Article

Not peer-reviewed version

Fuzzy Logic System for Classifying Multiple Sclerosis Patients: High, Medium or Low Responder to Interferon-Beta

Edgar Ponce de Leon-Sanchez *, Jorge Mendiola-Santibañez *, Omar Dominguez-Ramirez, Ana Herrera-Navarro, Alberto Vazquez-Cervantes, Hugo Jimenez-Hernandez, Horacio Senties-Madrid

Posted Date: 21 July 2023

doi: 10.20944/preprints202307.1478.v1

Keywords: fuzzy logic system; pipeline prediction model; multiple sclerosis

Article

Fuzzy Logic System for Classifying Multiple Sclerosis Patients: High, Medium or Low Responder to Interferon-Beta

Edgar R. Ponce de Leon-Sanchez [1,*,†], Jorge D. Mendiola-Santibañez [2,*,†], Omar A. Dominguez-Ramirez [3], Ana M. Herrera-Navarro [1], Alberto Vazquez-Cervantes [4], Hugo Jimenez-Hernandez [1] and Horacio Senties-Madrid [5]

1 Facultad de Informática, Universidad Autónoma de Querétaro, Querétaro 76230, Mexico; mherrera@uaq.mx
2 Facultad de Ingeniería, Universidad Autónoma de Querétaro, Querétaro 76010, Mexico; mendijor@uaq.mx
3 Centro de Investigación en Tecnologías de Información y Sistemas, Universidad Autónoma del Estado de Hidalgo, Pachuca 42039, Mexico; omar@uaeh.edu.mx
4 Centro de Ingeniería y Desarrollo Industrial, CIDESI, Querétaro 76125, México; alberto.vazquez@cidesi.edu.mx
5 Hospital HMG Coyoacán, CDMX 04380, Mexico; sentiesmadridh@gmail.com
* Correspondence: eponcedeleon13@alumnos.uaq.mx (E.R.P.d.L.-S.); mendijor@uaq.mx (J.D.M.-S.)
† These authors contributed equally to this work.

Abstract: Interferon-beta is one of the most widely prescribed disease-modifying therapies for multiple sclerosis patients. However, this treatment is only partially effective, and a significant proportion of patients do not respond to this drug. This paper proposes an alternative fuzzy logic system based on the opinion of a neurology expert to classify relapsing-remitting multiple sclerosis patients: high, medium, and low responder to interferon-beta. Also, a pipeline prediction model trained with biomarkers associated to interferon-beta response is proposed for predicting whether patients are potential candidates to be treated with this drug, in order to avoid ineffective therapies. The classification results shows that the fuzzy system presents a 100% efficiency compared with an unsupervised hierarchical clustering method (52%). So, the performance of the prediction model is evaluated, and a 0.8 testing accuracy is achieved. Hence, a pipeline model including data standardization, data compression, and a learning algorithm, can be a useful tool for getting reliable predictions about the response to interferon-beta.

Keywords: fuzzy logic system; pipeline prediction model; multiple sclerosis

1. Introduction

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system (CNS) [1]. Although MS can take several different forms, the most common type is relapsing-remitting MS (RRMS), characterized by alternating periods of remission and intensification of symptoms [2]. The etiology of MS can include several factors, such as genetic susceptibility, viral infections, and so on [3–5], which activate the immune system, generating immune dysregulation, and producing an immune attack against the myelin covering of the CNS [6].

Studies have shown that susceptibility to MS is genetically dependent [7], but the specific gene factors remain largely unknown. It is known that peripheral self-antigen-specific immune cells are activated during the antigen presentation process and they enter the CNS through the disrupted blood–brain barrier (BBB) [8]. The route of entry depends on the phenotype and activation state of T-cells. T-cells play important roles in cellular immunity [9]. T- cells are divided into helper T-cells (Th) and regulatory T-cells (Treg). Nishihara et al. [10] analysed the cellular mechanisms implicated in the migration of different CD4+ T helper cell subsets (Th1, Th2, Th17) through the BBB compared to the epithelial blood-cerebrospinal fluid barrier (BCSFB). After migration into the CNS, myelin-specific T-cells are re-activated by CNS-resident antigen-presenting cells (APCs), which leads to the growth

 doi:10.20944/preprints202307.1478.v1

or intensification of inflammatory response and demyelination [11]. Ma et al. [12] recapitulated the evolution of different T-cell subsets and their cytokines in the pathogenesis of MS.

The autoimmune etiology of MS has been the target of the therapeutic approach to patients. Treatment of MS can be divided into: treatment of MS symptoms, treatment of MS relapse, and treatment modifying disease progression. The main target of MS treatment is delaying the disease progression [13]. Interferon-beta (IFN-β) is one of the most widely prescribed disease-modifying therapies for RRMS patients. IFN-β has multiple pathways of action on immune system. IFN-β can inhibit the activated proliferation of T- cells, prevent the migration of activated immune cells through the BBB, also it inhibits the production of pro-inflammatory cytokines (e.g., IL-2, IL-12, IFN-γ), induces the increase in anti-inflammatory cytokines (e.g., IL-4, IL-5, IL-10 and TGF-β), and promotes re-myelination in the CNS [14,15]. IFN-β can also prevent the differentiation of inflammatory Th1/Th17 cells and change the phenotype of Th cells from inflammatory Th1 to anti-inflammatory Th2 cells. Studies have shown that IFN-β can significantly improve the clinical symptoms of patients, reduce the annual recurrence rate, and delay the progress of the disease [16]. However, IFN-β is only partially efficient, and a significant proportion of patients with MS do not respond to this treatment, with the proportion of non-responders ranging from 20 to 50% [17]. Hence, we propose a pipeline model based on potential biomarkers associated with response to IFN-β to predict whether MS patients are potential candidates to be treated with this drug. Studies have researched the effect of gene polymorphisms on therapeutic response to IFN-β, which can affect the efficacy of this drug. Bustamante et al. [18] analysed the relationship between single nucleotide polymorphisms (SNPs) disposed in type I IFN-induced genes, genes becoming to the toll-like receptor (TLR) pathway, and genes encoding neurotransmitter receptors, and the response to IFN-β treatment in MS patients. From seven selected SNPs two polymorphisms were exposed to be related to IFN-β response: rs2277302 (PELI3) and rs832032 (GABRR3). Martinez et al. [19] evaluated the effect of polymorphisms in some genes (CD46, CD58, FHIT, IRF5, GAPVD1, GPC5, GRBRB3, MxA, PELI3, and ZNF697) on response to IFN-β treatment between RRMS patients.

Genome-wide researches generate in large numbers of data and there is a need for soft computing methods (SCMs) such as artificial neural networks, fuzzy systems, evolutionary algorithms, or metaheuristic and swarm intelligence algorithms, that can deal with this amount of data [20]. Studies only have focused on MS diagnosis applying fuzzy systems. Ayangbekun & Jimoh [21] designed a fuzzy inference system for diagnosing five brain diseases: Alzheimer, Creutzfeldt-Jakob, Huntington, MS, and Parkinson. Hosseini et al. [22] developed a clinical decision support system (CDSS) to help specialists diagnose MS with a relapsing-remitting phenotype. Matinfar et al. [23] proposed an expert system for MS diagnosis based on clinical symptoms and demographic characteristics. Studies have applied machine and deep learning techniques to detect biomarkers, and to analyse MS progress. Ali et al. [24] proposed two models to identify the biomarkers of two autoimmune diseases, MS and rheumatoid arthritis, through microRNA analysis. The proposed models include complete pipelines of text mining methods, composed by conventional machine learning (ML) methods, and LSTM deep learning (DL). Viatkin et al. [25] revealed a system to measure finger joint angles based on camera image, for tracing the motion and limits of hand mobility in MS. Convolutional neural networks (CNN) based on different architectures were used to analyze the information from the camera.

