

# A Short Study on the Use of Genetic 2-Tuples Tuning for Fuzzy Rule Based Classification Systems in Imbalanced Data-Sets

Alberto Fernández  
Dept. of Computer Science and A.I.  
University of Granada, Spain  
alberto@decsai.ugr.es

María José del Jesus  
Dept. of Computer Science  
University of Jaén, Spain  
mjjesus@ujaen.es

Francisco Herrera  
Dept. of Computer Science and A.I.  
University of Granada, Spain  
herrera@decsai.ugr.es

## Abstract

*In this work our aim is to increase the performance of Fuzzy Rule Based Classifications Systems in the framework of imbalanced data-sets by means of the application of a genetic tuning step. We focus on the imbalanced data-set problem since it appears in many real application areas and, for this reason, it has become a relevant topic in the area of machine learning. This problem occurs when the number of examples that represents one of the concepts of interest (usually the most important) is much lower than that of the remaining ones.*

*We want to adapt the 2-tuples based genetic tuning approach to classification problems and to study the positive synergy between this method and the Chi et al.'s fuzzy learning method, which is a basic approach in order to build the initial Knowledge Base.*

*The experimental results show the improvement achieved by the 2-tuples based genetic tuning over the Fuzzy Rule Based Classification System in all types of imbalanced data, obtaining a better behaviour than the basic approach.*

## 1 Introduction

Fuzzy Rule Based Classification Systems (FRBCSs) are a very useful tool in the ambit of Machine Learning, since they provide a very interpretable model for the end user [10]. There are many real applications in which the FRBCS have been employed, including anomaly intrusion detection, medicine or image processing. In most of these areas the data used is highly skewed, i.e. the number of instances of one class is much lower than the instances of the other

classes. This situation is known as the imbalanced data-set problem, and it has been recently identified as one important problem in data mining [5].

Our previous work on the topic [8] showed the good behaviour obtained by FRBCSs in the ambit of imbalanced data-sets, by means of the application of a preprocessing step in order to balance the training data before the rule generation phase via a re-sampling procedure. We determined the robustness of this methodology specially when increasing the imbalance degree, achieving better performance than the C4.5 decision tree in highly imbalanced data-sets.

Our aim in this work is to increase the global performance of FRBCSs applying a genetic optimization of the Data Base membership functions (MFs) by means of a new linguistic rule representation model that was proposed in [1]. This model is based on the linguistic 2-tuples representation [9], and allows the lateral displacement of a label considering an unique parameter.

We will use for our experimental study a basic learning algorithm proposed by Chi et al. [6], that extends the well-known Wang and Mendel method to classification problems. Furthermore, as we mentioned before, it is necessary to apply a re-sampling procedure to prepare the training data for the learning process. Specifically, following the conclusions obtained in [8], we will employ the “Synthetic Minority Over-sampling Technique” (SMOTE) [4].

The rest of this work is organized as follows. In Section 2, we present the imbalanced data-set problem, describing the preprocessing technique used in our work, the SMOTE algorithm, and discussing the evaluation metric we have employed. In Section 3, we briefly describe the FRBCSs and we present the fuzzy rule learning methodology used in this study. Next, Section 4 introduces the 2-tuples tu-

ning approach and the evolutionary algorithm that tunes the FRBCS. In Section 5, we include our experimental analysis in imbalanced data-sets with different degrees of imbalance, where we compare the FRBCS with 2-tuples based genetic tuning with the basic FRBCS approach. Finally, in Section 6, some concluding remarks are pointed out.

## 2 Imbalanced Data-Sets in Classification

The problem of imbalanced data-sets in classification occurs when the data used is highly skewed, i.e. the number of instances of one class is much lower than the instances of the other classes. This situation has been recently identified as one important problem in data mining [5], because it is implicit in most real world applications, including telecommunications, finance, biology, and medicine.

We focus on the two class imbalanced data-sets, where there is only one positive and one negative class. We consider the positive class as the one with the lowest number of examples and the negative class the one with the highest number of examples. Furthermore, in this work we use the imbalance ratio (IR) [11], defined as the ratio of the number of instances of the majority class and the minority class, to organize the different data-sets according to their IR.

