density Fuzzification

The first layer of the model is responsible for the fuzzification process and the formation of Gaussian-type fuzzy neurons that will compose the model's input structure. All information will be fuzzified using a technique based on data clusters due to its density. Thus, the neurons formed in this layer, represent the data cluster and, in turn, assist in the construction of a more compact FNN architecture and with neurons more significant to the problem.

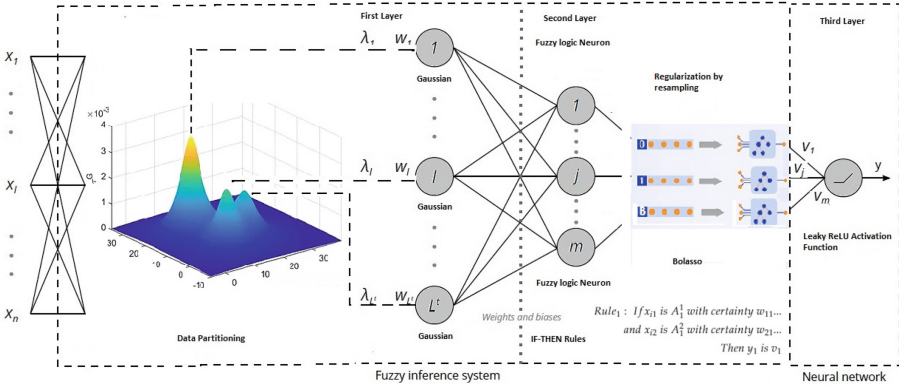


Fig. 1. Fuzzy neural network architecture

The objective of the clustering technique proposed by Hyde and Angelov [17] is to autonomously determine the clusters to be formed by the data, thus allowing the use of a simple fuzzification process and inspired by the behavior of the problem data. For this, the following concepts are fundamental, such as data distribution data density, radii and distances between reference points, among others.

The DDC works initially with the average of all the data to be evaluated. They are recursively calculated in order to decrease the complexity of the fuzzification approach and can be expressed by [17]:

$$\mu_0 = \frac{1}{N} \sum_{i=1}^N x_i \tag{1}$$

where  $x_i$  represents a sample from the problem, and  $N$  is the number of samples [17]. The next step is to calculate the sample density recursively and is expressed by [17]:

$$X_o = \frac{1}{N} \sum_{i=1}^N x_i^2 \tag{2}$$

The concepts reported in the Eqs. 1 and 2 provide the fundamental calculation in the construction of clusters. The density, as a result of the previous calculations, is also recursively calculated. Thus, it is expressed by [17]:

$$D_i = \frac{1}{1 + \|x_i - \mu_0\|^2 + X_0 - \|\mu_0\|^2} \tag{3}$$

Finally, in order to adjust the centers found in the grouping procedure, the following training algorithm is used [17]:

$$r_j^2 = \alpha r_0^2 + (1 - \alpha) \frac{1}{N_j} \sum_{i=1}^N \|x_i - \mu_j\|^2 \tag{4}$$

where  $\alpha$  it is seen as a learning parameter that can be set by the user or using data propagation. In this paper, the second option was defined so that the clustering approach was autonomous, based on the data essence.

Therefore, the clustering algorithm in this paper works by defining the sample with the highest density, transforming it into a center, and starting from there, defining the forms of clustering with the adjacent samples. Therefore, cluster C will have centers that will allow the construction of Gaussian neurons in the first layer of the model. They will be constructed based on the final centers found by the DDC approach and a sigma value defined in the unit interval, that is, between 0 and 1. For each input variable  $x_{ij}$ ,  $L$  neurons are defined  $A_{lj}$ ,  $l = 1, \dots, L$ , whose activation functions are composed of membership functions in the corresponding neurons. Therefore, it is represented by:

$$a_{jl} = \vartheta_{A_l}, \quad j = 1 \dots n, \quad l = 1 \dots L \tag{5}$$

where  $\vartheta$  is the membership degree. The number of Gaussian neurons in the first layer will be equal to the number of centers found by the fuzzification technique and these neurons can be expressed by:

$$gauss = (x_{ij}, c_l, \sigma_l) = \sum_{j=1}^n e^{-\frac{1}{2} \left( \frac{x_{ij} - c_l}{\sigma_l} \right)^2}, \quad for \quad l = 1 \dots L \tag{6}$$