Despite, the studies presented above have shown the efficiency of IFN-β to improve the clinical symptoms of MS patients, a proportion of patients do not respond to this treatment. Studies have analyzed the genome-wide in order to identify genetic factors associated to the response to IFN-β treatment. Gurevich et al. [26] identified a subgroup of secondary progressive MS (SPMS) patients presenting a gene expression signature similar to that of RRMS patients who are clinical responders to IFN-β treatment. SPMS patients were classified using unsupervised hierarchical clustering according to IFN inducible gene expression profile identified in RRMS clinical responders to treatment. Although, the hierarchical clustering method is easy to implement, it rarely provides the best solution due to lots of arbitrary decisions. Clarelli et al. [27] detected genetic factors that affect the long-term response to

 doi:10.20944/preprints202307.1478.v1

IFN-β. The found pathways associated to inflammatory processes and presynaptic membrane, i. e., the genes related to glutamatergic system (GRM3 and GRIK2), play a potential role in the response to IFN-β. Jin et al. [28] implemented a feature selection method based on differentially correlated edges (DCE) to identify the most relevant genes associated to the response to IFN-β treatment in RRMS patients. Between 23 identified genes, seven had a confidence score >2: CXCL9, IL2RA, CXCR3, AKT1, CSF2, IL2RB, and GCA. Because the analyzed data were unlabeled, the responder category was restricted to patients whose first relapse time was more than five years (60 months), resulting in nine responders and nine non-responders. So, seven patients were excluded from analysis. Hence, we attempt to address some of the issues above. Our proposed solution consists of the following stages: 1) Collect gene expression profiles, demographic, and clinical characteristics, associated with the response to IFN-β treatment, 2) Classify RRMS patients using a fuzzy logic system, 3) Implement a pipeline model including data preprocessing, data compression, and a learning algorithm to make predictions, and 4) Evaluate the prediction performance.

The main contributions of this paper are as follows:

- An alternative fuzzy system based on expert knowledge with linguistic rules to classify RRMS patients: high, medium, or low responder to IFN-β treatment.
- A pipeline prediction model including a data preprocessing technique, a transformation technique for data compression, and a learning algorithm for making predictions on new data. The prediction model is trained with biomarkers associated to the IFN-β response for predicting whether MS patients are potential candidates to be treated with this drug, in order to avoid ineffective therapies.

This paper is organized as follows. Section 2 explains the proposed research strategy. Section 3 provides the experimental results. Section 4 discusses the proposed contributions. Section 5 presents the conclusions.

2. Materials and Methods

The strategy followed in this research is described in the flowchart of Figure 1, which divides the proposal into four stages.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

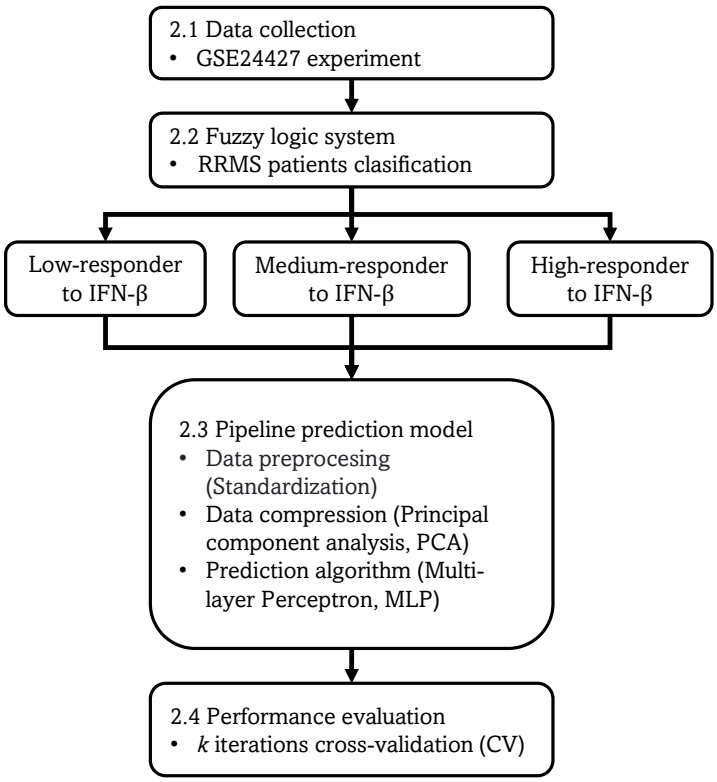

Figure 1. Proposed methodology. The gene data, demographic, and clinical characteristics are collected. Then, the RRMS patients are classified by the fuzzy logic system. A pipeline prediction model is implemented including data standardization, PCA for data compression, and MLP algorithm for making predictions. Finally, the k iterations CV is implemented for evaluating the model prediction performance.

2.1. Data Collection

The dataset is collected from the GSE24427 expression profiling by array experiment, available in the public repository of genomic data GEO [29]. Through GPL96 [HG-U133A] platform (Affymetrix Human Genome U133A Array), the genome-wide expression profiles of peripheral blood mononuclear cells from 25 RRMS patients were obtained. Patients were treated with subcutaneous IFN-beta-1b (Betaferon, Bayer Healthcare) at standard dose (250 µg every other day). Patient blood samples were drawn immediately before first, second, first month, 12th month, and 24th month IFN-β injection. The expression summary values were analyzed by GEO2R, an interactive web tool that allows viewing a specific gene expression through the profile graph tab. On the one hand, GPL96 platform lets to see demographic, and clinical characteristics of RRMS patients, which are used as input varibles for the proposed fuzzy system, and these are presented in Table 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 1. Demographic, and clinical characteristics.

Sample	Gender	Age	EDSS[1] 1st month	EDSS[1] 24th month
1	Female	63	4	5.5
2	Male	45	1	1
3	Female	25	1	1
4	Female	27	4	3.5
5	Female	51	3	2.5
6	Female	41	2	4.5
7	Female	44	4	3
8	Male	30	1.5	2
9	Female	26	4	3.5
10	Male	42	1	1
11	Male	29	2	2.5
12	Female	28	1.5	2.5
13	Female	48	1	1
14	Female	47	3.5	3
15	Female	42	2	3
16	Female	50	3.5	3.5
17	Male	37	1.5	4.5
18	Female	43	2	2
19	Male	54	3	2
20	Male	40	1	1
21	Female	48	2	2
22	Female	38	2	3
23	Male	18	1.5	2
24	Female	24	1	1
25	Male	38	1	1

[1] Expanded disability status scale.

On the other hand, through GPL96 platform, the expression values of 15 biomarkers associated with the response to IFN-β: IL-2, IL-12, IFN-γ, TNF-α, IL-4, IL-10, TGF-β, CD46, CD58, FHIT, IRF5, GAPVD1, GPC5, GRM3, and GRIK2 are collected, and integrated into an excel spreadsheet for training the proposed pipeline prediction model. For example, IL-2, IL-12, IL-4, and IL-10 cytokines expression values are displayed in Figure 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

