Linguistic Hedges Fuzzy Feature Selection for Differential Diagnosis of Erythemato-Squamous Diseases

Ahmad Taher Azar[1], Shaimaa A. El-Said[2],
Valentina Emilia Balas[3], and Teodora Olariu[4]

[1] Misr University for Science & Technology (MUST), 6th of October City, Egypt.
[2] Faculty of Engineering, Zagazig University, Zagazig, Sharkia, Egypt
[3] Aurel Vlaicu University of Arad, Romania
[4] Vasile Goldis Western University of Arad, Romania
Ahmad_t_azar@ieee.org, eng.sahmed@windowslive.com,
balas@drbalas.ro, olariu_teodora@yahoo.com

Abstract. The differential diagnosis of erythemato-squamous diseases is a real challenge in dermatology. In diagnosing of these diseases, a biopsy is vital. However, unfortunately these diseases share many histopathological features, as well. Another difficulty for the differential diagnosis is that one disease may show the features of another disease at the beginning stage and may have the characteristic features at the following stages. In this paper, a new Feature Selection based on Linguistic Hedges Neural-Fuzzy classifier is presented for the diagnosis of erythemato-squamous diseases. The performance evaluation of this system is estimated by using four training-test partition models: 50–50%, 60–40%, 70–30% and 80–20%. The highest classification accuracy of 95.7746% was achieved for 80–20% training-test partition using 3 clusters and 18 fuzzy rules, 93.820% for 50–50% training-test partition using 3 clusters and 18 fuzzy rules, 92.5234% for 70–30% training-test partition using 5 clusters and 30 fuzzy rules, and 91.6084% for 60–40% training-test partition using 6 clusters and 36 fuzzy rules. Therefore, 80–20% training-test partition using 3 clusters and 18 fuzzy rules are the best classification accuracy with RMSE of 6.5139e-013. This research demonstrated that the proposed method can be used for reducing the dimension of feature space and can be used to obtain fast automatic diagnostic systems for other diseases.

Keywords: Erythemato-Squamous Diseases, Soft Computing, Takagi-Sugeno-Kang (TSK) fuzzy inference system, Linguistic Hedge (LH), Feature selection (FS).

1 Introduction

Decision support systems helping physicians play an important role in medical decision making particularly in those situations where decision must be made effectively and reliably. The differential diagnosis of erythemato-squamous diseases is a difficult

V.E. Balas et al. (Eds.): Soft Computing Applications, AISC 195, pp. 487–500.
springerlink.com

problem in dermatology (Barati et al. 2011). There are six classes of erythemato-squamous diseases. They are psoriasis (C_1), seboreic dermatitis (C_2), lichen planus (C_3), pityriasis rosea (C_4), chronic dermatitis (C_5) and pityriasis rubra pilaris (C_6). Usually a biopsy is necessary for the correct and definite diagnosis but unfortunately these diseases share many histopathological features (Govenir and Emeksiz 2000). Another problem for the differential diagnosis is that a disease may show the features of another disease at the beginning stage and may have the characteristic features at the following stages. Patients were first evaluated clinically with 12 features. Afterwards, skin samples were taken for the evaluation of 22 histopathological features. The values of these histopathological features are determined by an analysis of the samples under a microscope (Guvenir et al. 1998). Therefore, various new methods have been used for diagnosis of Erythemato-Squamous diseases like Artificial Neural Network (ANN) (Kabari and Bakpo 2009; Übeyli 2009; Karabatak and Ince 2009), Support Vector Machines (SVM) (Abdi and Giveki 2012; Xie and Wang 2011; Xie et al. 2010, 2012; Ubeyli 2008; Nanni 2006), Fuzzy expert systems and neuro-fuzzy classification (Luukka 2011a, b; Lekkas and Mikhailov 2010; Parthiban and Subramanian 2009; Polat and Günes 2006; Ubeyli and Güler 2005), nearest neighbor classifier (NNC) (Polat and Günes 2006; Govenir and Emeksiz 2000), Instance-based classifiers (Gagliardi 2011), rough sets (Revett et al. 2009), decision trees classifiers (DTC) (Ozcift and Gulten 2012; Polat and Gunes 2009; Polat and Günes 2006; Guvenir et al. 1998).

Feature selection methods have been successfully used in many areas especially in medical diagnosis to reduce the size of features collected during the clinical testing and experiments (Chrysostomou 2008). The process of feature selection is very important because it reduces the dimensionality of the data and enables learning algorithms to operate faster (reduction of the computation time), more efficiently and therefore increases the accuracy of the resulting model (Guyon et al. 2006). Therefore, this paper presents a fuzzy feature selection (FS) method based on the linguistic hedges (LH) concept (Cetişli 2010a, b) for Erythemato-Squamous diseases classification. The proposed method is used to achieve a very fast, simple and efficient computer aided diagnosis (CAD) system. It helps to reduce the dimensionality of the used data set, to speed up the computation time of a learning algorithm and therefore simplifies the classification task.