### 3.2 Second Layer- Fuzzy Rules

The second layer is composed of fuzzy neurons capable of expressing knowledge relationships on the database through IF-THEN rules. For this purpose, fuzzy logical neurons are constructed through the aggregation of the Gaussian neurons of the first layer of the model. To perform the proper calculations, operators such as t-norm (product) and s-norm (probabilistic sum) [5] are applied to the aggregation operators, which in the specific case of this paper, is called uninorm [40].

The uninorm [40] is an aggregator that allows functions of t-norm and s-norm to be used within the same context, because depending on a term, sometimes the operator can perform calculations with the product, sometimes with the probabilistic sum. These factors facilitate the construction of more interpretable, contextual rules that represent the problem domain in a clear and precise way. The uninorm format used in this paper is defined by:

$$U(x, y) = \begin{cases} o T\left(\frac{x}{o}, \frac{y}{o}\right), & if \quad y \in [0, o] \\ o + (1 - o)S\left(\frac{x-o}{1-o}, \frac{y-o}{1-o}\right), & if \quad y \in [o, 1] \\ \delta(x, y), & otherside \end{cases} \tag{7}$$

and

$$\delta(x, y) = \begin{cases} max(x, y) & - if \quad g \in [0, 0.5] \\ min(x, y) & - if \quad g \in (0.5, 1] \end{cases}$$

where  $T$  is a *t-norm*,  $S$  is a *s-norm* and  $o$  is the activation of the fuzzy neuron randomly set between  $[0, 1]$ . In order for the uninorm to act on fuzzy neurons,

two preliminary steps are required. The first procedure is the transformation of the neuron input ( $a_i$ ) (which in the case of this paper is represented by the Gaussian neuron of the first layer) together with a weight value ( $w_i$ ), defined in a random between 0 and 1 with Eq. 7, for  $i = 1, \dots, L$ . The second step aggregates all the values resulting from the first step through the following Equation:

$$p(w, a, o) = wa + \bar{w}o \tag{8}$$

where  $\bar{w}$  represents the complement of  $w$ . The fuzzy neuron ( $\mathbf{z}$ ) that uses uninorm to aggregate neurons in the first layer with weights is called a unineuron [22] and can be described by:

$$z = UNI(w; a; o) = U_{i=1}^n p(w_i, a_i, o) \tag{9}$$

The  $\mathbf{z}$  neurons are seen as an aggregation of the first layer to obtain knowledge about the database. In this case, these neurons can be interpreted as a set of fuzzy rules of the IF-THEN type that can be expressed as:

$$\begin{aligned} Rule_z : IF x_{i1} \text{ is } A_1^1 \text{ with certainty } w_{11} \dots \\ THEN z_z = v_z \end{aligned} \tag{10}$$

The concept of certainty can also be extended to the rule consequent in Eq. 10.

$$THEN z = Class Y \text{ with certainty } w_Y \text{ (in } [0, 1]) \tag{11}$$

In binary problems, Eq. 11 could be seen as an auxiliary process to interpret fuzzy rules. Therefore, in a hypothetical example where  $z = Class 0$  with certainty 0.8 would mean that this rule has certainty 0.8 in *Class 0* and 0.2 in *Class 1*.

