Fuzzy Design of Nearest Prototype Classifier

Yanela Rodríguez Alvarez, Rafael Bello Pérez, Yailé Caballero Mota, Yaima Filiberto Cabrera, Yumilka Fernández Hernández, Mabel Frias Dominguez

1 Departamento de Computación, Universidad de Camagüey, Circunvalación Norte Km 5 ½, Camagüey, Cuba. {yanela.rodriguez, yaile.caballero, yumilka.fernandez, mabel.frias}@reduc.edu.cu

2 Departamento de Ciencias de la Computación, Universidad Central “Marta Abreu” de las Villas. Carretera a Camajuaní Km. 5 y ½, Santa Clara, Villa Clara, Cuba. rbellop@uclv.edu.cu

Abstract. In pattern classification problems, many works have been carried out with the aim of designing good classifiers from different perspectives. These works achieve very good results in many domains. However, in general they are very dependent on some crucial parameters involved in the design. An alternative is to use fuzzy relations to eliminate thresholds and make the development of classifiers more flexible. In this paper, a new method for solving data classification problems based on prototypes is proposed. Using fuzzy similarity relations for the granulation of the universe, similarity classes are generated and a prototype is built for each similarity class. In the new approach we replace the relation of similarity between two objects by a binary fuzzy relation, which quantifies the strength of the relationship in a range of [0;1]. Experimental results show that the performance of our method is superior to other methods.

Keywords: prototype generation; similarity relations; fuzzy-rough sets theory; classification.

1. Introduction

Theories of rough and fuzzy sets are distinct and complementary generalizations of set theory. A fuzzy set allows a membership value other than 0 and 1. A rough set uses three membership functions, a reference set and its lower and upper approximations in an approximation space. There are extensive studies on the relationships between rough and fuzzy sets *[1-4]. Many proposals have been made for the combination of rough and fuzzy sets.

Rough and fuzzy sets are complementary with each other in the sense that fuzzy sets model the ambiguous memberships between elements and classes while rough sets provide a way of approximating indefinable concept with a pair of definable sets within the universe. This observation motivated a lot of researchers to combine fuzzy sets and rough sets together and various kinds of fuzzy rough set models had been proposed in publications. One is the constructive approach that starts with the fuzzy relations on the
universe and the lower and upper approximation operators are constructed via these fuzzy relations [5]. The constructive approach of fuzzy rough sets is firstly proposed by Dubois and Prade [6, 7], and Radzikowska and Kerre [8] provide a more general approach of constructing fuzzy rough sets.

A recently proposed competitive scheme termed NPBASIR-CLASS method, concerning to work [9], has been adapted to solve different kinds of problems such as: classification problems on domains with hubness [10] and classification of imbalanced datasets [11]. This method NPBASIR-CLASS to build prototypes use the concepts of Granular Computation [12] and are based on the NPBASIR algorithm [13]. Granulation of a universe is performed using a relation of similarity which generates similarity classes of objects in the universe and for each similarity class one prototype is built. To build the similarity relation the method proposed in [14] is used. A similarity quality measure attempts to establish a correspondence between the granulation induced by the set of input attributes A and that induced by the decision attribute d.

NPBASIR-CLASS make use of crisp similarity relations to build the granulations corresponding to the A and d spaces, respectively. In other words, two objects of the universe will belong to the same similarity class if their degree of indiscernibility (or similarity) goes beyond a user-defined threshold in [0;1]. Hence, the similarity quality measure depends on two thresholds (one for the input attributes $\varepsilon_1$ and one for the decision attribute $\varepsilon_2$) which have a crucial importance in any subsequent data analysis step that leans upon the similarity classes. Besides, the threshold values will be application dependent, hence calling for a fine-tuning process to maximize the performance of the knowledge discovery process at hand.

Given that NPBASIR-CLASS is fairly sensitive to the values of the $\varepsilon_1$ and $\varepsilon_2$ similarity thresholds, in this paper we tackle this limitation by employing fuzzy sets to categorize their domains through fuzzy binary relations. We show how this facilitates the definition of the similarity relations (as there are fewer parameters to consider) and enhances the system’s interpretability without degrading, from a statistical perspective, the efficacy of the subsequent data mining tasks.