In order to deal with the imbalanced data-set problem we may apply internal approaches that create new algorithms or modify existing ones to take this problem into consideration [2] and external approaches that preprocess the data in order to diminish the effect caused by their class imbalance [3].

In our previous work on this topic [8] we analyzed the cooperation of some preprocessing methods with FRBCSs, showing a good behaviour for the oversampling methods, specially in the case of the SMOTE methodology [4]; according to this, we will employ in this work this preprocessing method.

In short, its main idea is to form new minority class examples by interpolating between several minority class examples that lie together. Thus, the overfitting problem is avoided and causes the decision boundaries for the minority class to spread further into the majority class space.

Regarding the empirical measure, instead of using accuracy, a more correct metric is considered. This is due to the fact that accuracy can lead to erroneous conclusions, since it doesn't take into account the proportion of examples for each class. Because of this, in this work we use the geometric mean of the true rates [2], which can be defined as

$$GM = \sqrt{\frac{TP}{TP + FN} \cdot \frac{TN}{FP + TN}} \quad (1)$$

where  $TP, TN, FP$  and  $FN$  stand for True Positives, True Negatives, False Positives and False Negatives respectively. This metric attempts to maximize the accuracy of each one

of the two classes with a good balance. It is a performance metric that links both objectives.

## 3 Fuzzy Rule Based Classification Systems and Fuzzy Learning Method

FRBCSs are a very useful tool in the ambit of Machine Learning, since they allow the inclusion of all the available information in system modeling, both the one that comes for expert knowledge and the one from empiric measures and mathematical models, deriving on a very interpretable model and therefore allowing the knowledge representation to be understandable for the system users.

Any classification problem consists of  $m$  training patterns  $x_p = (x_{p1}, \dots, x_{pn}, y_p)$ ,  $p = 1, 2, \dots, m$  from  $M$  classes where  $x_{pi}$  is the  $i$ th attribute value ( $i = 1, 2, \dots, n$ ) of the  $p$ -th training pattern.

In this work we use fuzzy rules of the following form for our FRBCSs:

$$\text{Rule } R_j : \text{ If } x_1 \text{ is } A_{j1} \text{ and } \dots \text{ and } x_n \text{ is } A_{jn} \quad (2) \\ \text{then Class} = C_j \text{ with } RW_j$$

where  $R_j$  is the label of the  $j$ th rule,  $x = (x_1, \dots, x_n)$  is an  $n$ -dimensional pattern vector,  $A_{ji}$  is an antecedent fuzzy set,  $C_j$  is a class label, and  $RW_j$  is the rule weight. We use triangular MFs as antecedent fuzzy sets.

Fuzzy learning methods are the basis to build a FRBCS. The algorithm used in this work is the method proposed in [6], that we have called the Chi et al.'s rule generation.

To generate the fuzzy RB this FRBCSs design method determines the relationship between the variables of the problem and establishes an association between the space of the features and the space of the classes by means of the following steps:

1. *Establishment of the linguistic partitions.* Once the domain of variation of each feature  $A_i$  is determined, the fuzzy partitions are computed.
2. *Generation of a fuzzy rule for each example*  $x_p = (x_{p1}, \dots, x_{pn}, C_p)$ . To do this is necessary:
  - 2.1 To compute the matching degree  $\mu(x_p)$  of the example to the different fuzzy regions using a conjunction operator (usually modeled with a minimum or product T-norm).
  - 2.2 To assign the example  $x_p$  to the fuzzy region with the greatest membership degree.
  - 2.3 To generate a rule for the example, whose antecedent is determined by the selected fuzzy region and whose consequent is the label of class of the example.
  - 2.4 To compute the rule weight.

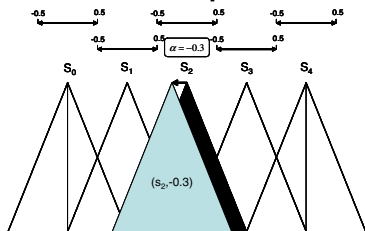
We must remark that rules with the same antecedent can be generated during the learning process. If they have the same class in the consequent we just remove one of the duplicated rules, but if they have a different class only the rule with the highest weight is kept in the RB.