The rest of this paper is organized as follows: in Section 2, a review of the classifiers that are considered in Erythemato-Squamous diseases diagnosis. Section 3 provides subjects and methods that are used in this study. In Section 4, a detailed description of the proposed feature selection method that is considered in Erythemato-Squamous diseases diagnosis is presented. Section 5 reports the results of experimental evaluations of the adaptive neural-fuzzy classifier and finally, in Section 6, conclusion and directions for future research are presented.

2 Related Works

Güvernir et al. (1998) developed a classification algorithm VFI5 called ''Voting Feature Intervals'' and it was applied for the differential diagnosis of erythemato-squamous diseases. Genetic algorithm for learning the feature weights was used with the Nearest Neighbor classification algorithm and was applied to determine the weights of the features to be used with the VFI5 algorithm which achieved 96.2% accuracy on the Dermatology dataset. Govenir and Emeksiz (2000) presented an expert system for differential diagnosis erythemato-squamous diseases incorporating decisions made by three classification algorithm: nearest neighbor classifier, naive Bayesian classifier and voting feature intervals-5. The features are assigned weights such that the irrelevant features have lower weights while the strongly relevant features are given higher weights. Govenir and Emeksiz (2000) used a genetic algorithm to learn the feature weights to be used with the classification algorithm. Further research by Übeyli and Güler (2005) showed a new approach based on adaptive neuro-fuzzy inference system (ANFIS) for the detection of erythemato-squamous diseases, the ANFIS model was assessed in terms of training performance and classification accuracies and showed to perform well in detecting erythemato-squamous diseases, the model achieved a total classification accuracy of 95.5% which is concluded to be good in comparison to the one of 85.5% achieved with the stand-alone neural network. Nanni (2006) investigated an ensemble of support vector machines (SVM) classifier based on random subspace (RS) and feature selection for the diagnosis of erythemato-squamous diseases, and tested it on a real-world dataset. The results improved the average predictive accuracy obtained by a ‘‘stand-alone’’ SVM or by a RS ensemble of SVM (Nani 2006). Polat and Günes (2006) developed a novel method for differential diagnosis of erythemato-squamous disease. Their method was based on fuzzy weighted pre-processing, k-NN (nearest neighbor) based weighted pre-processing and decision tree classifier. Luukka and Leppälampi (2006) presented a new approach based on similarity classifier with generalized mean and applied it for medical data. The presented method was managed to detect erythemato-squamous diseases with a good mean classification accuracy of 97.02%. In recent published study by Luukka (2011a), fuzzy entropy measures were used in feature selection. This method successfully managed to discard the non-important features in the datasets; this has positively facilitated the classification task which was done by using the similarity based on Lukasiewicz structure where a mean accuracy of 98.28% was achieved. In Luukka (2011b) study, a new nonlinear fuzzy robust principal component analysis (NFRPCA) algorithm was developed to get data into more feasible form. This new nonlinear fuzzy robust principal component analysis algorithm achieved a classification accuracy of 97.09 % accuracy with dermatology data. Ozcift and Gulten (2012) proposed a new multi-class feature selection method based on Rotation Forest meta- learner algorithm and was tested on Erythemato-Squamous diseases dataset. The Rotation Forest selection based features achieved accuracies between 98% and 99% in various classifiers. Xie et al. (2012) proposed hybrid feature selection algorithms based on generalized F-score (GF) and SVM. The new hybrid feature selection algorithms combined the strengths of filters and wrappers to uncover the optimal feature subset with the best diagnostic efficiency. Experimental results showed that the proposed hybrid methods construct efficient diagnosis classifiers with high average accuracy when compared with traditional algorithms.

3 Subjects and Methods

The dataset was taken from UCI (University of California, Irvine) machine learning repository (UCI 2012). There are 366 records in this database and each record has 34 attributes; 33 of which are linear (continuous) valued and one of them is nominal (discrete). After the 9 instances are removed from the dataset due to 8 missing values

Table 1. Erythemato-squamous diseases class distribution

Class Number	Class Name	No. of instances
1	Psoriasi	110
2	Seboreic dermatitis	60
3	Lichen planus	71
4	Pityriasis rosea	48
5	Cronic dermatitis	48
6	Pityarisis rubra pilaris	20
Total		357