As the relationship of  $\mathbf{z}$  neurons occurs through the fuzzification process and the determination of centers, it is concluded that the model that makes up the first two layers of the model can be seen as a fuzzy inference system based on the density of the data evaluated in the model. Where the values of  $\mathbf{v}$  are defined by training the model and represent the weights that connect the fuzzy inference system with the third layer. Another view for the value of  $\mathbf{v}$  is that it represents the weight of that rule in the context of the expected output.

### 3.3 Third Layer- Artificial Neural Network

The third layer of the model merges the fuzzy rules and provides the expected answers to the problem. It is composed of an artificial neuron (can also be seen as a Singleton), which receives the  $\mathbf{z}$  neurons of the second layer as input and performs the due calculations with a set of weights  $\mathbf{v}$  obtained analytically. The neuron activation function, responsible for the necessary calculations for the final responses of the model, uses the Leaky ReLU approach [25]. Therefore, the neurons of the third layer can be expressed by:

$$y = \text{sign} \sum_{j=0}^l f_{\omega}(z_l, v_l) \tag{12}$$

where  $\omega$  is defined by the leaky ReLU activation function and the sign function represents the signal function. This activation function introduces a  $\beta$  factor to prevent neurons from being discarded when analyzing the problem. Its function is expressed by [25]:

$$f_{\text{LeakyReLU}}(\mathbf{z}, \beta) = \text{max}(\beta \mathbf{z}, \mathbf{z}) \tag{13}$$

where, in this paper,  $\beta = 0.001$ .

The process carried out in the third layer can also be seen as the stage of defuzzification of the model, allowing fuzzy relations to return to crisp values.

### 3.4 Training

The training of the model is based on the concepts of Extreme Learning Machine, where the parameters of the hidden layer are defined randomly, and the weights used by the neural aggregation network (third layer) are calculated analytically, through the Moore Penrose pseudoinverse. Thus, we can express the obtainment of the weights that connect the fuzzy inference system to the neural aggregation network by the Equation:

$$v = Z^+ y \tag{14}$$

where  $Z^+$  is the Moore Penrose pseudoinverse of  $\mathbf{Z}$ , which is defined by:

$$Z = \begin{bmatrix} \text{UNI}(w_1, a_1 + o_1) & \dots & \text{UNI}(w_l, a_1 + o_l) \\ \text{UNI}(w_1, a_2 + o_1) & \dots & \text{UNI}(w_l, a_2 + o_l) \\ \text{UNI}(w_1, a_n + o_1) & \dots & \text{UNI}(w_l, a_n + o_l) \end{bmatrix}_{N \times l} \tag{15}$$

The model proposed in this paper needs the initial radii value in the fuzzification process and the sigma value of neurons in the first layer. The other values are defined according to the training algorithm. That makes the model simple and easy to adapt to solve the Parkinson’s problem.

The fuzzification technique can generate some redundant neurons. To avoid this problem, a resampling technique linked to the LARS [11] model is used to select the best neurons, which consequently will be the best rules. This technique, called bolasso [4], was proposed by Bach and has been widely used in hybrid models for this purpose.

It combines several random replications with fuzzy rules and assesses their relevance to the model’s expected outputs. Thus it is possible to define a subgroup of fuzzy candidate rules. At each replication, a different number of candidate rules are selected, and a consensus threshold defines the final selection of fuzzy rules. For example, if in 16 replications with the base, four rules were the most significant in 50% of the replications (that is, 8), they are selected to compose the final model. This selection is made before the weights are

generated, so it is guaranteed that the weights that connect the cloudy inference system and the model's output layer are generated with the significant rules of the problem.

### 4 Parkinson's Detection Test

In order to verify the capacity of the model proposed in this paper, standard classification tests will be performed with a database that was worked on by Naranjo et al. [28]. Initially, the database has 48 features. However, for these experiments, the ID dimension, which is responsible for identifying the people in the experiment, was discarded. The collected records were 240, and the database is balanced, as there are 120 people identified with Parkinson's and 120 others without the disease. Figure 2 presents some of the dimensions present in the database, where the blue colors represent healthy people, and the red color identifies a person with Parkinson's.