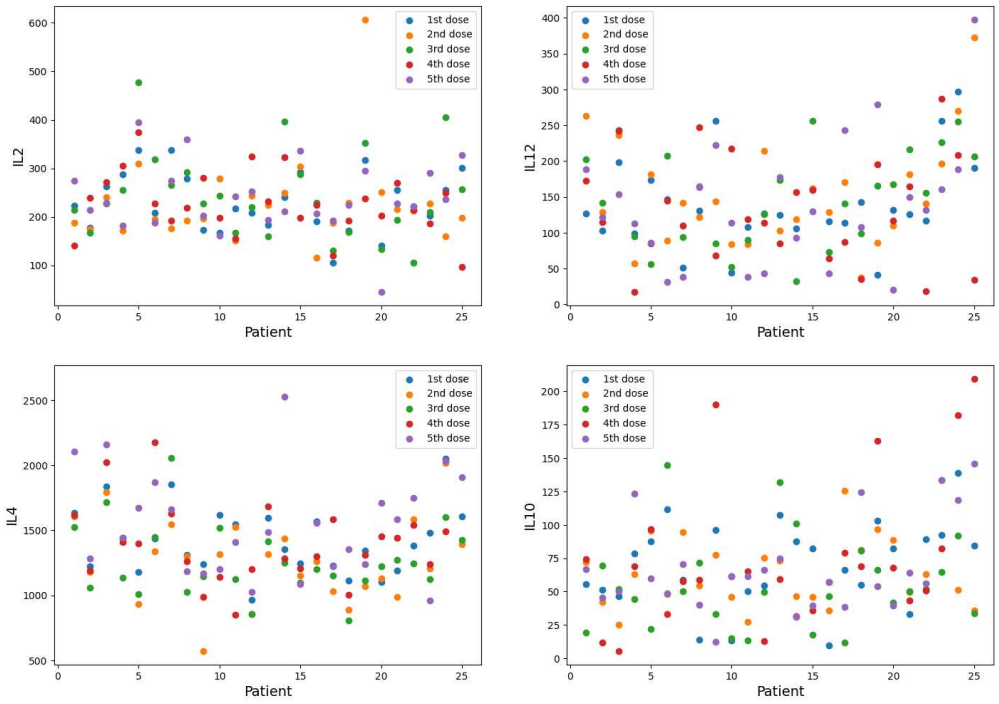

Figure 2. IL-2, IL-12, IL-4, and IL-10 cytokines. The expression values of 25 patients corresponding to five doses: before 1st, 2nd, 1st month, 12th month, and 24th month IFN-β injection.

2.2. *Fuzzy Logic System*

Fuzzy systems are structures based on fuzzy sets and fuzzy logic theories for processing inaccurate information [30]. Their main property includes symbolic knowledge representation in a form of fuzzy conditional (if-then) rules. The typical structure of a fuzzy system is described in Figure 3.

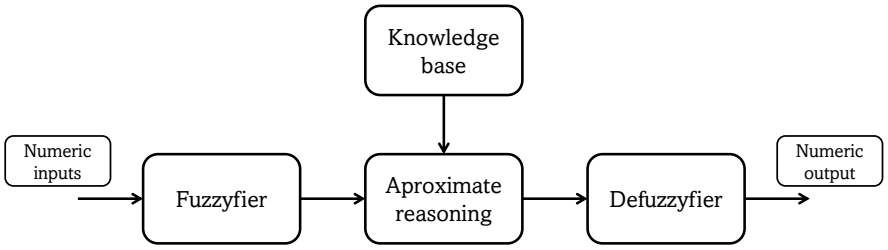

Figure 3. Fuzzy system structure.

The fuzzyfier maps the values of input variables $x_0 = [x_{01}, x_{02}, \ldots, x_{0N}]^T \in \mathbb{X} \subset \mathbb{R}^N$ into an N-dimensional fuzzy set A (linguistic values of the output variable) defined on the universe $\mathbb{X}$, by means of the approximate reasoning (inference engine) using expert knowledge, which is represented as a set of fuzzy conditional rules (knowledge base) [31]. The fuzzyfier can be defined as the membership function $\mu_A(x)$ of the fuzzy set A. Demographic, and clinical characteristics of RRMS patients are used as input variables for the fuzzyfier. The result of approximate reasoning is a fuzzy set $B(y)$, which can be related to a particular linguistic label. The defuzzyfier computes a representative numerical output $y_0 \in \mathbb{Y}$ from the result of fuzzy set $B(y)$ defined on the universe $\mathbb{Y}$. The numerical output is computed using the center of gravity (COG) method [32].

The structure of the proposed fuzzy system is based on the Mamdani–Assilan fuzzy system (MAFS) [33], which uses a set of conditional fuzzy rules of the canonical form of Equation 1, which can

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

be determined by a human expert. The proposed fuzzy system is designed through the Fuzzy Logic Designer App of MATLAB R2023a software.

$$\mathcal{R} = \{R^i\}_{i=1}^{N} = \{if\ and_{n=1}^{N}\ (X_n\ is\ L_{A_n}^{(i)}),\ then\ Y\ is\ L_B^{(i)}\}_{i=1}^{I}. \tag{1}$$

In this paper, we define the linguistic variables describing the demographic characteristics, including gender and age, and the clinical characteristics, including expanded disability status scale (EDSS) 1st month, and EDSS 24th month; $\mathcal{N}_1$ = "mean gender", $\mathcal{N}_2$ = "mean age", $\mathcal{N}_3$ = "mean EDSS 1st month", $\mathcal{N}_4$ = "mean EDSS 24th month". The sets of possible linguistic values are collections of different labels describing the gender, age, EDSS 1st month, and EDSS 24th month: L_1 = "female", "male", L_2 = "pediatric", "adult", "elderly", L_3 = "low", "medium", "high", and L_4 = "low", "medium", "high". To each one of the labels we assign the fuzzy sets $A_N^{(i)}$, defined on the universes $\mathbb{X}_N$, which represents the range of possible values, respectively. The whole description of the defined linguistic variables is presented in Table 2.

Table 2. Linguistic variables description.

Membership Function	Fuzzy Set	Universe of discourse	Parameters and Type
$\mu_{A_1^{(1)}}(gender)$	$A_1^{(1)}$	$\mathbb{X}_1$: [0 a 1]	Female: [-0.75 ; -0.083 ; 0.083 ; 0.75] Trapezoidal Male: [0.25 ; 0.916 ; 1.083 ; 1.75] Trapezoidal
$\mu_{A_2^{(1)}}(age)$	$A_2^{(1)}$	$\mathbb{X}_2$: [0 a 100] years	Pediatric: [-37.5 ; -4.167 ; 4.167 ; 37.5] Trapezoidal Adult: [8.333 ; 50 ; 91.666] Triangular Elderly: [62.5 ; 95.83 ; 104.2 ; 137.5] Trapezoidal
$\mu_{A_3^{(1)}}(EDSS^1\ 1st\ month)$	$A_3^{(1)}$	$\mathbb{X}_3$: [0 a 10] units	Low: [-3.75 ; -0.416 ; 1.0 ; 5.0] Trapezoidal Med: [1.0 ; 5.0 ; 9.0] Triangular High: [5.0 ; 9.0 ; 10.42 ; 13.75] Trapezoidal
$\mu_{A_4^{(1)}}(EDSS^1\ 24th\ month)$	$A_4^{(1)}$	$\mathbb{X}_4$: [0 a 10] units	Low: [-3.75 ; -0.416 ; 1.0 ; 5.0] Trapezoidal Med: [1.0 ; 5.0 ; 9.0] Triangular High: [5.0 ; 9.0 ; 10.42 ; 13.75] Trapezoidal

[1] Expanded disability status scale.

For example, the graphics of the membership functions $\mu_{A_2^{(1)}}(age)$, and $\mu_{A_4^{(1)}}(EDSS\ 24th\ month)$ of the fuzzy sets $A_2^{(1)}$, and $A_4^{(1)}$ are displayed at Figures 4 and 5, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

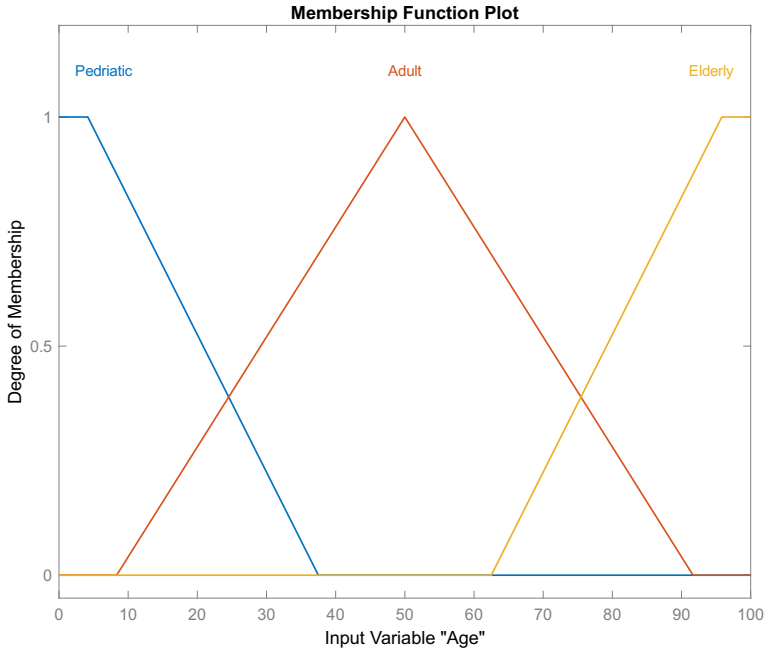

Figure 4. Set of linguistic values, which are three labels describing the "age" input variable, corresponding to fuzzy set $A_2^{(1)}$.