The present study examines the combination of rough and fuzzy sets from the perspective of design of Nearest Prototype Classifiers. The key contribution of the proposed work is the development of a novel classification approach that uses Nearest Prototype and then combines the advantages of both rough and fuzzy sets in the classification process. Also, in this case the weights of the similarities between the two objects according to each input attributes have been optimized using Particle Swarm Optimization., which conducts a search over the space of similarity weights for the input attributes and uses the similarity quality measure as the fitness of a candidate weight vector. The performance comparison with state-of-the-art methods proposed in the literature is presented in the experimental part.
2. Related Works

The based method, NPBASIR-CLASS [9], uses a similarity relation $R$ and a set of instances $X = \{X_1, X_2, \ldots, X_n\}$, each of which is described by a vector of $m$ descriptive features and belongs to one of $k$ classes $C = \{c_1, c_2, \ldots, c_k\}$. The similarity relation $R$ is constructed according to the method proposed in [14]; this is based on finding the relation that maximizes the quality of the similarity measure. In this case, the relation $R$ is sought that generates a granulation considering the $m$ descriptive features, as similar as possible to the granulation according to the classes.

Both two granulations are built using the crisp binary relations $R_1$ and $R_2$ defined in Equation (1) and Equation (2):

\[
xR_1y \iff F_1(x, y) \geq \varepsilon_1
\]

\[
xR_2y \iff F_2(x, y) \geq \varepsilon_2
\]

where $x$ and $y$ are two objects in $U$, $F_1$ is a similarity function over the input attributes, $F_2$ is a similarity function over the decision attribute and $\varepsilon_1$, $\varepsilon_2$ are their respective similarity thresholds. The similarity function $F_1$ and $F_2$ usually takes the form in Equation (3) and (4):

\[
F_1(x, y) = \sum_{i=1}^{N} w_i \cdot \delta_i(x_i, y_i)
\]

\[
F_2(x, y) = \delta_d(x_d, y_d)
\]

where $N = |A|$ is the number of input attributes, $\delta_i(\cdot, \cdot)$ is the similarity function of two objects $x$ and $y$ with respect to the $i$-th attribute and $w_i$ its associated weight. The goal then is to find the relations $R_1$ and $R_2$ such that $R_1(x, y)$ and $R_2(x, y)$ are as similar as possible. To do so, the following sets are defined:

\[
N_1(x) = y \in U : xR_1y
\]

\[
N_2(x) = y \in U : xR_2y
\]

where $N_1(x)$ and $N_2(x)$ contain the objects that are similar enough to $x$ according to the input attribute set $A$ and the decision attribute $d$, respectively. The similarity degree $0 \leq \phi(x) \leq 1$ between both sets for a given object $x$ is hence quantified as shown in Equation (7):

\[
\phi(x) = \frac{|N_1(x) \cap N_2(x)|}{0.5|N_1(x)| + 0.5|N_2(x)|} = \frac{2|N_1(x) \cap N_2(x)|}{|N_1(x)| + |N_2(x)|}
\]

The similarity quality measure $\Theta(\cdot)$ of a decision system DS of $M = |DS|$ objects is then defined in Equation (8):

\[
\Theta(SD) = \frac{\sum_{x \in SD} \phi(x)}{|U|}
\]
Filiberto et al [15] review different applications of the similarity quality measure in machine learning, e.g. improving classifiers like Multi-Layer Perceptron [14, 16] and K-nearest neighbors (KNN) [14, 17], in prototype construction [18] and the induction of classification rules [13].

To determine the weight vector \( \mathbf{w} = (w_1, \ldots, w_N) \) in the similarity function \( F_1 \) of Equation (3), the authors in [16] employ Particle Swan Optimization PSO to navigate the numerical space of all candidate weight vectors. Each particle is represented by a position vector \( p \) and a velocity vector \( v \). The position encodes a candidate weight vector \( \mathbf{w} \) whose fitness value is precisely the similarity quality measure \( \theta (DS) \) in Equation (8) induced by \( \mathbf{w} \) over the decision system DS. The attributewise similarity is defined as shown in Equation (9).