Table 2. Description of Erythemato-squamous diseases attributes

Clinical Attributes	Histopathological Attributes
Att.1: erythema Att.2: scaling Att.3: definite borders Att.4: itching Att.5: koebner phenomenon Att.6: polygonal papules Att.7: follicular papules Att.8: oral mucosal involvement Att.9: knee and elbow involvement Att.10: scalp involvement Att.11: family history Att.34: age	Att.12: melanin incontinence Att.13: eosinophils in infiltrate Att.14: PNL infiltrate Att.15: brosis of the papillary dermis Att.16: exocytosis Att.17: acanthosis Att.18: hyperkeratosis Att.19: parakeratosis Att.20: clubbing of the rete ridges Att.21: elongation of the rete ridges Att.22: thinning of the suprapapillary epidermis Att.23: pongiform pustule Att.24: munro microabscess Att.25: focal hypergranulosis Att.26: disappearance of the granular layer Att.27: vascularization and damage of basal layer Att.28: spongiosis Att.29: saw-tooth appearance of retes Att.30: follicular horn plug Att.31: perifollicular parakeratosis Att.32: inflammatory mononuclear infiltrate Att. 33: band-like infiltrate

and one zero value of age, there are 357 instances. Distribution according to class variable of this dataset is given in Table 1. Each record contains 12 clinical features and 22 histopathological features (see Table 2). The family history feature in the dataset has the value 1 if any of these diseases has been observed in the family and 0 otherwise. The age feature simply represents the age of the patient. Every other feature (clinical and histopathological) was given a degree in the range of 0–3. Here, 0 indicates that the feature was not present; a 3 indicates the largest amount possible and 1, 2 indicate the relative intermediate values.

4 Adaptive Neuro-Fuzzy Classifiers

The usage of ANFIS (Jang 1993; Jang and Sun 1995, Jang et al. 1997) for classifications is unfavorable. For example, if there are three classes labeled as 1, 2 and 3. The ANFIS outputs are not integer. For that reason the ANFIS outputs are rounded, and determined the class labels. But, sometimes, ANFIS can give 0 or 4 class labels. These situations are not accepted. As a result ANFIS is not suitable for classification problems. In this section, adaptive neuro-fuzzy classifier is discussed in details. In these models, k-means algorithm is used to initialize the fuzzy rules. Also, Gaussian membership function is only used for fuzzy set descriptions, because of its simple derivative expressions.

4.1 Adaptive Neuro-Fuzzy Classifier with Linguistic Hedges (ANFCLH)

Adaptive neuro-fuzzy classifier (ANFC) with Linguistic hedges (Cetişli 2010a, b) is based on fuzzy rules. Linguistic hedges are applied to the fuzzy sets of rules, and are adapted by Scaled Conjugate Gradient (SCG) algorithm. By this way, some distinctive features are emphasized by power values, and some irrelevant features are damped with power values. The power effects in any feature are generally different for different classes. The using of linguistic hedges increases the recognition rates. A fuzzy classification rule that has two inputs$\{x_1, x_2\}$ and one output y is defined with LHs as IF x_1 is A_1 with p_1 hedge AND x_2 is A_2 with p_2 hedge THEN y is C_1 class, where A_1 and A_2 denote linguistic terms that are defined on X_1 and X_2 feature space; p_1 and p_2 denote linguistic hedges, respectively; C_1 denotes the class label of the output y. Fig. 1 shows the ANFCLH architecture.

The feature space with two inputs $\{x_1, x_2\}$ is partitioned into three classes $\{C_1, C_2, C_3\}$, in the Figure. The feature space $X_1 \times X_2$ is separated into fuzzy regions (Jang et al. 1997). This technique is based on zero-order Sugeno fuzzy model (Takagi and Sugeno 1985). The crisp outputs of fuzzy rules are determined by weighted average operator (Jang et al. 1997). In this classifier, the nodes in the same layer have the same type of node functions. The layers and their properties are given as follows:

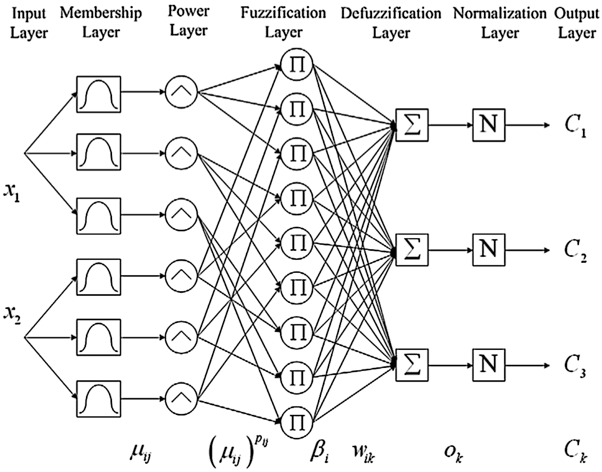

Fig. 1. A neuro-fuzzy classifier with LHs (Cetişli 2010a)