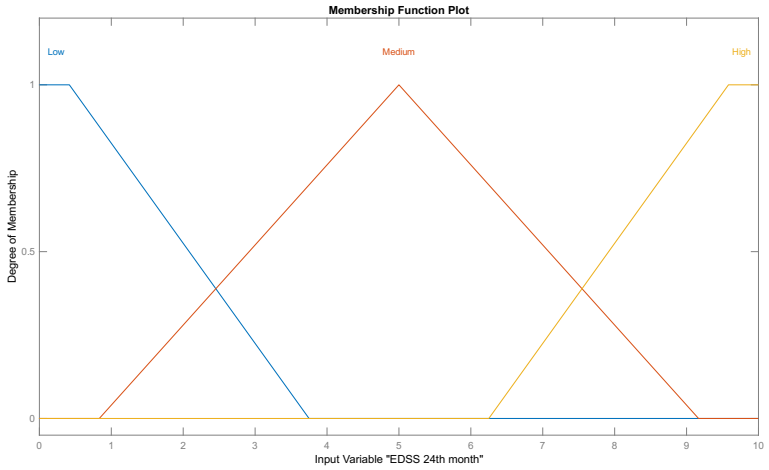

Figure 5. Set of linguistic values, which are three labels describing the "EDSS 24th month" input variable, corresponding to fuzzy set $A_4^{(1)}$.

The fuzzy conditional rules (knowledge base) are meant to decide about the influence of the input variables on response to IFN-β treatment. Tables 3, and 4 display the 36 defined rules according to the opinion of a neurology expert.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 3. Fuzzy rules definition (first part).

#	Rule
1	If gender is female and age is adult and EDSS 1st month is low and EDSS 24th month is low then response to IFNb is medium
2	If gender is female and age is adult and EDSS 1st month is medium and EDSS 24th month is medium then response to IFNb is medium
3	If gender is female and age is adult and EDSS 1st month is high and EDSS 24th month is high then response to IFNb is medium
4	If gender is female and age is elderly and EDSS 1st month is low and EDSS after 24th month is low then response to IFNb is medium
5	If gender is female and age is elderly and EDSS 1st month is medium and EDSS 24th month is medium then response to IFNb is medium
6	If gender is female and age is elderly and EDSS 1st month is high and EDSS 24th month is high then response to IFNb is medium
7	If gender is male and age is adult and EDSS 1st month is low and EDSS 24th month is low then response to IFNb is medium
8	If gender is male and age is adult and EDSS 1st month is medium and EDSS 24th month is medium then response to IFNb is medium
9	If gender is male and age is adult and EDSS 1st month is high and EDSS 24th month is high then response to IFNb is medium
10	If gender is male and age is elderly and EDSS 1st month is low and EDSS 24th month is low then response to IFNb is medium
11	If gender is male and age is elderly and EDSS 1st month is medium and EDSS 24th month is medium then response to IFNb is medium
12	If gender is male and age is elderly and EDSS 1st month is high and EDSS 24th month is high then response to IFNb is medium
13	If gender is female and age is adult and EDSS 1st month is low and EDSS 24th month is medium then response to IFNb is low
14	If gender is female and age is adult and EDSS 1st month is low and EDSS 24th month is high then response to IFNb is low
15	If gender is female and age is adult and EDSS 1st month is medium and EDSS 24th month is high then response to IFNb is low
16	If gender is female and age is elderly and EDSS 1st month is low and EDSS 24th month is medium then response to IFNb is low
17	If gender is female and age is elderly and EDSS 1st month is low and EDSS 24th month is high then response to IFNb is low
18	If gender is female and age is elderly and EDSS 1st month is medium and EDSS 24th month is high then response to IFNb is low

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 4. Fuzzy rules definition (second part).

#	Rule
19	If gender is male and age is adult and EDSS 1st month is low and EDSS 24th month is medium then response to IFNb is low
20	If gender is male and age is adult and EDSS 1st month is low and EDSS 24th month is high then response to IFNb is low
21	If gender is male and age is adult and EDSS 1st month is medium and EDSS 24th month is high then response to IFNb is low
22	If gender is male and age is elderly and EDSS 1st month is low and EDSS 24th month is medium then response to IFNb is low
23	If gender is male and age is elderly and EDSS 1st month is low and EDSS 24th month is high then response to IFNb is low
24	If gender is male and age is elderly and EDSS 1st month is medium and EDSS 24th month is high then response to IFNb is low
25	If gender is female and age is adult and EDSS 1st month is high and EDSS 24th month is medium then response to IFNb is high
26	If gender is female and age is adult and EDSS 1st month is high and EDSS 24th month is low then response to IFNb is high
27	If gender is female and age is adult and EDSS 1st month is medium and EDSS 24th month is low then response to IFNb is high
28	If gender is female and age is elderly and EDSS 1st month is high and EDSS 24th month is med then response to IFNb is high
29	If gender is female and age is elderly and EDSS 1st month is high and EDSS 24th month is low then response to IFNb is high
30	If gender is female and age is elderly and EDSS 1st month is medium and EDSS 24th month is low then response to IFNb is high
31	If gender is male and age is adult and EDSS 1st month is high and EDSS 24th month is medium then response to IFNb is high
32	If gender is male and age is adult and EDSS 1st month is high and EDSS 24th month is low then response to IFNb is high
33	If gender is male and age is adult and EDSS 1st month is medium and EDSS 24th month is low then response to IFNb is high
34	If gender is male and age is elderly and EDSS 1st month is high and EDSS 24th month is medium then response to IFNb is high
35	If gender is male and age is elderly and EDSS 1st month is high and EDSS 24th month is low then response to IFNb is high
36	If gender is male and age is elderly and EDSS 1st month is medium and EDSS 24th month is low then response to IFNb is high

2.3. Pipeline Prediction Model

A pipeline is a tool for setting a learning model including a data preprocessing technique, for instance standardization for feature scaling, a transformation technique, such as PCA for data compression, and a learning algorithm, like MLP for making predictions on new data. The structure of the proposed pipeline is shown in Figure 6.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

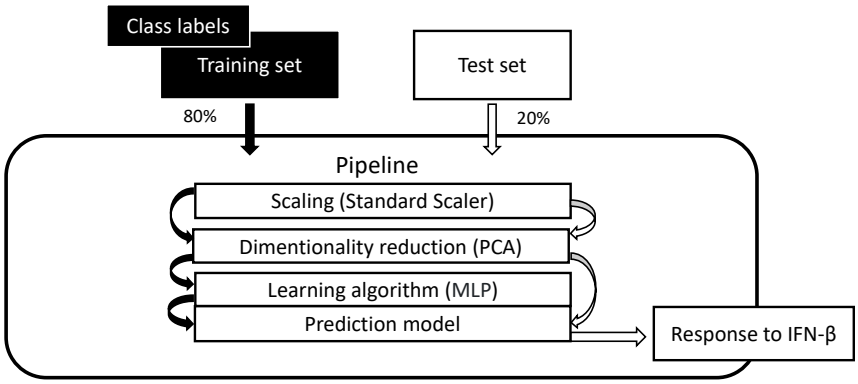

Figure 6. Structure of proposed pipeline model.