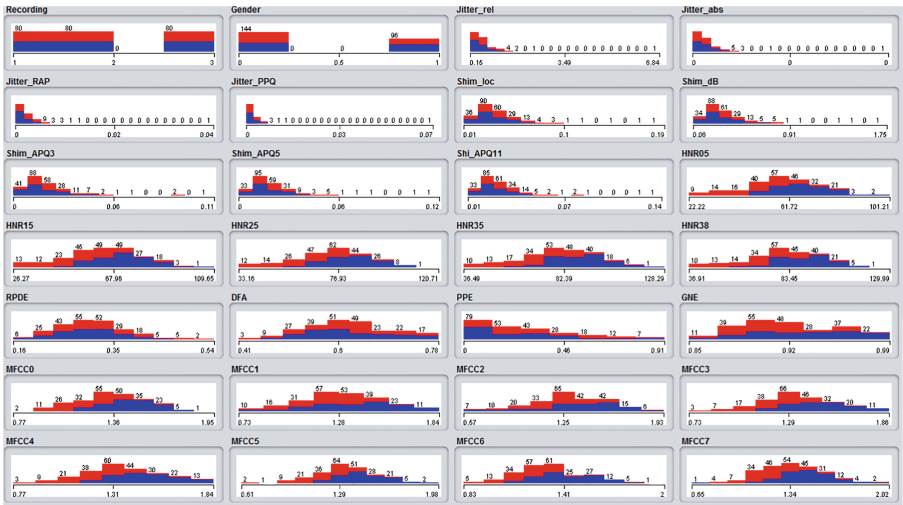


Fig. 2. Parkinson dataset- example (Color figure online)

To verify the model's ability to identify people with Parkinson's disease, the dataset will be divided into 70% for training the model, and the remaining 30% will be used to evaluate the results. All samples were normalized and were selected at random. To avoid trends in the test results, 30 repetitions were performed for the model that is part of the test.

As it is a database with binary outputs, the evaluation criteria of an intelligent model are well known by the academic community. The following criteria assess the model's ability to correct the diagnosis, the number of false positives and false negatives.



$$accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (16)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (17)$$

$$specificity = \frac{TN}{TN + TP} \quad (18)$$

$$AUC = \frac{1}{2} (sensitivity + specificity) \quad (19)$$

where,  $TP$  = true positive,  $TN$  = true negative,  $FN$  = false negative and  $FP$  = false positive.

The model proposed in this paper used the following values:  $\sigma = 1.9$  and  $radii = 0.08$  defined through a preliminary 10 k-fold procedure for the following value range of  $\sigma = [1.5, 1.6, 1.7, 1.8, 1.9, 2.0]$  and  $radii = [0.05, 0.06, 0.07, 0.08, 0.09]$ . For the experiments in this paper, the replication values were defined as 16, consensus threshold = 60% after cross-validation and 10 -fold tests with the database for training the model<sup>1</sup>. The other models used in the test are listed below:

**SVM** - Support vector machine algorithm [37] is to find a hyperplane in an  $N$ -dimensional space that distinctly classifies the data points.<sup>2</sup>

**MLP** - Multilayer Perceptron [26]. It uses training based on the backpropagation technique and has a hidden layer.<sup>3</sup>

**NB** - The Naive Bayes [19] algorithm is a probabilistic classifier based on the Bayes Theorem.<sup>4</sup>

**C4.5** - Generating a pruned or unpruned C4.5 decision tree [31].<sup>5</sup>

**RNT** - Random Tree [1] is use for constructing a tree that considers  $K$  randomly chosen attributes at each node.<sup>6</sup>

<sup>1</sup> For replication values between 8 and 32 and for the consensus threshold between 40 to 80%.

<sup>2</sup> In all tests that the model was used, one should consider the linear kernel. For more information on the code used, visit <https://la.mathworks.com/help/stats/support-vector-machines-for-binary-classification.html>.