Layer 1: In this layer, the membership grade of each input to specified fuzzy region is measured. Gaussian function is employed as MF due to smooth partial derivatives of its parameters, and has less parameter. The Gaussian MF is given as follows (Cetişli 2010a):

$$\boxed{\mu_{ij}(x_{sj}) = \exp\left(-0.5\frac{(x_{sj}-c_{ij})^2}{\sigma_{ij}^2}\right)} \tag{1}$$

where $\mu_{ij}(x_{sj})$ represents the membership grade of the *i*th rule and the *j*th feature; x_{sj} denotes the *s*th sample and the *j*th feature of input matrix $X\{X \in R^{NxD}\}$; c_{ij} and σ_{ij} are the center and the width of Gaussian function, respectively.

Layer 2: In this layer, the secondary meanings of fuzzy sets are calculated with their LHs as in Eq. (2) (Cetişli 2010a):

$$\boxed{\alpha_{ijs} = [\mu_{ij}(x_{sj})]^{p_{ij}}} \tag{2}$$

where α_{ijs} denotes the modified membership grades of μ_{ij} (x_{sj}); p_{ij} denotes the LH value of the *i*th rule and the *j*th feature.

Layer 3: The degree of fulfillment of the fuzzy rule for x s sample is determined in this layer. It is also called as the firing strength of rule. So, the B_{is} firing strength of the *i*th rule for D number of features is defined as in Eq. (3) (Cetişli 2010a):

$$\boxed{\mathrm{B}_{is} = \prod_{j=1}^{D}\alpha_{ijs}} \tag{3}$$

Layer 4: In this layer, the weighted outputs are calculated as in Eq. (4) (Cetişli 2010a), and every rule can affect each class according to their weights. However, if a rule controls a specific class region, the weight between this rule output and the specific class is to be bigger than the other class weights. Otherwise, the class weights are fairly small:

$$\boxed{O_{sk} = \sum_{i=1}^{U} \beta_{is} w_{ik}} \tag{4}$$

where w_{ik} represents the degree of belonging to the kth class that is controlled with the ith rule; O_{sk} denotes the weighted output for the sth sample that belongs to the kth class, and U is the number of rules.

Layer 5: If the summation of weights is bigger than 1, the outputs of the network should be normalized in the last layer as follows (Cetişli 2010a):

$$\boxed{h_{sk} = \frac{O_{sk}}{\sum_{l=1}^{k} O_{sl}} = \frac{O_{sk}}{\delta_s}, \delta_s = \sum_{l=1}^{k} O_{sl}} \tag{5}$$

Where h_{sk} represents the normalized degree of the sth sample that belongs to the kth class; and K is the number of classes. After then, the class label (C_s) of sth sample is determined by the maximum h_{sk} value as in Eq. (6) (Cetişli 2010a):

$$\boxed{C_s = \max_{k=1,2,\ldots,k} \{h_{sk}\}} \tag{6}$$

The antecedent parameters of the network $\{c, \sigma, p\}$ could be adapted by any optimization method. In this study, scaled conjugate gradient (SCG) method is used to adapt the network parameters (Moller 1993). The cost function that is used in the SCG method is determined from the least mean squares of the difference target and the calculated class value (Jang et al. 1997; Sun and Jang 1993). According to the above definition, the cost function E is defined as in Eq. (7) (Cetişli 2010a):

$$\boxed{E = \frac{1}{N}\sum_{s=1}^{N} E_s,\ E_s = \frac{1}{2}\sum_{k=1}^{K}(t_{sk} - h_{sk})^2} \tag{7}$$

4.2 Linguistic Hedges Neural-Fuzzy Classifier with Selected Features (LHNFCSF)

Cetişli (2010a, b) presented a fuzzy feature selection (FS) method based on the LH concept. It uses the powers of fuzzy sets for feature selection. The values of LHs can be used to show the importance degree of fuzzy sets. When this property is used for classification problems, and every class is defined by a fuzzy classification rule, the LHs of every fuzzy set denote the importance degree of input features. If the LHs values of features are close to concentration values, these features are more important or relevant, and can be selected. On the contrary, if the LH values of features are close to dilation values, these features are not important, and can be eliminated. According to the LHs value of features, the redundant, noisily features can be eliminated, and significant features can be selected. In this technique (Cetişli 2010b), if linguistic hedge values of classes in any feature are bigger than 0.5 and close to 1, this feature is relevant, otherwise it is irrelevant. The program creates a feature selection and a rejection criterion by using power values of features. There are two selection criteria, one is the selection of features that have the biggest hedge value for any class and the other is the selection of features that have a bigger hedge value for every class, because any feature cannot be selective for every class. For that reason, a selective function should be described from the hedge values of any feature as in Eq. (8) (Cetişli 2010b):

$$\boxed{p_j = \prod_{i=1}^{k} p_{ij}} \tag{8}$$

where P_j denotes the selection value of the jth feature, and K is the number of classes. The Feature selection and classification algorithms were discussed in detail in Cetişli (2010b).