PCA is a technique of dimensionality reduction, which transforms data from a high-dimensional space to a space of lower dimensions. The dimension reduction is achieved by selecting the principal components (directions of maximum variance) as a basis set for the new space [34]. Applications of PCA include analysis of genome data and gene expression levels. For extracting the principal components, the data are standardized, then the covariance matrix is built to store the pairwise covariances between features. For example, the covariance between two features x_j and x_k can be computed by:

$$\sigma_{jk} = \frac{1}{n}\sum_{i=1}^{n}(x_j^{(i)} - \mu_j)(x_k^{(i)} - \mu_k) \tag{2}$$

where μ_j and μ_k are the representative samples of j and k features, respectively ($\mu_k, \mu_j = 0$, because the standardization). The eigenvectors of the covariance matrix represent the principal components and the eigenvalues define the magnitude of the eigenvectors, so the eigenvalues have to be ordered by decreasing the magnitude [35]. The ratio of a explained variance of an eigenvalue λ_j is the fraction of the eigenvalue and the total sum of the eigenvalues:

$$\frac{\lambda_j}{\sum_{j=1}^{d}\lambda_j}. \tag{3}$$

MLP is a supervised learning algorithm, which uses the backpropagation technique for learning. The structure of MLP consists of an input layer of neurons that receive the $X = x_1, x_2, \ldots, x_m$ sample inputs, one or more hidden layers of neurons that convert the values from the previous layer to a weighted linear summation $w_1x_1 + w_2x_2 + \ldots + w_mx_m$, followed by a non-linear activation function which is used to learn the weights, and then the output layer that predicts the class label of the samples [36]. During the learning stage, MLP compares the true class labels with the continuous output values of the non-linear activation function to compute the prediction error and update the weights. The hyperparameters of MLP are arbitrarily set as follows: solver='sgd', activation='tanh', and learning_rate_init=0.01.

2.4. Performance Evaluation

One of the key steps in building a ML or DL model is estimating its performance with new data. A model can suffer under fitting (high bias) if the model is too simple, or can suffer over fitting (high variance) if the model is too complex for the subjacent training data [35]. In order to get an acceptable bias-variance rate, the k iterations CV technique is implemented, which can obtain reliable estimates of the model's generalization performance.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

In the k iterations CV, the training dataset is randomly splitted into k iterations without replacement, where $k-1$ iterations are used for model training and one iteration for performance evaluation. This process is repeated k times to obtain k models and performance estimates. Then, the average performance of the models is computed based on the independent iterations to obtain a performance estimate E:

$$E = \frac{1}{k}\sum_{i=1}^{k} E_i. \tag{4}$$

Typically, the k iterations CV is used for model fitting, to find the optimal values of the hyperparameters that produce satisfactory generalization performance. Also, the confusion matrix (CM) is computed, which reports the count of the predictions of a classifier [37]: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

3. Results

In this paper, a fuzzy logic system based on MAFS was implemented to classify RRMS patients: high, medium and low responders to IFN-β treatment. Also, for comparison purposes, a hierarchical clustering technique is implemented to classify the same patients. After the dataset outputs were labeled, the gene features were used as training inputs for the proposed pipeline prediction model.

3.1. Fuzzy Logic System

At fuzzification stage, the membership values were computed for each one of the input variables. Tables 5, and 6 display the computed values of each membership function for all the samples.

Table 5. Fuzzification results (gender, and age).

Sample	$\mu_{Female}(Gender)$	$\mu_{Male}(Gender)$	$\mu_{Pediatric}(Age)$	$\mu_{Adult}(Age)$	$\mu_{Elderly}(Age)$
1	1	0	0	0.687	0.015
2	0	1	0	0.88	0
3	1	0	0.375	0.4	0
4	1	0	0.315	0.448	0
5	1	0	0	0.975	0
6	1	0	0	0.784	0
7	1	0	0	0.856	0
8	0	1	0.225	0.479	0
9	1	0	0.345	0.424	0
10	0	1	0	0.808	0
11	0	1	0.255	0.496	0
12	1	0	0.258	0.472	0
13	1	0	0	0.952	0
14	1	0	0	0.928	0
15	1	0	0	0.808	0
16	1	0	0	1	0
17	0	1	0.015	0.688	0
18	1	0	0	0.832	0
19	0	1	0	0.903	0
20	0	1	0	0.76	0
21	1	0	0	0.952	0
22	1	0	0	0.712	0
23	0	1	0.585	0.232	0
24	1	0	0.405	0.376	0
25	0	1	0	0.712	0

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 6. Fuzzification results (EDSS 1st month, and EDSS 24th month).

Sample	$\mu_{Low}(EDSS^{1})$	$\mu_{Med}(EDSS^{1})$	$\mu_{High}(EDSS^{1})$	$\mu_{Low}(EDSS^{2})$	$\mu_{Med}(EDSS^{2})$	$\mu_{High}(EDSS^{2})$
1	0.25	0.75	0.0	0.0	0.875	0.125
2	1.0	0.0	0.0	1.0	0.0	0.0
3	1.0	0.0	0.0	1.0	0.0	0.0
4	0.25	0.75	0.0	0.375	0.625	0.0
5	0.5	0.5	0.0	0.625	0.375	0.0
6	0.75	0.25	0.0	0.125	0.875	0.0
7	0.25	0.75	0.0	0.5	0.5	0.0
8	0.875	0.125	0.0	0.75	0.25	0.0
9	0.25	0.75	0.0	0.25	0.75	0.0
10	1.0	0.0	0.0	1.0	0.0	0.0
11	0.75	0.25	0.0	0.625	0.375	0.0
12	0.875	0.125	0.0	0.625	0.375	0
13	1.0	0.0	0.0	1.0	0.0	0.0
14	0.375	0.625	0.0	0.5	0.5	0.0
15	0.75	0.25	0.0	0.5	0.5	0.0
16	0.375	0.625	0.0	0.375	0.625	0.0
17	0.875	0.125	0.0	0.125	0.875	0.0
18	0.75	0.25	0.0	0.75	0.25	0.0
19	0.5	0.5	0.0	0.75	0.25	0.0
20	1.0	0.0	0.0	1.0	0.0	0.0
21	0.75	0.25	0.0	0.75	0.25	0.0
22	0.75	0.25	0.0	0.5	0.5	0.0
23	0.875	0.125	0	0.75	0.25	0.0
24	1.0	0.0	0.0	1.0	0.0	0.0
25	1.0	0.0	0.0	1.0	0.0	0.0

[1] 1st month. [2] 24th month.

At approximate reasoning stage, each one of the 36 inference rules from the knowledge base were evaluated with the obtained membership values from Tables 5, and 6. For example, considering the input values of #7 sample (gender: "female", age: "44", EDSS 1st month: "4", and EDSS 24th month: "3"), the inference engine results are shown in Table 7. In this case, only four rules: 1, 2, 13, and 27 had an inference result different of zero. Figure 7 displays the evaluation graph of previous inference rules.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 7. Inference results for the #7 sample.