<sup>3</sup> In all tests the following settings were used: batch size = 100, hidden layers = 1, learning rate = 0.3, momentum = 0.2, validation Threshold = 20.

<sup>4</sup> In all tests the following settings were used: batch size = 100, use kernet Estimator = false, Supervised Discretization = false.

<sup>5</sup> In all tests the following settings were used: batch size = 100, confidence factor = 0.25, min. num. obj = 2, num. Folds = 3, seed = 1, reduce Error pruning = false, unpruned = false, laplace = false.

<sup>6</sup> In all tests the following settings were used: batch size = 100, breakTiesRandomly = false, KValue = 0, num. Folds = 3, seed = 1, minNum = 1.0, allowUnclassifiedInstances = false, maxDepth = 0.

The values in parentheses are standard deviations. Simulations were performed on a Core (TM) 2 Duo CPU, 2.27 GHz, with 3-GB RAM. Time is represented by the sum of training time and test (seconds) in each of the models. Neurons represent the most representative neurons after the pruning or regularization of the models.

**Table 2.** Results Optical Interconnection database.

Models	Accuracy	AUC	Sens.	Spec.	Time
This Paper	<b>80.88 (4.43)*</b>	0.8204 (0.05)*	<b>0.8534 (0.65)*</b>	0.7875 (0.09)	10.11 (0.42)
SVM	79.93 (4.20)*	0.7969 (0.10)	0.8244 (0.04)*	0.7695 (0.06)	129.94 (16.22)
MLP	77.69 (4.95)*	0.8599 (0.04)*	0.7935 (0.07)	0.9263 (0.10)*	5.63 (0.06)
NB	78.51 (4.56)*	0.8693 (0.04)*	0.8213 (0.06)*	0.9173 (0.05)*	<b>0.01 (0.01)</b>
C-4.5	73.47 (4.97)	0.7386 (0.06)	0.7426 (0.08)	0.7346 (0.65)	<b>0.01 (0.01)</b>
RNT	79.81 (3.32)*	<b>0.8807 (0.04)*</b>	0.8343 (0.06)*	<b>0.9271 (0.21)*</b>	0.12 (0.01)

The results presented in the Table 2 present the total accuracy more significant than the other models used in the test and also surpass the original dataset. All test results were evaluated using a statistical test (ANOVA) [15]. With a 95% probability, we can say that all results marked with an ‘\*’ are statistically equal concerning the equitable performance of the factors collected between the models analyzed in the test. All premises (normality of residues, homoscedasticity, and independence) were not violated.

Therefore, it proves the efficiency of fuzzy neural networks in solving problems. It should be noted that the total fuzzy rules generated initially revolved around 46 fuzzy relations. After the regularization technique, around 5 to 6 rules are used in the 30 executions performed during the procedures reported in the tests. Another relevant factor reported in the results is that the Random Tree algorithm has results very close to the hybrid model, also standing out in AUC and specificity. The model proposed in this paper had better results in sensitivity (the ability of the diagnostic test to detect truly positive individuals, that is, to correctly diagnose patients with Parkinson’s disease). So we can say that the model has the best ability to identify people who have the disease. The following fuzzy rule was extracted from the dataset in an experiment with a final accuracy of 84% probability of defining a patient with a Parkinson’s diagnosis.