5 Results and Discussions

In this section, the performance analysis is presented. The simulations were performed by using an Intel (R) Core (TM) i3 CPU 530-2.93 GHz personal computer and a Microsoft Windows 7 64-bit operating system. The core of the NFC calculations was implemented by using the MATLAB software package.

5.1 Training and Testing Phases of Classifier

The collection of well-distributed, sufficient, and accurately measured input data is the basic requirement in order to obtain an accurate model. The classification process starts by obtaining a data set (input-output data pairs) and dividing it into a training set and testing data set. The training data set is used to train the NFC, whereas the test data set is used to verify the accuracy and effectiveness of the trained NFC. Once the model structure and parameters have been identified, it is necessary to validate the quality of the resulting model. In principle, the model validation should not only validate the accuracy of the model, but also verify whether the model can be easily interpreted to give a better understanding of the modeled process. It is therefore important to combine data-driven validation, aiming at checking the accuracy and robustness of the model, with more subjective validation, concerning the interpretability of the model. There will usually be a challenge between flexibility and interpretability, the outcome of which will depend on their relative importance for a given application. While, it is evident that numerous cross-validation methods exist, the choice of the suitable cross-validation method to be employed in the NFC is based on a trade-off between maximizing method accuracy and stability and minimizing the operation time. To avoid overfitting problems during modeling process, the data set were randomly split into four training-test partition model sets: 50–50%, 60–40%, 70–30% and 80–20%. The number of training and test data for each of classes is given in Table 3.

Table 3. The number of training and test data for each of class

Erythemato-Squamous Class	50-50%	60-40%	70-30%	80-20%
Class-1: Psoriasi	51-59	61-49	75-35	86-24
Class-2: Seboreic dermatitis	41-19	49-11	53-7	58-2
Class-3: Lichen planus	37-34	47-24	53-18	58-13
Class-4: Pityriasis rosea	18-30	19-29	22-26	34-14
Class-5: Cronic dermatitis	23-25	28-20	35-13	35-13
Class-6: Pityarisis rubra pilaris	9-11	10-10	12-8	15-5
Total	179-178	214-143	250-107	286-71

In feature reduction stage of the ANFCLH for diagnosis of erythemato-squamous diseases, the feature extraction and the feature reduction processes are performed. The number of fuzzy rules is determined according to the number of classes. According to the feature selection algorithm, features 1-18 are common relevant features for each class while features from 19-34 are irrelevant for each class. The LHNFCSF classification results during training and testing phases of erythemato-squamous diseases are shown in Table 4 and also represented graphically in Fig. 2.

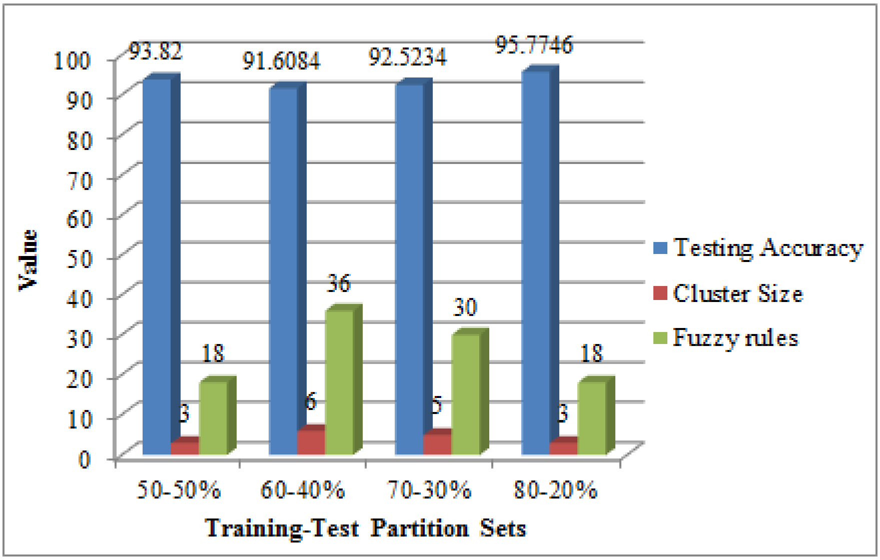

Fig. 2. LHNFCSF Classification results of erythemato-squamous diseases based on feature selection, fuzzy rules and cluster size for each class using for 80–20% training-test partition

Although the number of selected features is reduced to 18 from 34 features, the classification is still successful as seen in Table 4. The number of cluster size was opted between 1 and 6 to check the optimum topology for each model. With this assumption, different trials were performed to find the best Neural-fuzzy in each method using different cluster sizes with different fuzzy rules.