#### 4 Genetic Tuning of the Fuzzy Rule Based Classification Systems

In this work we will employ the 2-tuples based genetic tuning for classification problems, adapting the previous work on the topic [1] in order to obtain good models of FRBCSs to increase the performance of the initial Knowledge Base (KB). In this approach a new rule representation model based on the linguistic 2-tuples representation [9] is used. This representation allows the lateral displacement of the labels considering only one parameter (slight displacements to the left/right of the original MFs).

The symbolic translation of a linguistic term is a number within the interval  $[-0.5, 0.5]$  that expresses the domain of a label when it is moving between its two lateral labels. An example is illustrated in Figure 1 where we show the symbolic translation of a label represented by the pair  $(S_2, -0.3)$  together with the lateral displacement of the corresponding MF.

Figure 1. Lateral Displacement of a MF.



There are two different possibilities as how to perform the lateral tuning, the most interpretable one and the most accurate one:

- *Global Tuning of the Semantics (GTS)*. In this case, the tuning is applied to the level of linguistic partition. In this way, the pair  $(X_i, \text{label})$  takes the same tuning value in all the rules where it is considered.
- *Local Tuning of the Rules (LTR)*. In this case, the tuning is applied at the rule level. The pair  $(X_i, \text{label})$  is tuned in a different way for each rule, based on the quality measures associated to the tuning method.

In order to apply the 2-tuples based genetic tuning, we will consider the use of a specific genetic algorithm to design the proposed learning method, the CHC algorithm [7]

which presents a good trade-off between diversity and convergence, being a good choice in problems with complex search spaces.

This genetic model makes use of a mechanism of “Selection of Populations”.  $M$  parents and their corresponding offspring are put together to select the best  $M$  individuals to take part of the next population (with  $M$  being the population size). Furthermore, no mutation is applied during the recombination phase, Instead, when the population converges or the search stops making progress, the population is re-initialized.

The components needed to design this process are explained below. They are: coding scheme, initial gene pool, chromosome evaluation, crossover operator (together with an incest prevention) and restarting approach.

1. *Coding Scheme*: A real coding is considered. In the case of the GTS the chromosome is represented by the joint of the parameters of the fuzzy partitions, while for the LTR is the joint of the rule parameters.
2. *Chromosome Evaluation*: The fitness function must be in accordance with the framework of imbalanced datasets. Thus, we will use, as presented in Section 2, the geometric mean of the true rates.
3. *Initial Gene Pool*: To make use of the available information, the initial FRBCS is included in the population as an initial solution. To do so, the initial pool is obtained with the first individual having all genes with value ‘0.0’, and the remaining individuals generated at random in  $[-0.5, 0.5]$ .
4. *Crossover Operator*: We consider the Parent Centric BLX (PCBLX) operator, which is based on the BLX- $\alpha$ . We consider the incest prevention mechanism in order to apply the PCBLX operator. Two parents are crossed if their hamming distance divided by 2 is above a predetermined threshold,  $L$ . Since we consider a real coding scheme, we have to transform each gene considering a Gray Code (binary code) with a fixed number of bits per gene (*BITSGENE*). In this way, the threshold value is initialized as:

$$L = (\#Genes \cdot BITSGENE)/4.0$$

where  $\#Genes$  stands for the total length of the chromosome.  $L$  is decremented by *BITSGENE* when there are no new individuals in the next generation.

5. *Restarting approach*: When the threshold value  $L$  is lower than zero, all the chromosomes are regenerated at random within the interval  $[-0.5, 0.5]$ . Furthermore, the best global solution found is included in the population to increase the convergence of the algorithm.