IF *Gender* is female with certainty 0.33 AND/OR *Jitterrel* is very high with certainty 0.90 AND/OR *Jitterabs* is very small with certainty 0.12 AND/OR *JitterRAP* is very small with certainty 0.81 AND/OR *JitterPPQ* is very small with certainty 0.12 AND/OR *Shimloc* is small with certainty 0.15 AND/OR *ShimdB* is very high with certainty 0.94 AND/OR *ShimAPQ3* is very small with certainty 0.04 AND/OR *ShimAPQ5* is very small with certainty 0.74 AND/OR *ShiAPQ11* is very small with certainty 0.73 AND/OR *HNR05* is very small with certainty 0.72 AND/OR *HNR15* is very small with certainty 0.90 AND/OR *HNR25* is very small with certainty 0.15 AND/OR *HNR35* is very small with certainty 0.83 AND/OR *HNR38* is very small with certainty 0.04 AND/OR *RPDE* is medium

with certainty 0.72 AND/OR *DFA* is medium with certainty 0.07 AND/OR *PPE* is medium with certainty 0.05 AND/OR *GNE* is very small with certainty 0.76 AND/OR *MFCC0* is medium with certainty 0.87 AND/OR *MFCC1* is medium with certainty 0.98 AND/OR *MFCC2* is small with certainty 0.65 AND/OR *MFCC3* is very small with certainty 0.32 AND/OR *MFCC4* is very small with certainty 0.03 AND/OR *MFCC5* is high with certainty 0.04 AND/OR *MFCC6* is small with certainty 0.65 AND/OR *MFCC7* is very high with certainty 0.45 AND/OR *MFCC8* is very high with certainty 0.62 AND/OR *MFCC9* is very small with certainty 0.05 AND/OR *MFCC10* is small with certainty 0.08 AND/OR *MFCC11* is very small with certainty 0.14 AND/OR *MFCC12* is medium with certainty 0.11 AND/OR *Delta0* is medium with certainty 0.10 AND/OR *Delta1* is extremely high with certainty 0.64 AND/OR *Delta2* is very high with certainty 0.77 AND/OR *Delta3* is very high with certainty 0.65 AND/OR *Delta4* is very small with certainty 0.75 AND/OR *Delta5* is small with certainty 0.66 AND/OR *Delta6* is very small with certainty 0.15 AND/OR *Delta7* is medium with certainty 0.54 AND/OR *Delta8* is medium with certainty 0.02 AND/OR *Delta9* is extremely high with certainty 0.32 AND/OR *Delta10* is very high with certainty 0.16 AND/OR *Delta11* is very high with certainty 0.34 AND/OR *Delta12* is extremely high with certainty 0.16 THEN *Status* is *Parkinson* with certainty 0.18.

## 5 Conclusion

The results of the tests carried out corroborate that the fuzzy neural network proposed in this paper can act efficiently in the identification of patients with Parkinson's. That corroborates the high accuracy of the model in identifying intricate patterns within a database, and the fuzzification technique based on the data essence demonstrates that the constructed fuzzy neurons can efficiently represent characteristics of the problem. This paper encourages new works to be elaborated for the construction of rules more representative of the Parkinson's problem, at the same time that it provides the evolution of comparative techniques using all dimensions of the problem. In future work, it is expected to expand the techniques to be used in hybrid models, such as new fuzzification, training, and defuzzification techniques.

**Acknowledgment.** The authors acknowledge the support by the Austrian Science Fund (FWF): contract number P32272-N38, acronym IL-EFS.

## References

1. Aldous, D.: The continuum random tree. I. Ann. Probab. 1–28 (1991)
2. Ali, L., Zhu, C., Zhou, M., Liu, Y.: Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection. Expert Syst. Appl. **137**, 22–28 (2019)