Table 4. The LHNFCSF classification results of erythemato-squamous diseases dataset

training-test partition sets	Features	Cluster size for each class	Training accuracy	Testing Accuracy	RMSE	No. of Rules	Epoch
50-50%	1-18	1	98.8827	91.0112	0.0259	6	75
		2	100	89.8876	4.8452e-013	12	325
		3	**100**	**93.820**	**5.5845e-013**	**18**	**300**
		4	100	92.6966	1.4583e-011	24	150
		5	100	91.0112	7.0032e-013	30	175
		6	100	91.573	5.9842e-013	36	175
60-40%	1-18	1	95.3271	90.2098	0.0755192	6	75
		2	100	88.1119	6.7298e-013	12	425
		3	99.065	90.9091	0.0198467	18	125
		4	100	90.9091	1.5839e-011	24	250
		5	100	88.8112	8.3029e-013	30	375
		6	**100**	**91.6084**	**0.000382048**	**36**	**75**
70-30%	1-18	1	98.8	86.9159	0.0241916	6	100
		2	100	90.6542	1.6108e-011	12	325
		3	100	92.5234	6.4936e-013	18	275
		4	100	91.5888	6.1730e-013	24	300
		5	**100**	**92.5234**	**9.2392e-013**	**30**	**275**
		6	100	89.7196	7.7009e-013	36	150
80-20%	1-18	1	98.951	92.9577	0.0289586	6	125
		2	99.6503	94.3662	0.00669146	12	825
		3	**100**	**95.7746**	**1.3057e-005**	**18**	**675**
		4	99.6503	94.3662	0.00657387	24	425
		5	100	91.5493	7.4613e-013	30	250
		6	100	92.9577	1.1043e-006	36	400

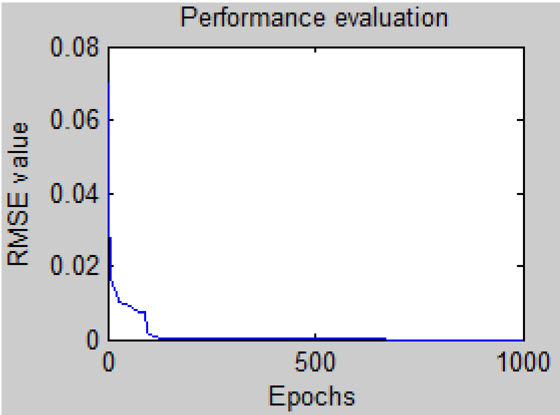

Fig. 3. Performance Evaluation of LHNFCSF after selection of relevant features using three clusters for 80–20% training-test partition

Among the four training-test partition sets, the highest classification accuracy of 95.7746% was achieved for 80–20% training-test partition using 3 clusters and 18 fuzzy rules, 93.820% for 50–50% training-test partition using 3 clusters and 18 fuzzy rules, 92.5234% for 70–30% training-test partition using 5 clusters and 30 fuzzy rules, and 91.6084% for 60–40% training-test partition using 6 clusters and 36 fuzzy rules. Therefore, 80–20% training-test partition using 3 clusters and 18 fuzzy rules are the best classification accuracy with RMSE of 6.5139e-013 as shown in Fig. 3.

It's noted from Table 4 also that the selected features according to four different partitions were always the same. The results indicated that, the selected features increase the recognition rate for test set. It means that some overlapping classes can be easily distinguished by selected features. The LH values of erythemato-squamous diseases for 6 features only as example are given in Table 5. After the classification step, it can be seen that some of the hedge values are bigger than 1, because the hedge values are not constrained in the classification step. Each class for ANFCLH is intuitively defined with 18 fuzzy rules. The neural fuzzy classifier surface of feature 1 and feature 2 using 80–20% training-test partition is shown in Fig. 4

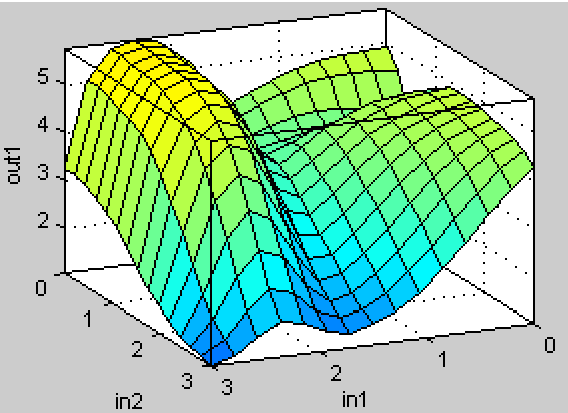

Fig. 4. Neural Fuzzy classifier surface using three clusters for 80–20% training-test partition