**Table 1. Summary Description for Imbalanced Data-Sets**

Data-set	#Ex.	#Atts.	Class (min., maj.)	% Class	IR
<i>Data-sets with Low Imbalance (1.5 to 3 IR)</i>					
Glass2	214	9	(build-window-non_float-proc, remainder)	(35.51, 64.49)	1.82
EcoliCP-IM	220	7	(im, cp)	(35.00, 65.00)	1.86
Wisconsin	683	9	(malignant, benign)	(35.00, 65.00)	1.86
Pima	768	8	(tested-positive, tested-negative)	(34.84, 66.16)	1.90
Iris1	150	4	(Iris-Setosa, remainder)	(33.33, 66.67)	2.00
Glass1	214	9	(build-window-float-proc, remainder)	(32.71, 67.29)	2.06
Yeast2	1484	8	(NUC, remainder)	(28.91, 71.09)	2.46
Vehicle2	846	18	(Saab, remainder)	(28.37, 71.63)	2.52
Vehicle3	846	18	(bus, remainder)	(28.37, 71.63)	2.52
Vehicle4	846	18	(Opel, remainder)	(28.37, 71.63)	2.52
Haberman	306	3	(Die, Survive)	(27.42, 73.58)	2.68
<i>Data-sets with Medium Imbalance (3 to 9 IR)</i>					
GlassNW	214	9	(non-window glass, remainder)	(23.83, 76.17)	3.19
Vehicle1	846	18	(van, remainder)	(23.64, 76.36)	3.23
Ecoli2	336	7	(im, remainder)	(22.92, 77.08)	3.36
New-thyroid3	215	5	(hypo, remainder)	(16.89, 83.11)	4.92
New-thyroid2	215	5	(hyper, remainder)	(16.28, 83.72)	5.14
Ecoli3	336	7	(pp, remainder)	(15.48, 84.52)	5.46
Segment1	2308	19	(brickface, remainder)	(14.26, 85.74)	6.01
Glass7	214	9	(headlamps, remainder)	(13.55, 86.45)	6.38
Yeast4	1484	8	(ME3, remainder)	(10.98, 89.02)	8.11
Ecoli4	336	7	(iMU, remainder)	(10.88, 89.12)	8.19
Page-blocks	5472	10	(remainder, text)	(10.23, 89.77)	8.77
<i>Data-sets with High Imbalance (higher than 9 IR)</i>					
Vowel0	988	13	(hid, remainder)	(9.01, 90.99)	10.10
Glass3	214	9	(Ve-win-float-proc, remainder)	(8.78, 91.22)	10.39
Ecoli5	336	7	(om, remainder)	(6.74, 93.26)	13.84
Glass5	214	9	(containers, remainder)	(6.07, 93.93)	15.47
Abalone9-18	731	8	(18, 9)	(5.65, 94.25)	16.68
Glass6	214	9	(tableware, remainder)	(4.20, 95.80)	22.81
YeastCYT-POX	482	8	(POX,CYT)	(4.15, 95.85)	23.10
Yeast5	1484	8	(ME2, remainder)	(3.43, 96.57)	28.41
Yeast6	1484	8	(ME1, remainder)	(2.96, 97.04)	32.78
Yeast7	1484	8	(EXC, remainder)	(2.49, 97.51)	39.16
Abalone19	4174	8	(19, remainder)	(0.77, 99.23)	128.87

## 5 Experimental Study

In this study, our intention is to show the improvement achieved in FRBCSs by applying a 2-tuples genetic tuning approach. Thus, we will compare the performance of the FRBCS, both for the basic scheme and with 2-tuples genetic tuning, performing a global comparison employing a large amount of imbalanced data-sets.

Specifically, we have considered thirty three data-sets from UCI with different IR, as shown in Table 1, where we denote the number of examples (#Ex.), number of attributes (#Atts.), class name of each class (minority and majority), class attribute distribution and IR. This table is in ascendance order according to the IR. Data-sets with more than two classes have been modified by taking one against the others or by contrasting one class with another.

In order to reduce the effect of imbalance, we will employ the SMOTE preprocessing method [4] for all our experiments, considering only the 1-nearest neighbour to generate the synthetic samples, and balancing both classes to the 50% distribution.

In the remaining of this section, we will first present the experimental framework and all the parameters employed in this study and then we will show the results and all the

statistical study for the FRBCS approach with the 2-tuples based genetic tuning.

### 5.1 Experimental Set-Up

To develop the different experiments we consider a 5-*folder cross-validation model*, i.e., 5 random partitions of data with a 20%, and the combination of 4 of them (80%) as training and the remaining one as test. For each data-set we consider the average results of the five partitions. Furthermore, Wilcoxon’s Signed-Ranks Test [12] is used for statistical comparison of our empirical results. In all cases the level of confidence ( $\alpha$ ) will be set at 0.05.