\[
\partial_i(x_i, y_i) = \begin{cases} 
1 - \frac{|x_i - y_i|}{\text{Max}(n_j) - \text{Min}(n_j)} & \text{if } A_i \text{ is numerical} \\
1 & \text{if } A_i \text{ is nominal } \land x_i = y_i \\
0 & \text{if } A_i \text{ is nominal } \land x_i \neq y_i 
\end{cases} 
\] (9)

where \( A_i \) is the \( i \)-th attribute of the decision system DS and \( x_i, y_i \) are the values of \( A_i \) in objects \( x \) and \( y \), respectively. The particles update their velocities and positions as per the standard rules in Equations (10) and (11).

\[
v_i(t + 1) = w \cdot v_i(t) + c_1(t) \cdot r_1(t) \cdot (p_{best}(t) - p_i(t)) + c_2(t) \cdot r_2(t) \cdot (g_{best}(t) - p_i(t)) 
\] (10)

\[
p_i(t + 1) = p_i(t) + v_i(t + 1) 
\] (11)

where \( t \) is the current iteration number, \( p_i \) and \( v_i \) are the particle’s position and velocity in the \( i \)-th dimension, \( w \) is the inertia weight, \( c_1 \) and \( c_2 \) are two acceleration constants, \( r_1 \) and \( r_2 \) are two uniformly distributed random numbers in \([0; 1]\), \( p_{best}(t) \) and \( g_{best}(t) \) are the respective best solutions found by the particle itself and the swarm as a whole. This iterative update process takes place until the stop criterion is met, e.g., reaching a maximum number of iterations. At the end of the PSO search, the \( g_{best} \) particle encodes in its position the best weight vector \( \mathbf{w} \) found, which is then plugged into the similarity function \( F_1 \) in Equation (3), thus allowing the calculation of the similarity relation \( R_1 \) depicted in Equation (1).

3. **Our proposed**

The impact of the similarity thresholds \( \varepsilon_1 \) and \( \varepsilon_2 \) upon the NPBSIR-CLASS approach was assessed and discussed in previously sections as well. In general, this scheme was found to be fairly sensitive to small variations of these thresholds, thus causing the PSO metaheuristic to converge to different local optima in the weight space, hence affecting the stability of this technique. In an attempt to overcome this limitation and increase the overall system interpretability without degrading its performance, this paper employs fuzzy sets to characterize the domain of both similarity thresholds.
Taking into account [19], we propose the use of fuzzy binary relations in lieu of crisp binary relations. Zadeh [18] defined a fuzzy binary relation $R$ between $X \subseteq U$ and $Y \subseteq U$ as a fuzzy subset of $X \times Y$, i.e. the similarity $\mu_R(x,y)$ between any two objects of the universe can take values in the interval $[0; 1]$. The value of $\mu_R(x,y)$ represents the strength of the relationship between $x$ and $y$ [19]. The use of fuzzy relations makes the computational methods more tolerant and flexible to imprecision, especially in the case of mixed data (numerical and nominal variables). Taking these criteria into account, the use of fuzzy relations in the NPBASIR-CLASS algorithm is proposed in this paper.

Using a fuzzy approach to build the similarity quality measure requires the rewriting of the binary relations $R_1$ and $R_2$, defined in Equation (1) and Equation (2), as fuzzy binary relations defined in Equation (12) and Equation (13):

$$x R_1^f y \iff F_1(x,y) \text{ is High}_1$$

$$x R_2^f y \iff F_2(x,y) \text{ is High}_2$$

where $R_1^f$ and $R_2^f$, are the fuzzy binary relations for the input attributes $A$ and the decision attribute $d$, respectively and High$_1$ and High$_2$ are fuzzy sets that quantify the similarity between objects $x$ and $y$ in their respective attribute spaces. Both fuzzy sets are modeled after sigmoidal membership functions whose analytical expressions are given in Equations (14) and (15) and schematic depictions provided in Figures 1(a) and 1(b).

$$\mu_{\text{High}_1}(x) = \begin{cases} 0 & \text{if } x \leq 0.70 \\ \frac{2(x-0.70)^2}{1+2(x-0.70)^2} & \text{otherwise} \end{cases}$$

$$\mu_{\text{High}_2}(x) = \begin{cases} 0 & \text{if } x < 0.75 \\ \frac{2\left(\frac{x-0.75}{0.90-0.75}\right)^2}{1-2\left(\frac{x-0.75}{0.90-0.75}\right)^2} & \text{if } 0.75 \leq x \leq 0.85 \\ 1 & \text{if } 0.85 \leq x \leq 0.90 \\ \frac{2\left(\frac{x-0.75}{0.90-0.75}\right)^2}{1-2\left(\frac{x-0.75}{0.90-0.75}\right)^2} & \text{otherwise} \end{cases}$$