#	Rule	Inference Engine
1	If gender is "female" and age is "adult" and EDSS 1st month is "low" and EDSS 24th month is "low" then response to IFNb is "medium"	min(1.0, 0.856, 0.25, 0.5) = 0.25
2	If gender is "female" and age is "adult" and EDSS 1st month is "medium" and EDSS 24th month is "medium" then response to IFNb is "medium"	min(1.0, 0.856, 0.75, 0.5) = 0.5
13	If gender is "female" and age is "adult" and EDSS 1st month is "low" and EDSS 24th month is "medium" then response to IFNb is "low"	min(1.0, 0.856, 0.25, 0.5) = 0.25
27	If gender is "female" and age is "adult" and EDSS 1st month is "medium" and EDSS 24th month is "low" then response to IFNb is "high"	min(1.0, 0.856, 0.75, 0.5) = 0.5

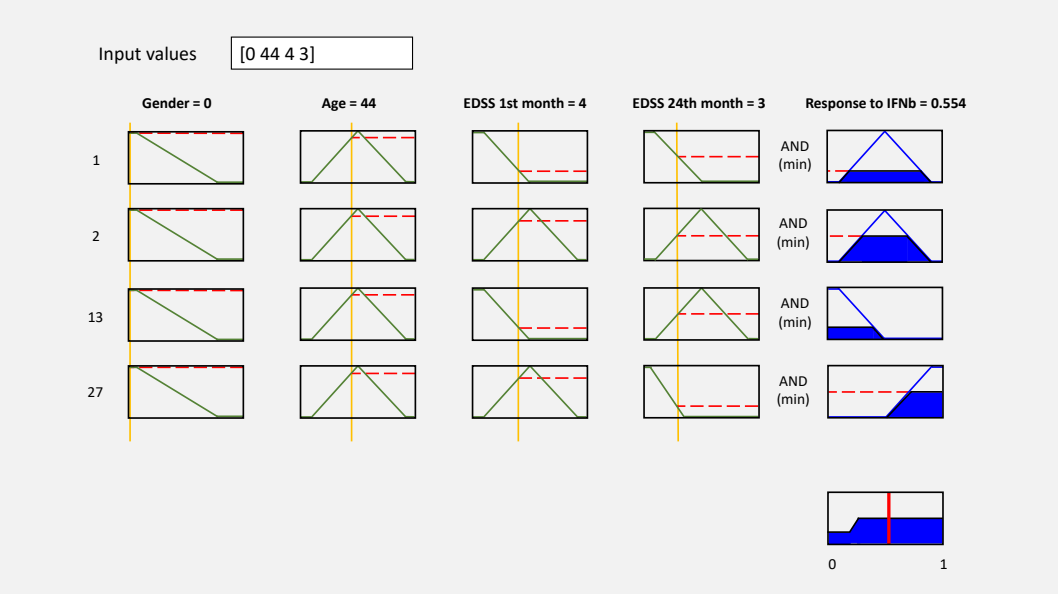

Figure 7. Evaluation graph of 1, 2, 13, and 27 inference rules.

Finally, at deffuzification stage, the numerical outputs were calculated using the COG method. A classification of high, medium, and low responder to IFN-β drug was carried out with three different methods: 1) Opinion of a neurology expert, 2) Proposed fuzzy system, and 3) Agglomerative clustering model. The results are displayed in Table 8.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 8. Classification of response to IFN-β. The resulting numerical values of defuzzification less than 0.5 are considered as low-responder (LR), equal to 0.5 as medium-responder (MR), and greater than 0.5 as high-responder (HR). For comparison purposes, the input data of Table 1 were preprocessed by Standard-Scaler technique, and they were used to train a prediction model of agglomerative clustering (n_clusters=3).

Sample	Expert opinion	Fuzzy System (Deffuzification)	Agglomerative Clustering
1	LR	0.459 $\Rightarrow$ LR	HR
2	MR	0.5 $\Rightarrow$ MR	LR
3	MR	0.5 $\Rightarrow$ MR	MR
4	HR	0.529 $\Rightarrow$ HR	HR
5	HR	0.527 $\Rightarrow$ HR	HR
6	LR	0.337 $\Rightarrow$ LR	HR
7	HR	0.554 $\Rightarrow$ HR	HR
8	LR	0.474 $\Rightarrow$ LR	LR
9	HR	0.53 $\Rightarrow$ HR	HR
10	MR	0.5 $\Rightarrow$ MR	LR
11	LR	0.472 $\Rightarrow$ LR	LR
12	LR	0.445 $\Rightarrow$ LR	MR
13	MR	0.5 $\Rightarrow$ MR	MR
14	HR	0.527 $\Rightarrow$ HR	HR
15	LR	0.446 $\Rightarrow$ LR	MR
16	MR	0.5 $\Rightarrow$ MR	HR
17	LR	0.302 $\Rightarrow$ LR	HR
18	MR	0.5 $\Rightarrow$ MR	MR
19	HR	0.554 $\Rightarrow$ HR	LR
20	MR	0.5 $\Rightarrow$ MR	LR
21	MR	0.5 $\Rightarrow$ MR	MR
22	LR	0.446 $\Rightarrow$ LR	MR
23	LR	0.463 $\Rightarrow$ LR	LR
24	MR	0.5 $\Rightarrow$ MR	MR
25	MR	0.5 $\Rightarrow$ MR	LR

As Table 6 shows, 100% of outputs were correctly labeled by the proposed fuzzy system, while 52% were correctly labeled by agglomerative clustering, regarding to an expert opinion.

3.2. Pipeline Prediction Model

Once the dataset output labels were classified, the pipeline prediction model was implemented for making predictions on new data. First, the gene expression values were scaled by the Standard-Scaler technique. Then, PCA technique was used to reduce the dimensionality of gene dataset by compressing it into a new subspace, only the subset of the eigenvectors (principal components) that contained more information (maximum variance) were selected. Figure 8 shows the results of the explained variance ratio of the eigenvalues.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

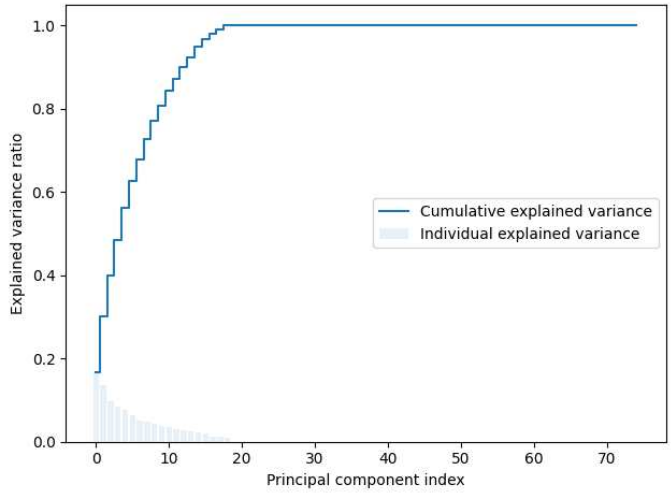

Figure 8. Explained variance ratio.

The resulting graph indicates that the first principal component by itself accounts approximately 20% of the variance. Furthermore, the first two combined principal components represent almost 40% of the total variance.

Figure 9 shows the graph used to determine the optimal value of the number of principal components (n_components=13) for PCA technique to achieve the highest testing accuracy of the MLP prediction algorithm.

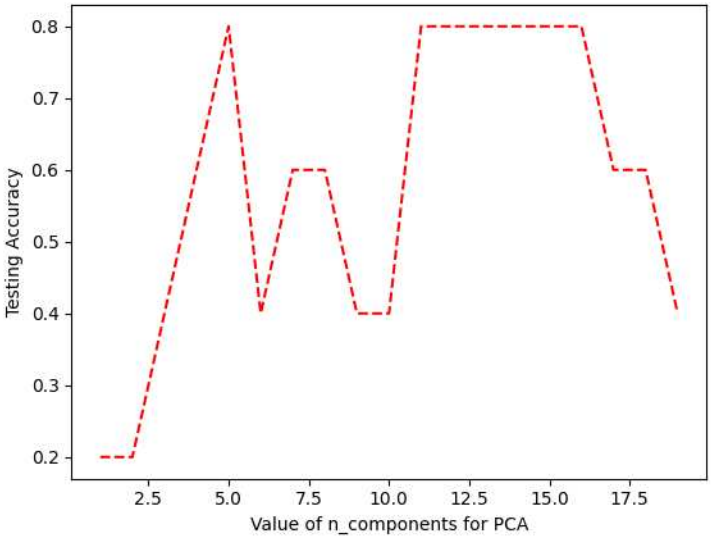

Figure 9. Optimal value of n_components.

3.3. Performance Evaluation

In this paper, the $k = 8$ iterations CV technique was implemented for evaluating the prediction model performance. Table 9 presents the CV accuracy results for each fold. The maximun CV accuracy was achieved at 7th, and 8th folds, and the average performance estimate was 0.521 +/- 0.327.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1478.v1

Table 9. *K* iterations cross-validation results.