3. Arora, S., Baghai-Ravary, L., Tsanas, A.: Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice. *J. Acoust. Soc. Am.* **145**(5), 2871–2884 (2019)
4. Bach, F.R.: Bolasso: model consistent lasso estimation through the bootstrap. In: *Proceedings of the 25th International Conference on Machine Learning*, pp. 33–40 (2008)
5. Balasubramaniam, J., Rao, C.J.M.: On the distributivity of implication operators over T and S norms. *IEEE Trans. Fuzzy Syst.* **12**(2), 194–198 (2004)
6. Boller, F., Mizutani, T., Roessmann, U., Gambetti, P.: Parkinson disease, dementia, and Alzheimer disease: clinicopathological correlations. *Ann. Neurol.: Off. J. Am. Neurol. Assoc. Child Neurol. Soc.* **7**(4), 329–335 (1980)
7. Brabenec, L., Mekyska, J., Galaz, Z., Rektorova, I.: Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *J. Neural Transm.* **124**(3), 303–334 (2017)
8. de Campos Souza, P.V., Guimaraes, A.J., Araujo, V.S., Rezende, T.S., Araujo, V.J.S.: Incremental regularized data density-based clustering neural networks to aid in the construction of effort forecasting systems in software development. *Appl. Intell.* **49**(9), 3221–3234 (2019)
9. de Campos Souza, P.V., Torres, L.C.B.: Regularized fuzzy neural network based on or neuron for time series forecasting. In: Barreto, G.A., Coelho, R. (eds.) *NAFIPS 2018. CCIS*, vol. 831, pp. 13–23. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-95312-0\\_2](https://doi.org/10.1007/978-3-319-95312-0_2)
10. de Campos Souza, P.V., Torres, L.C.B., Guimarães, A.J., Araujo, V.S.: Pulsar detection for wavelets soda and regularized fuzzy neural networks based on and neuron and robust activation function. *Int. J. Artif. Intell. Tools* **28**(01), 1950003 (2019)
11. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., et al.: Least angle regression. *Ann. Stat.* **32**(2), 407–499 (2004)
12. Gelb, D.J., Oliver, E., Gilman, S.: Diagnostic criteria for Parkinson disease. *Arch. Neurol.* **56**(1), 33–39 (1999)
13. Guimarães, A.J., Araujo, V.J.S., Araujo, V.S., Batista, L.O., de Campos Souza, P.V.: A hybrid model based on fuzzy rules to act on the diagnosed of Autism in adults. In: MacIntyre, J., Maglogiannis, I., Iliadis, L., Pimenidis, E. (eds.) *AIAI 2019. IAICT*, vol. 559, pp. 401–412. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-19823-7\\_34](https://doi.org/10.1007/978-3-030-19823-7_34)
14. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. *ACM SIGKDD Explor. Newslett.* **11**(1), 10–18 (2009)
15. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, And Prediction*. SSS. Springer, New York (2009). <https://doi.org/10.1007/978-0-387-84858-7>
16. Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: theory and applications. *Neurocomputing* **70**(1–3), 489–501 (2006)
17. Hyde, R., Angelov, P.: Data density based clustering. In: *2014 14th UK Workshop on Computational Intelligence (UKCI)*, pp. 1–7. IEEE (2014)
18. Ishihara, L., Brayne, C.: A systematic review of depression and mental illness preceding Parkinson's disease. *Acta Neurol. Scand.* **113**(4), 211–220 (2006)
19. John, G.H., Langley, P.: Estimating continuous distributions in Bayesian classifiers. In: *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 338–345. Morgan Kaufmann Publishers Inc. (1995)