From the High$_1$ and High$_2$ fuzzy sets we can build the fuzzy versions of $N_1(x)$ and $N_2(x)$ by replacing Equations (5) and (6) with Equations (16) and (17).

$$N_1^f(x) = \{ (y, \mu_{\text{High}_1}(F_1(x,y))) \forall y \in U \}$$

$$N_2^f(x) = \{ (y, \mu_{\text{High}_2}(F_2(x,y))) \forall y \in U \}$$

The fuzzy similarity degree between both sets [20] for an object $\bar{x}$ is calculated as the similarity between the fuzzy sets $N_1^f(x)$ and $N_2^f(x)$ according to Equation (18):

$$\phi^f(\bar{x}) = \frac{\sum_{i=1}^{M} (1-|\mu_{\text{High}_1}(\bar{x}_i) - \mu_{\text{High}_2}(\bar{x}_i)|)}{M}$$
Fig. 1. The sigmoidal membership functions for the High_1 and High_2 fuzzy sets

With the fuzzy similarity degree in Equation (18), we can formalize the fuzzy similarity quality measure of a decision system DS of \( M = |DS| \) objects as shown in Equation (19):

\[
\Theta^F (DS) = \frac{\sum_{i=1}^{M} \phi^F (x)}{M} \quad (19)
\]

With the advent of fuzzy logic into the similarity-quality-based framework, numerical thresholds are no longer required as part of the definition of the similarity relation. Therefore, we reduce the number of degrees of freedom (algorithm parameters) under consideration, increase system interpretability (since the High_1 and High_2 fuzzy sets are intuitive for any user).

4. Experimental Results

In this section, we present some experimental facts regarding the proposed NPBASIR-FUZZY classification methodology and compare it with other state-of-the-art classification algorithms in terms of accuracy. The proposed NPBASIR-FUZZY classification methodology has been deployed on 17 datasets from UCI Machine Learning repository [21]. The 17 datasets and their description with number of attributes, instances, continuous attributes, discrete attributes, classes and missing values are shown below in the Table 1. These data sets include a wide range of domains and a variety of data characteristics such as number of classes, instances, and attributes. The prediction accuracy has been measured by applying a 10-fold cross-validation where each dataset is randomly partitioned into 10 approximately equally sized subsets (or folds or tests). The induction algorithm is executed 10 times; in each time it is trained on the data that is outside one of the subsets and the generated classifier is tested on that subset. The estimated accuracy for each cross-validation fold is a random variable that depends on the random partitioning of the data. So, for each dataset, we repeated 10- fold cross-validation 10 times. The estimated accuracy is the average over the ten 10- fold cross-validations.
Table 1. Description of data sets