We will employ the following configuration for the FRBCS: product T-norm as conjunction operator, together with the Penalized Certainty Factor approach for the rule weight and FRM of the winning rule. We have selected this FRBCS model as it achieved a good performance in our previous study on imbalanced data-sets [8]. Because it is not clear what level of granularity must be employed for the Chi FRBCS, we will use both 3 and 5 labels per variable.

Finally, we indicate the values that have been considered for the parameters of the genetic tuning:

- Population Size: 50 individuals.

- Number of evaluations: 5,000 *· number of variables*.
- Bits per gene (Gray codification): 30 bits.

## 5.2 Analysis of the 2-Tuples Based Genetic Tuning on Fuzzy Rule Based Classification Systems

As we said previously, our aim is to emphasize the necessity in the use of the 2-tuples based genetic tuning to improve the FRBCS performance by means of the comparison of the results obtained by the Chi et al.'s method.

Experimental results are shown in Table 2, where we show by columns the two different approaches considered, that is, the Chi et al.'s algorithm with 3 and 5 labels, noted as Chi-3 and Chi-5 respectively. In addition, there are three different results for each method: the first row contains the results when applying the basic scheme (Base) and the second and third rows contains the results for the global and local 2-tuples based genetic tuning, named as GTS and LTR.

**Table 2. Summary results table for the FRBCSs comparison in imbalanced data-sets**

Approach	Chi-3		Chi-5	
	$GM_{Tr}$	$GM_{Tst}$	$GM_{Tr}$	$GM_{Tst}$
Base	84.95 ± 1.48	80.48 ± 6.24	90.24 ± 0.94	79.57 ± 6.00
GTS	92.70 ± 1.01	83.77 ± 6.84	94.56 ± 1.05	78.76 ± 7.97
LTR	94.33 ± 1.08	83.77 ± 7.32	96.24 ± 0.85	80.26 ± 6.98

For the FRBCSs analysis we must select which granularity is preferred for the Chi method, whether 3 or 5 labels. For this purpose Table 3 presents a Wilcoxon's test where we compare the results for each approach (with the two types of genetic tuning) using the two different number of fuzzy partitions.

**Table 3. Wilcoxon's test to compare Chi using different granularity levels.  $R^+$  corresponds to 3 labels and  $R^-$  to 5 labels**

Comparison	$R^+$	$R^-$	Hypothesis ( $\alpha = 0.05$ )	p-value
Chi-3-GTS vs. Chi-5-GTS	522.5	38.5	Rejected for Chi-3-GTS	0.000
Chi-3-LTR vs. Chi-5-LTR	471.0	90.0	Rejected for Chi-3-LTR	0.001

The main conclusion obtained in this table is that when we choose 5 labels per variable, we get a high over-fitting for the 2-tuples based genetic tuning and, in this case, the choice of a lower level of granularity allows the achievement of better results.

We must emphasize that our work was focused on determining the positive behaviour of the 2-tuples based genetic tuning for FRBCS in the field of imbalanced data-sets. We show in Table 4 the complete results for the Chi et al.'s fuzzy

method with 3 labels per variable, which has been the selected approach for this study. In this table we can see that in most cases the 2-tuples based genetic tuning obtains better results in performance than the basic FRBCS.

This study is analyzed by a Wilcoxon Test (Table 5), where we show that the 2-tuples tuning improves the behaviour of the simple KB, both in the global and local approaches. Therefore, we stress the goodness of the 2-tuples methodology for the tuning of the MF in imbalanced data-sets, both for the whole rule set (global approach) and for each fuzzy rule (local approach).

**Table 5. Wilcoxon's Test to compare the simple FRBCS approaches ( $R^+$ ) with the use of 2-tuples based genetic tuning ( $R^-$ )**

Comparison	$R^+$	$R^-$	Hypothesis ( $\alpha = 0.05$ )	p-value
Chi-3 vs. Chi-3-GTS	99.5	461.5	Rejected for Chi-3-GTS	0.001
Chi-3 vs. Chi-3-LTR	100.5	460.5	Rejected for Chi-3-LTR	0.001

Finally, comparing both tuning approaches (GTS and LTR) neither of them is better in performance, so initially there is no argument to support which one is preferable. Nevertheless, since the GTS works at the Data Base level, the global interpretability of the final FRBCS is maintained and, in this case, it achieves a good trade-off between accuracy and interpretability.