20. Junio Guimarães, A., Vitor de Campos Souza, P., Jonathan Silva Araújo, V., Silva Rezende, T., Souza Araújo, V.: Pruning fuzzy neural network applied to the construction of expert systems to aid in the diagnosis of the treatment of cryotherapy and immunotherapy. *Big Data Cogn. Comput.* **3**(2) (2019). <https://doi.org/10.3390/bdcc3020022>, <https://www.mdpi.com/2504-2289/3/2/22>
21. Karan, B., Sahu, S.S., Mahto, K.: Parkinson disease prediction using intrinsic mode function based features from speech signal. *Biocybern. Biomed. Eng.* **40**(1), 249–264 (2020)
22. Lemos, A., Caminhas, W., Gomide, F.: New uninorm-based neuron model and fuzzy neural networks. In: 2010 Annual Meeting of the North American Fuzzy Information Processing Society, pp. 1–6. IEEE (2010)
23. Lin, C.T., et al.: Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks. *IEEE Trans. Circuits Syst. I Regul. Pap.* **53**(11), 2469–2476 (2006)
24. Lin, C.T., Lee, C.: *Neural Fuzzy Systems: A Neuro-fuzzy Synergism to Intelligent Systems*. Prentice-Hall, Inc. (1996)
25. Maas, A.L., Hannun, A.Y., Ng, A.Y.: Rectifier nonlinearities improve neural network acoustic models. In: *Proceedings of ICML*, vol. 30, p. 3 (2013)
26. McClelland, J.L., Rumelhart, D.E., Group, P.R., et al.: *Parallel Distributed Processing*, vol. 2. MIT Press, Cambridge (1987)
27. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning (ICML 2010)*, pp. 807–814 (2010)
28. Naranjo, L., Pérez, C.J., Campos-Roca, Y., Martín, J.: Addressing voice recording replications for Parkinson's disease detection. *Expert Syst. Appl.* **46**, 286–292 (2016)
29. Naranjo, L., Pérez, C.J., Martín, J., Campos-Roca, Y.: A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications. *Comput. Methods Programs Biomed.* **142**, 147–156 (2017)
30. Nilashi, M., bin Ibrahim, O., Ahmadi, H., Shahmoradi, L.: An analytical method for diseases prediction using machine learning techniques. *Comput. Chem. Eng.* **106**, 212–223 (2017)
31. Quinlan, J.R.: *C4. 5: programs for machine learning*. Elsevier (2014)
32. Silva Araújo, V.J., Guimarães, A.J., de Campos Souza, P.V., Rezende, T.S., Araújo, V.S.: Using resistin, glucose, age and BMI and pruning fuzzy neural network for the construction of expert systems in the prediction of breast cancer. *Mach. Learn. Knowl. Extr.* **1**(1), 466–482 (2019). <https://doi.org/10.3390/make1010028>, <https://www.mdpi.com/2504-4990/1/1/28>
33. Smith, M.E., Ramig, L.O., Dromey, C., Perez, K.S., Samandari, R.: Intensive voice treatment in Parkinson disease: laryngostroboscopic findings. *J. Voice* **9**(4), 453–459 (1995)
34. Soares, E., Costa, P., Costa, B., Leite, D.: Ensemble of evolving data clouds and fuzzy models for weather time series prediction. *Appl. Soft Comput.* **64**, 445–453 (2018). <https://doi.org/10.1016/j.asoc.2017.12.032>, <http://www.sciencedirect.com/science/article/pii/S1568494617307573>
35. Souza, P.V.C., et al.: Using hybrid systems in the construction of expert systems in the identification of cognitive and motor problems in children and young people. In: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–6, June 2019. <https://doi.org/10.1109/FUZZ-IEEE.2019.8858906>
36. Vaiciukynas, E., Verikas, A., Gelzinis, A., Bacauskiene, M.: Detecting Parkinson's disease from sustained phonation and speech signals. *PloS One* **12**(10) (2017)

37. Vapnik, V.: *The Nature of Statistical Learning Theory*. Springer, Heidelberg (2013). <https://doi.org/10.1007/978-1-4757-3264-1>
38. Viswanathan, R., et al.: Complexity measures of voice recordings as a discriminative tool for Parkinson's disease. *Biosensors* **10**(1), 1 (2020)
39. Wood-Kaczmar, A., Gandhi, S., Wood, N.: Understanding the molecular causes of Parkinson's disease. *Trends Mole. Med.* **12**(11), 521–528 (2006)
40. Yager, R.R., Rybalov, A.: Uninorm aggregation operators. *Fuzzy Sets Syst.* **80**(1), 111–120 (1996)
41. Zhang, H.H., et al.: Classification of Parkinson's disease utilizing multi-edit nearest-neighbor and ensemble learning algorithms with speech samples. *Biomed. Eng. Online* **15**(1), 122 (2016)