<table>
<thead>
<tr>
<th>Datasets</th>
<th>#Att.</th>
<th>#Inst.</th>
<th>#Cont.</th>
<th>#Disc.</th>
<th>#Class</th>
<th>Missing values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast-cancer</td>
<td>10</td>
<td>286</td>
<td>0</td>
<td>10</td>
<td>2</td>
<td>9/0.3%</td>
</tr>
<tr>
<td>Breast-w</td>
<td>9</td>
<td>683</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>16/2%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>8</td>
<td>768</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Flags</td>
<td>30</td>
<td>194</td>
<td>10</td>
<td>20</td>
<td>8</td>
<td>None</td>
</tr>
<tr>
<td>Haberman</td>
<td>4</td>
<td>306</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Heart-Statlog</td>
<td>13</td>
<td>270</td>
<td>13</td>
<td>0</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Ionosphere</td>
<td>35</td>
<td>351</td>
<td>34</td>
<td>1</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Iris</td>
<td>4</td>
<td>150</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>None</td>
</tr>
<tr>
<td>Mfeat-fourier</td>
<td>76</td>
<td>2000</td>
<td>76</td>
<td>0</td>
<td>10</td>
<td>None</td>
</tr>
<tr>
<td>Optdigits</td>
<td>64</td>
<td>5620</td>
<td>64</td>
<td>0</td>
<td>10</td>
<td>None</td>
</tr>
<tr>
<td>Pima</td>
<td>8</td>
<td>768</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Postoperative-patient-data</td>
<td>9</td>
<td>90</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>3 / 3%</td>
</tr>
<tr>
<td>Primary-tumor</td>
<td>18</td>
<td>339</td>
<td>0</td>
<td>18</td>
<td>21</td>
<td>224 / 66%</td>
</tr>
<tr>
<td>Solar-flare_1</td>
<td>13</td>
<td>323</td>
<td>0</td>
<td>13</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Solar-flare_2</td>
<td>13</td>
<td>1066</td>
<td>0</td>
<td>13</td>
<td>2</td>
<td>None</td>
</tr>
<tr>
<td>Sponge</td>
<td>46</td>
<td>76</td>
<td>3</td>
<td>43</td>
<td>3</td>
<td>22 / 29%</td>
</tr>
<tr>
<td>Waveform</td>
<td>40</td>
<td>5000</td>
<td>40</td>
<td>0</td>
<td>3</td>
<td>None</td>
</tr>
</tbody>
</table>

This time we chose of the state-of-the-art classification algorithms which have been included within the KEEL [22] software tool, an open source software for multi-stage analysis in data mining. Specifically, 4 algorithms were selected from those that appear within the family of the Fuzzy Instance Based Learning: FRKNNA-C [23], FRNN-C [24], FRNN_FRS-C [25], FRNN_VQRS-C [25]. These algorithms following the fuzzy-rough nearest-neighbor approach based on the fuzzy-rough sets theory, like as our proposed. Table 2 shows, for each method, the classification accuracy (%) achieved by the state-of-the-art classification algorithms and our proposed, tested on 17 UCI datasets.

Table 2. Classification accuracy (%)

<table>
<thead>
<tr>
<th>Datasets</th>
<th>FRKNNA-C</th>
<th>FRNN-C</th>
<th>FRNN_FRS-C</th>
<th>FRNN_VQRS-C</th>
<th>NPBASIR-FUZZY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast-cancer</td>
<td>65.78</td>
<td>71.7</td>
<td>66.13</td>
<td>64.74</td>
<td>74.51</td>
</tr>
<tr>
<td>Breast-w</td>
<td>97.07</td>
<td>88.14</td>
<td>95.47</td>
<td>95.32</td>
<td>97.22</td>
</tr>
<tr>
<td>Diabetes</td>
<td>72.52</td>
<td>66.41</td>
<td>70.06</td>
<td>70.06</td>
<td>74.61</td>
</tr>
<tr>
<td>Flags</td>
<td>51.05</td>
<td>54.13</td>
<td>29.89</td>
<td>26.79</td>
<td>60.21</td>
</tr>
<tr>
<td>Haberman</td>
<td>71.55</td>
<td>73.53</td>
<td>64.04</td>
<td>64.04</td>
<td>73.84</td>
</tr>
<tr>
<td>Heart-Statlog</td>
<td>76.67</td>
<td>81.85</td>
<td>75.93</td>
<td>74.44</td>
<td>82.22</td>
</tr>
<tr>
<td>Ionosphere</td>
<td>64.1</td>
<td>0</td>
<td>76.07</td>
<td>64.1</td>
<td>80.33</td>
</tr>
<tr>
<td>Iris</td>
<td>92.67</td>
<td>88.67</td>
<td>95.33</td>
<td>95.33</td>
<td>96.00</td>
</tr>
<tr>
<td>Mfeat-fourier</td>
<td>78.35</td>
<td>77.5</td>
<td>78.8</td>
<td>78.65</td>
<td>79.00</td>
</tr>
<tr>
<td>Optdigits</td>
<td>9.86</td>
<td>0</td>
<td>9.86</td>
<td>9.54</td>
<td>89.57</td>
</tr>
<tr>
<td>Pima</td>
<td>71.11</td>
<td>66.67</td>
<td>70.6</td>
<td>70.6</td>
<td>73.83</td>
</tr>
</tbody>
</table>
Postoperative-patient-data  62.22  71.11  63.33  60  71.11
Primary-tumor  36.9  43.34  33.34  15.65  40.99
Solar-flare_1  97.84  97.84  97.22  94.73  97.84
Sponge  92.5  95  55  55.18  92.5
Waveform  72.94  83.6  72.2  72.28  81.46