Table 5. The LH values of erythemato-squamous for every class after selection of relevant features (values for only 6 from 18 relevant features)

Class	F1	F2	F3	F4	F5	F6
C1	1.004206	1.021716	1.042981	1.011113	1.031387	1.000265
C2	1.027196	0.9907	0.984645	1.039615	1.011038	1.000047
C3	0.981362	0.992032	0.984563	0.984836	0.998278	0.955953
C4	1.02159	1.038587	1.006401	1.013554	1.042992	1.0066
C5	0.969889	0.993738	0.990568	0.991844	1.009644	1.003435
C6	1.03143	1.005849	1.001836	1.008994	1.00656	1

6 Conclusion

In this study, the proposed Feature Selection based on Linguistic Hedges Neural-Fuzzy classifier managed to discard redundant and irrelevant features in erythemato-squamous disease dataset. The computation time was reduced and this has positively impacted to achieve highest classification accuracy of 95.7746% for 80–20% training-test partition using 3 clusters and 18 fuzzy rules. Experimental results showed that when the linguistic hedge value of the fuzzy classification set in any feature is close to 1, this feature is relevant for that class, otherwise it may be irrelevant. The results demonstrate that using of Linguistic Hedges in daptive neural-fuzzy classifier improves the success of the classifier. Also, when the ANFCLH is compared with the other methods, it can be noted that this classifier uses less parameter than the others with the best recognition rates. The results strongly suggest that the proposed method not only help to reduce the dimensionality of large data sets but also to speed up the computation time of a learning algorithm and therefore simplify the classification task. It can aid in the diagnosis of erythemato-squamous diseases and can be very helpful to the physicians for their final decision on their patients. The future investigation will pay much attention to evaluate Neural-Fuzzy classifier with Linguistic Hedges in other medical diagnosis problems like micro array gene selection, internet, and other data mining problems. Therefore, the impressive results may be obtained with the proposed method and improving the performance of NFCs using high-performance computing techniques. In addition, the combination of the approaches mentioned above will yield an efficient fuzzy classifier for a lot of applications.

References

1. Abdi, M.J., Giveki, D.: Automatic detection of erythemato-squamous diseases using PSO–SVM based on association rules. Eng. Appl. Artif. Intel. (2012), doi:10.1016/j.engappai.2012.01.017
2. Barati, E., Saraee, M., Mohammadi, A., et al.: A Survey on Utilization of Data Mining Approaches for Dermatological (Skin) Diseases Prediction. Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Health Informatics (JSHI): March Edition, 1–11 (2011)
3. Cetişli, B.: Development of an adaptive neuro-fuzzy classifier using linguistic hedges: Part 1. Expert Systems with Applications 37(8), 6093–6101 (2010a)
4. Cetişli, B.: The effect of linguistic hedges on feature selection: Part 2. Expert Systems with Applications 37(8), 6102–6108 (2010b)
5. Chrysostomou, K.A.: The Role of Classifiers in Feature selection Selection: Number vs Nature, PhD dissertation, School of Information Systems. Computing and mathematics, Brunel University (2008)
6. Diamantidis, N.A., Karlis, D., Giakoumakis, E.A.: Unsupervised stratification of cross-validation for accuracy estimation. Artif. Intell. 116(1-2), 1–16 (2000)
7. Francois, D., Rossi, F., Wertz, V.: Verleysen M Resampling methods for parameter-free and robust feature selection with mutual information. Neurocomputing 70, 1276–1288 (2007)