## 6 Conclusions

In this work we have adapted the 2-tuples based genetic tuning to classification problems with imbalanced data-sets, in order to increase the performance of simple FRBCSs.

Our empirical and statistical results have supported our aim and thus, we have shown that the genetic tuning improves the behaviour of the FRBCS in imbalanced data-sets. Thus, we can conclude that the tuning step is a necessity, since it helps FRBCSs to obtain better results in this scenario.

**Acknowledgements** This work had been supported by the Spanish Ministry of Science and Technology under Projects TIN-2005-08386-C05-01 and TIN-2005-08386-C05-03.

## References

- [1] R. Alcalá, J. Alcalá-Fdez, and F. Herrera. A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. *IEEE Transactions on Fuzzy Systems*, in 15(4):616–635, 2007.
- [2] R. Barandela, J. S. Sánchez, V. García, and E. Rangel. Strategies for learning in class imbalance problems. *Pattern Recognition*, 36(3):849–851, 2003.

**Table 4. Detailed results table for Chi et al.’s learning method with 3 labels: basic approach and 2-tuples based genetic tuning**

Dataset	Chi-3		Chi-3-GTS		Chi-3-LTR	
	$GM_{Tr}$	$GM_{Tst}$	$GM_{Tr}$	$GM_{Tst}$	$GM_{Tr}$	$GM_{Tst}$
Data-Sets with Low Imbalance						
EcoliCP-IM	95.49	92.27 ± 5.93	99.09	<b>95.08 ± 3.80</b>	99.42	93.75 ± 4.48
Haberman	66.21	58.91 ± 6.03	73.38	58.09 ± 5.11	78.33	<b>59.99 ± 3.53</b>
Iris1	100.0	<b>100.0 ± 0.0</b>	100.0	<b>100.0 ± 0.0</b>	100.0	<b>100.0 ± 0.0</b>
Pima	72.31	66.80 ± 5.93	81.72	<b>72.77 ± 5.31</b>	86.49	72.29 ± 3.30
Vehicle3	88.10	85.54 ± 3.36	97.96	<b>92.12 ± 2.08</b>	99.43	92.04 ± 1.53
Wisconsin	98.07	88.91 ± 2.13	99.12	93.10 ± 0.97	99.19	<b>94.92 ± 1.07</b>
Yeast2	68.33	67.69 ± 1.91	76.41	71.19 ± 2.71	78.96	<b>71.23 ± 1.95</b>
Glass1	66.57	64.06 ± 3.51	88.29	80.63 ± 4.90	90.72	<b>81.28 ± 5.30</b>
Glass2	75.37	64.90 ± 6.91	85.70	65.27 ± 7.10	90.88	<b>74.26 ± 4.36</b>
Vehicle2	76.47	70.92 ± 4.34	85.84	<b>74.44 ± 2.81</b>	90.06	71.38 ± 3.38
Vehicle4	75.52	69.22 ± 4.89	86.98	<b>72.26 ± 3.81</b>	88.95	69.94 ± 4.02
Mean	80.22	75.38 ± 4.09	88.59	79.54 ± 3.51	91.13	<b>80.10 ± 2.99</b>
Data-Sets with Medium Imbalance						
Ecoli2	87.92	<b>85.28 ± 9.77</b>	94.59	84.03 ± 12.7	96.23	82.63 ± 10.9
GlassNW	94.05	85.83 ± 3.04	98.03	89.77 ± 5.60	99.12	<b>90.26 ± 2.88</b>
New-Thyroid2	92.32	87.44 ± 8.11	99.93	<b>99.16 ± 0.77</b>	100.0	96.21 ± 4.09
New-Thyroid3	94.70	89.81 ± 10.7	99.86	<b>97.16 ± 4.06</b>	100.0	95.99 ± 5.58
Page-Blocks	80.60	79.91 ± 4.29	90.97	<b>89.13 ± 1.55</b>	87.05	85.06 ± 10.3
Segment1	95.45	94.99 ± 0.45	99.78	98.83 ± 0.69	99.94	<b>99.11 ± 0.83</b>
Vehicle1	88.23	86.41 ± 3.06	96.35	87.42 ± 1.80	97.99	<b>89.69 ± 1.13</b>
Ecoli3	89.66	88.01 ± 5.45	93.69	87.31 ± 5.22	94.60	<b>88.06 ± 7.42</b>
Yeast4	91.37	90.13 ± 4.09	94.85	91.62 ± 2.91	96.09	<b>91.75 ± 2.75</b>
Ecoli4	89.24	87.58 ± 4.08	96.36	<b>91.40 ± 5.19</b>	96.66	90.26 ± 5.60
Glass7	95.04	83.87 ± 9.82	98.19	<b>84.98 ± 7.33</b>	98.19	82.86 ± 7.03
Mean	90.78	87.21 ± 5.72	96.60	<b>90.98 ± 4.35</b>	96.90	90.17 ± 5.33
Data-Sets with High Imbalance						
Abalone9-18	69.80	63.93 ± 11.0	83.23	<b>71.96 ± 11.7</b>	87.16	69.70 ± 15.4
Abalone19	70.39	<b>62.96 ± 8.27</b>	83.81	62.58 ± 17.6	87.60	54.26 ± 18.0
Ecoli5	94.04	91.27 ± 7.43	99.17	<b>92.36 ± 7.31</b>	99.56	91.18 ± 6.68
Glass3	58.00	47.67 ± 10.2	90.29	69.72 ± 15.1	96.13	<b>74.98 ± 15.2</b>
Yeast5	83.44	<b>82.99 ± 3.10</b>	90.47	82.07 ± 4.82	91.83	77.38 ± 9.95
Vowel0	98.56	98.37 ± 0.61	99.94	98.03 ± 2.92	99.96	<b>99.38 ± 1.23</b>
YeastCYT-POX	75.66	72.75 ± 15.0	84.49	71.83 ± 14.7	88.73	<b>75.64 ± 16.2</b>
Glass5	95.15	84.96 ± 13.8	99.63	88.78 ± 12.6	99.94	<b>91.68 ± 12.2</b>
Glass6	94.15	<b>81.56 ± 12.6</b>	99.33	71.72 ± 41.7	99.82	78.26 ± 43.8
Yeast6	94.67	93.41 ± 5.35	98.02	<b>95.85 ± 2.65</b>	98.96	95.02 ± 2.79
Yeast7	88.43	<b>87.50 ± 10.5</b>	93.57	83.76 ± 8.18	94.95	84.06 ± 8.62
Mean	83.85	78.85 ± 8.90	92.90	80.79 ± 12.7	94.97	81.05 ± 13.6
All Data-Sets						
Mean	84.95	80.48 ± 6.24	92.70	<b>83.77 ± 6.84</b>	94.33	<b>83.77 ± 7.32</b>