To statistically validate the experimental results, nonparametric tests for multiple comparisons with a control method were used. In Table III we show the average ranks obtained by each method in the Friedman test. The P-value computed by Iman and Davenport Test is equal $0.000007910702$. The resulting p-value = $0.004989577133 < \alpha = 0.05$, this indicates that there are indeed significant performance differences in the group. Besides, Table 4 show the Holm post hoc comparison (Friedman) where p-values are obtained in by applying post hoc methods over the results of Friedman procedure. The results reported in Table 4 reject all null hypotheses, hence confirming the superiority of the control method.

Table 3. Average rankings of Friedman test

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPBASIR-FUZZY</td>
<td>1.5</td>
</tr>
<tr>
<td>FRKNNA-C</td>
<td>2.9118</td>
</tr>
<tr>
<td>FRNN-C</td>
<td>3.0294</td>
</tr>
<tr>
<td>FRNN-FRS-C</td>
<td>3.5294</td>
</tr>
<tr>
<td>FRNN-VQRS-C</td>
<td>4.0294</td>
</tr>
</tbody>
</table>

Table 4. Post Hoc comparison Table with NPBASIR-FUZZY as the control method for $\alpha=0.05$

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>z</th>
<th>p</th>
<th>Holm</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>FRNN-VQRS-C</td>
<td>4.664005</td>
<td>0.000003</td>
<td>0.0125</td>
</tr>
<tr>
<td>3</td>
<td>FRNN-FRS-C</td>
<td>3.74205</td>
<td>0.00183</td>
<td>0.016667</td>
</tr>
<tr>
<td>2</td>
<td>FRNN-C</td>
<td>2.820096</td>
<td>0.004801</td>
<td>0.025</td>
</tr>
<tr>
<td>1</td>
<td>FRKNNA-C</td>
<td>2.603165</td>
<td>0.009237</td>
<td>0.05</td>
</tr>
</tbody>
</table>

From the above analysis we can conclude that by replacing the similarity thresholds $\varepsilon_1$, $\varepsilon_2$ in the definition of the crisp binary relations in Equation (1) and Equation (2) with the High_1 and High_2 fuzzy sets in the fuzzy binary relations shown in Equation (11) and Equation (12), the accuracy of the classification is not statistically deteriorated. Additionally, the number of parameters in the automatic similarity learning procedure is reduced and the overall system interpretability is enhanced.

5. Conclusions

This paper introduces a fuzzy modeling of one of the Nearest Prototype Classifiers, in the base method NPBASIR-CLASS we change the numerical thresholds by a fuzzy
approach in the definition of the binary relations that underlie any similarity-based in-
ference process. This modification improving the interpretability of the system and re-
ducing the number of parameters that will need to be fine-tuned. This will, in turn, shorten the time necessary to test, validate and deploy the system. The experimental study carried out on several datasets, four learning algorithms based on the fuzzy-rough sets theory and non-parametric statistical methods have been used to compare and an-
alyze the accuracy of the algorithms. The previously study permit concludes that the proposed method NPBASIR-FUZZY outperforms the rest of the algorithms with regard to the accuracy results on test data. NPBASIR-FUZZY is a compact and interpretable Nearest Prototype Classifier.

Future research may focus on simulating this approach on prototypes selection based on similarity relations for classification problems, taking NPBASIR SEL-CLASS [10] as base method, and no limit the simulation for fuzzy-rough NN approach only.

6. Acknowledgment

This research has been partially sponsored by VLIR-UOS Network University Co-
operation Programme - Cuba.

7. References

Yao, Y., Combination of rough and fuzzy sets based on α-level sets, in Rough sets and Data Mining. 1997, Springer. p. 301-321.


Filiberto, Y., et al. A method to build similarity relations into extended Rough Set Theory. in Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on. 2010. IEEE.


Filiberto, Y., et al. Improving the MLP learning by using a method to calculate the initial weights of the network based on the quality of similarity measure. in Mexican International Conference on Artificial Intelligence. 2011. Springer.