Fold	CV Accuracy
1	0.333
2	0.667
3	0.333
4	0.333
5	0.500
6	0.000
7	1.000
8	1.000

The input data (1,875 samples) were divided into 80% *X_train* (1,500 samples) and 20% *X_test* (375 samples) to avoid over fitting. In addition, the output labels (25 samples) were divided into 80% *y_train* (20 samples) and 20% *y_test* (5 samples) for validation. The CM was computed with test and predicted data, the results are shown in Figure 10.

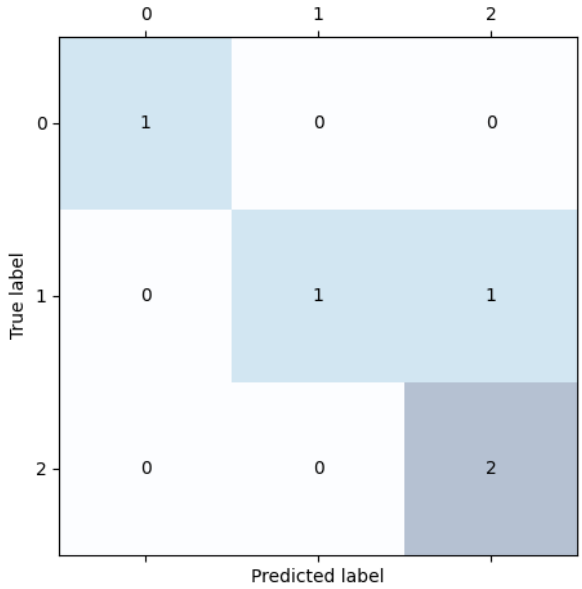

Figure 10. Confusion matrix results. (0) High-responder to IFN-β, (1) Low-responder to IFN-β, and (2) Medium-responder to IFN-β.

The CM results represents one high-responder patient that was correctly predicted as high-responder, one low-responder patient that was correctly predicted as low-responder, one low-responder patient that was wrongly predicted as medium-responder, and two medium-responders patients that were correctly predicted as medium-responders. Based on previous results, the prediction model achieved a 0.8 testing accuracy.

4. Discussion

While binary logic generates only two output types: [0, 1], fuzzy inference engines use an approximate reasoning based on generalized rules of inference. Hence, fuzzy systems are convenient methods for decision support, due to their capability to process inaccurate information. In this paper, an alternative fuzzy system based on expert knowledge was implemented for decision support in classification of the response to IFN-β treatment of RRMS patients. Demographic, and clinical characteristics were used as input variables to fuzzy system. As shown in Table 8, the classification of the proposed fuzzy system had better results than the agglomerative clustering, because the latter

 doi:10.20944/preprints202307.1478.v1

does not consider the intrinsic properties of the data, it just uses the distance between the data points to group them into clusters. A software issue in the fuzzy system design was to set small number of input variables. The greater number of variables, the greater data processing time.

Once the dataset output labels were classified by the fuzzy system, a pipeline prediction model was implemented including data standardization, data compression through PCA technique, and a MLP learning algorithm. The pipeline model was trained with 15 biomarkers associated with the response to IFN-β for predicting whether RRMS patients are potential candidates to be treated with this drug. As shown in Figure 9, by setting thirteen principal components for PCA, a 0.8 testing accuracy was achieved. The use of PCA technique for data compression provides some advantages: 1) The reduced dimension has the property of keeping the most of the useful information while reducing noise and other undesirable data, 2) The time and memory that used in data processing are smaller, 3) It provides a way to understand and visualize the structure of complex datasets. The use of the k iterations CV technique helps to get a good bias-variance rate. The highest CV accuracy was achieved at 7th, and 8th folds, as shown in Table 9. One disadvantage in evaluating the prediction model performance was that the test samples size was too small. Therefore, the number of iterations for CV technique was limited to eight.

5. Conclusions

In general, IFN-β treatment effectively reduce the rate of relapse and delay the progression of neurological disability in MS patients. However, a percentage of patients do not respond, or partially respond to this drug. In this paper, the proposed fuzzy system based on the opinion of an expert demonstrated high efficiency in decision support, and it can be a useful tool in labeling classes such as classification of the response to IFN-β therapy.

Although genome research is complex, there are machine and deep learning methods for instance the proposed pipeline model that can effectively deal the gene data for getting reliable predictions to guide specialists in the selection of MS patients who may have the greatest benefit from IFN-β treatment. Biomarkers in particular IL-2, IL-12, IFN-γ, TNF-α, IL-4, IL-10, TGF-β, CD46, CD58, FHIT, IRF5, GAPVD1, GPC5, GRM3, and GRIK2 can be convenient predictive variables for improving the comprehension of the influence of IFN-β therapy in MS patients.

Author Contributions: Conceptualization, E.R.P.d.L.-S., and A.M.H.-N.; methodology, E.R.P.d.L.-S., and A.M.H.-N.; software, E.R.P.d.L.-S.; validation, H.S.-M.; formal analysis, J.D.M.-S., O.A.D.-R., A.V.-C., and H.J.-H.; investigation, E.R.P.d.L.-S.; resources, J.D.M.-S, O.A.D.-R., A.M.H.-N., A.V.-C., and H.J.-H.; writing—original draft preparation, E.R.P.d.L.-S.; writing—review and editing, E.R.P.d.L.-S., and J.D.M.-S.; supervision, J.D.M.-S., O.A.D.-R., A.V.-C., and H.J.-H.; project administration, E.R.P.d.L.-S., J.D.M.-S, and O.A.D.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The implemented pseudo codes, and the collected dataset are available at https://github.com/ponceraf2020/Pipeline-model.git (accessed on 30 May 2023).

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

References

1. Milo, R.; Miller, A. Revised diagnostic criteria of multiple sclerosis. *Autoimmunity reviews* **2014**, *13*, 518–524.
2. Martynova, E.; Khaibullin, T.; Salafutdinov, I.; Markelova, M.; Laikov, A.; Lopukhov, L.; Liu, R.; Sahay, K.; Goyal, M.; Baranwal, M.; et al. Seasonal Changes in Serum Metabolites in Multiple Sclerosis Relapse. *International Journal of Molecular Sciences* **2023**, *24*, 3542.
3. Tarlinton, R.E.; Martynova, E.; Rizvanov, A.A.; Khaiboullina, S.; Verma, S. Role of viruses in the pathogenesis of multiple sclerosis. *Viruses* **2020**, *12*, 643.
4. Zarghami, A.; Li, Y.; Claflin, S.B.; van der Mei, I.; Taylor, B.V. Role of environmental factors in multiple sclerosis. *Expert Review of Neurotherapeutics* **2021**, *21*, 1389–1408.