[3] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard. A study of the behaviour of several methods for balancing machine learning training data. *SIGKDD Explorations*, 6(1):20–29, 2004.

[4] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligent Research*, 16:321–357, 2002.

[5] N. V. Chawla, N. Japkowicz, and A. Kolcz. Editorial: special issue on learning from imbalanced data sets. *SIGKDD Explorations*, 6(1):1–6, 2004.

[6] Z. Chi, H. Yan, and T. Pham. *Fuzzy algorithms with applications to image processing and pattern recognition*. World Scientific, 1996.

[7] L. J. Eshelman. *Foundations of Genetic Algorithms*, chapter The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, pages 265–283. Morgan Kaufman, 1991.

[8] A. Fernández, S. García, M. J. del Jesus, and F. Herrera. A study of the behaviour of linguistic fuzzy rule based classification systems in the framework of imbalanced data-sets.

*Fuzzy Sets and Systems*, In Press, doi:10.1016/j.fss.2007.12.023, 2008.

[9] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.

[10] H. Ishibuchi, T. Nakashima, and M. Nii. *Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining*. Springer-Verlag, 2004.

[11] A. Orrioles-Puig and E. Bernadó-Mansilla. Evolutionary rule-based systems for imbalanced datasets. *Soft Computing*, In Press (doi: 10.1007/s00500-008-0319-7), 2008.

[12] D. Sheskin. *Handbook of parametric and nonparametric statistical procedures*. Chapman & Hall/CRC, 2003.