8. Gagliardi, F.: Instance-based classifiers applied to medical databases: diagnosis and knowledge extraction. Artif. Intell. Med. 52(3), 123–139 (2011)
9. Govenir, H.A., Emeksiz, N.: An expert system for the differential diagnosis of erythemato-squamous diseases. Expert Syst. Appl. 18(1), 43–49 (2000)
10. Güvenir, H.A., Demiroz, G., Ilter, N.: Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals. Artif. Intell. Med. 13(3), 147–165 (1998)
11. Guyon, I., Steve, G., Nikravesh, M., Zadeh, L.A.: Feature Extraction: Foundations and Applications. Springer, Heidelberg (2006)
12. Jang, J.S.R.: ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics 23(3), 665–685 (1993)
13. Jang, J.S.R., Sun, C.T.: Neuro-fuzzy modeling and control. Proceedings of the IEEE 83(3), 378–406 (1995)
14. Jang, J.S.R., Sun, C.T., Mizutani, E.: Neuro-Fuzzy and soft computing. Prentice-Hall, Englewood Cliffs (1997)
15. Kabari, L.G., Bakpo, F.S.: Diagnosing skin diseases using an artificial neural network. In: Proceeding of 2nd International Conference on Adaptive Science & Technology (ICAST 2009), pp. 187–191. IEEE Press, New York (2009)
16. Karabatak, M., Ince, M.C.: A new feature selection method based on association rules for diagnosis of erythemato-squamous diseases. Expert Syst. Appl. 36(10), 12500–12505 (2009)
17. Lekkas, S., Mikhailov, L.: Evolving fuzzy medical diagnosis of Pima Indians diabetes and of dermatological diseases. Artif. Intell. Med. 50(2), 117–126 (2010)
18. Luukka, P.: Feature selection using fuzzy entropy measures with similarity classifier. Expert Syst. Appl. 38(4), 4600–4607 (2011a)
19. Luukka, P.: A New Nonlinear Fuzzy Robust PCA Algorithm and Similarity Classifier in Classification of Medical Data Sets. International Journal of Fuzzy Systems 13(3), 153–162 (2011b)
20. Luukka, P., Leppälampi, T.: Similarity classifier with generalized mean applied to medical data. Comput. Biol. Med. 36(9), 1026–1040 (2006)
21. Moller, M.F.: A scaled conjugate gradient algorithm for fast supervised learning. Neural Networks 6, 525–533 (1993)
22. Nanni, L.: An ensemble of classifiers for the diagnosis of erythemato-squamous diseases. Neurocomputing 69(7-9), 842–845 (2006)
23. Ozcift, A., Gulten, A.: A Robust Multi-Class Feature Selection Strategy Based on Rotation Forest Ensemble Algorithm for Diagnosis of Erythemato-Squamous Diseases. J. Med. Syst. 36(2), 941–949 (2012)
24. Parthiban, L., Subramanian, R.: An intelligent agent for detection of erythemato-squamous diseases using Co-active Neuro-Fuzzy Inference System and genetic algorithm. In: International Conference on Intelligent Agent & Multi-Agent Systems, IAMA, pp. 1–6. IEEE Press, New York (2009)
25. Polat, K., Gunes, S.: A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems. Expert Systems with Applications 36(2), 1587–1592 (2009)
26. Polat, K., Günes, S.: The effect to diagnostic accuracy of decision tree classifier of fuzzy and k-NN based weighted pre-processing methods to diagnosis of erythema-to-squamous diseases. Digital Signal Processing 16(6), 922–930 (2006)
27. Revett, K., Gorunescu, F., Salem, A.B., El-Dahshan, E.S.: Evaluation of the feature space of an erythematosquamous dataset using rough sets. Ann. Univers. Craiova, Romania, Math. Comput. Sci. Ser. 36(2), 123–130 (2009)

28. Salzberg, S.L.: On comparing classifiers: Pitfalls to avoid and a recommended approach. Data Mining and Knowledge Discovery 1, 317–327 (1997)
29. Sun, C.T., Jang, J.S.R.: A neuro-fuzzy classifier and its application. In: Proc. of IEEE Int. Conf. on Fuzzy Systems, vol. 1, pp. 94–98. IEEE Press, San Francisco (1993)
30. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man. Cybern. 15(1), 116–132 (1985)
31. Xie, J., Lei, J., Xie, W., Gao, X., Shi, Y., Liu, X.: Novel Hybrid Feature Selection Algorithms for Diagnosing Erythemato-Squamous Diseases. In: He, J., Liu, X., Krupinski, E.A., Xu, G. (eds.) HIS 2012. LNCS, vol. 7231, pp. 173–185. Springer, Heidelberg (2012)
32. Xie, J., Wang, C.: Using support vector machines with a novel hybrid feature selection method for diagnosis of erythemato-squamous diseases. Expert Syst. Appl. 38(5), 5809–5815 (2011)
33. Xie, J.Y., Xie, W.X., Wang, C.X., et al.: A novel hybrid feature selection method based on IFSFFS and SVM for the diagnosis of erythemato-squamous diseases. In: JMLR Workshop and Conference Proceedings. Workshop on Applications of Pattern Analysis, vol. 11, pp. 142–151. MIT Press, Windsor (2010)
34. UCI, Machine Learning Repository, `http://archive.ics.uci.edu/ml/index.html` (accessed on July 12, 2012)
35. Übeyli, E.D.: Combined neural networks for diagnosis of erythemato-squamous diseases. Expert Syst. Appl. 36(3), 5107–5112 (2009)
36. Ubeyli, E.D.: Multiclass support vector machines for diagnosis of erythemato-squamous diseases. Expert Syst. Appl. 35(4), 1733–1740 (2008)
37. Ubeyli, E.D., Güler, I.: Automatic detection of erythemato-squamous diseases using adaptive neuro-fuzzy inference systems. Computers in Biology and Medicine 35(5), 147–165 (2005)