5. Dominguez-Mozo, M.I.; Perez-Perez, S.; Villarrubia, N.; Costa-Frossard, L.; Fernandez-Velasco, J.I.; Ortega-Madueño, I.; Garcia-Martinez, M.A.; Garcia-Calvo, E.; Estevez, H.; Luque Garcia, J.L.; et al. Herpesvirus antibodies, vitamin d and short-chain fatty acids: their correlation with cell subsets in multiple sclerosis patients and healthy controls. *Cells* **2021**, *10*, 119.
6. Rodríguez Murúa, S.; Farez, M.F.; Quintana, F.J. The immune response in multiple sclerosis. *Annual Review of Pathology: Mechanisms of Disease* **2022**, *17*, 121–139.
7. Chen, X.; Hou, H.; Qiao, H.; Fan, H.; Zhao, T.; Dong, M. Identification of blood-derived candidate gene markers and a new 7-gene diagnostic model for multiple sclerosis. *Biological Research* **2021**, *54*, 1–12.
8. Pinheiro, M.A.L.; Kooij, G.; Mizee, M.R.; Kamermans, A.; Enzmann, G.; Lyck, R.; Schwaninger, M.; Engelhardt, B.; de Vries, H.E. Immune cell trafficking across the barriers of the central nervous system in multiple sclerosis and stroke. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2016**, *1862*, 461–471.
9. Liu, Y.; Zheng, M.; Ma, Z.; Zhou, Y.; Huo, J.; Zhang, W.; Liu, Y.; Guo, Y.; Zhou, X.; Li, H.; et al. Design, synthesis, and evaluation of PD-L1 degraders to enhance T cell killing activity against melanoma. *Chinese Chemical Letters* **2022**, p. 107762.
10. Nishihara, H.; Soldati, S.; Mossu, A.; Rosito, M.; Rudolph, H.; Muller, W.A.; Latorre, D.; Sallusto, F.; Sospedra, M.; Martin, R.; et al. Human CD4+ T cell subsets differ in their abilities to cross endothelial and epithelial brain barriers in vitro. *Fluids and Barriers of the CNS* **2020**, *17*, 1–18.
11. Mundt, S.; Mrdjen, D.; Utz, S.G.; Greter, M.; Schreiner, B.; Becher, B. Conventional DCs sample and present myelin antigens in the healthy CNS and allow parenchymal T cell entry to initiate neuroinflammation. *Science immunology* **2019**, *4*, eaau8380.
12. Ma, X.; Ma, R.; Zhang, M.; Qian, B.; Wang, B.; Yang, W. Recent Progress in Multiple Sclerosis Treatment Using Immune Cells as Targets. *Pharmaceutics* **2023**, *15*, 728.
13. Szpakowski, P.; Ksiazek-Winiarek, D.; Glabinski, A. Targeting Antigen-Presenting Cells in Multiple Sclerosis Treatment. *Applied Sciences* **2021**, *11*, 8557.
14. Mirandola, S.R.; Hallal, D.E.; Farias, A.S.; Oliveira, E.C.; Brandão, C.O.; Ruocco, H.H.; Damasceno, B.P.; Santos, L.M. Interferon-beta modifies the peripheral blood cell cytokine secretion in patients with multiple sclerosis. *International immunopharmacology* **2009**, *9*, 824–830.
15. Kay, M.; Hojati, Z.; Dehghanian, F. The molecular study of IFNβ pleiotropic roles in MS treatment. *Iranian journal of neurology* **2013**, *12*, 149.
16. Cohan, S.L.; Hendin, B.A.; Reder, A.T.; Smoot, K.; Avila, R.; Mendoza, J.P.; Weinstock-Guttman, B. Interferons and multiple sclerosis: lessons from 25 years of clinical and real-world experience with intramuscular interferon beta-1a (Avonex). *CNS drugs* **2021**, *35*, 743–767.
17. Río, J.; Nos, C.; Tintoré, M.; Borrás, C.; Galán, I.; Comabella, M.; Montalban, X. Assessment of different treatment failure criteria in a cohort of relapsing–remitting multiple sclerosis patients treated with interferon β: Implications for clinical trials. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **2002**, *52*, 400–406.
18. Bustamante, M.F.; Morcillo-Suárez, C.; Malhotra, S.; Rio, J.; Leyva, L.; Fernández, O.; Zettl, U.K.; Killestein, J.; Brassat, D.; García-Merino, J.A.; et al. Pharmacogenomic study in patients with multiple sclerosis: responders and nonresponders to IFN-β. *Neurology-Neuroimmunology Neuroinflammation* **2015**, *2*.
19. Martínez-Aguilar, L.; Pérez-Ramírez, C.; del Mar Maldonado-Montoro, M.; Carrasco-Campos, M.I.; Membrive-Jiménez, C.; Martínez-Martínez, F.; García-Collado, C.; Calleja-Hernández, M.Á.; Ramírez-Tortosa, M.C.; Jiménez-Morales, A. Effect of genetic polymorphisms on therapeutic response in multiple sclerosis relapsing-remitting patients treated with interferon-beta. *Mutation Research/Reviews in Mutation Research* **2020**, *785*, 108322.
20. KARLIK, B. Soft computing methods in bioinformatics: a comprehensive review. *Mathematical and Computational Applications* **2013**, *18*, 176–197.
21. Ayangbekun, O.; Jimoh Ibrahim, A. Fuzzy logic application to brain diseases diagnosis. *Journal of Emerging Trends in Computing and Information Sciences* **2015**, *6*, 144–148.
22. Hosseini, A.; Asadi, F.; Arani, L.A. Development of a knowledge-based clinical decision support system for multiple sclerosis diagnosis. *Journal of Medicine and Life* **2020**, *13*, 612.
23. Matinfar, F.; Golpaygani, A.T. A fuzzy expert system for early diagnosis of multiple sclerosis. *Journal of Biomedical Physics & Engineering* **2022**, *12*, 181.

24. Ali, N.M.; Shaheen, M.; Mabrouk, M.S.; Aborizka, M. Machine Learning-Based Models for Detection of Biomarkers of Autoimmune Diseases by Fragmentation and Analysis of miRNA Sequences. *Applied Sciences* **2022**, *12*, 5583.
25. Viatkin, D.; Garcia-Zapirain, B.; Méndez Zorrilla, A. Deep learning techniques applied to predict and measure finger movement in patients with multiple sclerosis. *Applied Sciences* **2021**, *11*, 3137.
26. Gurevich, M.; Miron, G.; Falb, R.Z.; Magalashvili, D.; Dolev, M.; Stern, Y.; Achiron, A. Transcriptional response to interferon beta-1a treatment in patients with secondary progressive multiple sclerosis. *BMC neurology* **2015**, *15*, 1–8.
27. Clarelli, F.; Liberatore, G.; Sorosina, M.; Osiceanu, A.; Esposito, F.; Mascia, E.; Santoro, S.; Pavan, G.; Colombo, B.; Moiola, L.; et al. Pharmacogenetic study of long-term response to interferon-β treatment in multiple sclerosis. *The Pharmacogenomics Journal* **2017**, *17*, 84–91.
28. Tao Jin, Chi Wang, S.T. Feature selection based on differentially correlated gene pairs reveals the mechanism of IFN-therapy for multiple sclerosis. *Bioinformatics and Genomics* **2020**, *8*. https://doi.org///dx.doi.org/10.7717/peerj.8812.
29. National Center for Biotechnology Information (NCBI) - Gene Expression Omnibus (GEO) Database **2010**. https://doi.org/https://www.ncbi.nlm.nih.gov/geo/geo2r.
30. Czabanski, R.; Jezewski, M.; Leski, J. Introduction to fuzzy systems. *Theory and applications of ordered fuzzy numbers: A tribute to professor Witold Kosiński* **2017**, pp. 23–43.
31. Rutkowska, D. *Neuro-fuzzy architectures and hybrid learning*; Vol. 85, Springer Science & Business Media, 2001.
32. Van Leekwijck, W.; Kerre, E.E. Defuzzification: criteria and classification. *Fuzzy sets and systems* **1999**, *108*, 159–178.
33. Prokopowicz, P.; Czerniak, J.; Mikołajewski, D.; Apiecionek, Ł.; Ślezak, D. *Theory and applications of ordered fuzzy numbers: a tribute to Professor Witold Kosiński*; Springer Nature, 2017.
34. Sanguansat, P. *Principal Component Analysis: Engineering Applications*; BoD–Books on Demand, 2012.
35. Mirjalili, V.; Raschka, S. *Python machine learning*; Marcombo, 2020.
36. Casalino, G.; Castellano, G.; Consiglio, A.; Nuzziello, N.; Vessio, G. MicroRNA expression classification for pediatric multiple sclerosis identification. *Journal of Ambient Intelligence and Humanized Computing* **2021**, pp. 1–10.
37. Salamai, A.A.; El-kenawy, E.S.M.; Abdelhameed, I. Dynamic voting classifier for risk identification in supply chain 4.0. *CMC-COMPUTERS MATERIALS & CONTINUA* **2021**, *69*, 3749–3766.