Microarray Feature Selection and Dynamic Selection of Classifiers for Early Detection of Insect Bite Hypersensitivity in Horses

Conference Paper · April 2019

CITATIONS
0

READS
50

6 authors, including:

Alexandre Maciel Guerra
University of Nottingham
2 PUBLICATIONS **5** CITATIONS
SEE PROFILE

Grazziela P. Figueredo
University of Nottingham
65 PUBLICATIONS **155** CITATIONS
SEE PROFILE

Fernando J. Von Zuben
University of Campinas
346 PUBLICATIONS **6,621** CITATIONS
SEE PROFILE

Eliane Marti
Universität Bern
174 PUBLICATIONS **2,952** CITATIONS
SEE PROFILE

Microarray Feature Selection and Dynamic Selection of Classifiers for Early Detection of Insect Bite Hypersensitivity in Horses

Alexandre Maciel-Guerra
School of Computer Science
The University of Nottingham
Nottingham, UK
Alexandre.MacielGuerra@nottingham.ac.uk

Grazziela P. Figueredo
School of Computer Science
The University of Nottingham
Nottingham, UK
pmzgf@nottingham.ac.uk

Fernando J. Von Zuben
School of Electrical and Computer Engineering
University of Campinas
Campinas, Brazil
vonzuben@dca.fee.unicamp.br

Eliane Marti
Clinical Research and Veterinary Public Health
University of Bern
Bern, Switzerland
eliane.marti@vetsuisse.unibe.ch

Jamie Twycross
School of Computer Science
The University of Nottingham
Nottingham, UK
pszjpt@nottingham.ac.uk

Marcos J. C. Alcocer
School of Biosciences
The University of Nottingham
Sutton Bonington, UK
Marcos.Alcocer@nottingham.ac.uk

***Abstract*—Microarrays can be employed to better characterise allergies, as interactions between antibodies and allergens in mammals can be monitored. Once the joint dynamics of these elements in both healthy and diseased animals are understood, a model to predict the likelihood of an individual having allergic reactions can be defined. We investigate the potential use of Dynamic Selection (DS) methods to classify protein microarray data, with a case study of equine insect bite hypersensitivity (IBH) disease. To the best of our knowledge DS has not yet been applied to these data types. Since most microarrays datasets have a low number of samples, we hypothesise that DS models will produce satisfactory results due to their ability to perform better when compared to traditional ensemble techniques for similar data. We focus on three research questions: 1) What is the potential of DS for microarray data classification and how does it compare with existing classical classification methods results? 2) how do DS methods perform for the IBH dataset? and 3) does feature selection improve DS performance for this data? A wrapper using backward elimination and embedded with a regularized extreme learning machine are adopted to identify the more relevant features influencing the onset of the disease. Results from traditional classifiers are compared to 21 different DS methods before and after performing feature selection. Our results indicate that DS methods do not outperform single and static classifiers on this high-dimensional dataset and their performance also does not improved after feature selection.**

***Index Terms*—dynamic selection, feature selection, protein microarrays, allergies, insect bite hypersensitivity**

I. Introduction

Protein microarrays are a powerful tool employed in allergy diagnostics, as it monitors interactions between the immune system and allergens. In microarray data, there is information regarding the fluorescence of binding signals, which are proportional to the concentration of an antibody's reaction to each spot containing allergens in the microarray. As healthy and unhealthy animals are expected to mount different immune responses to allergens, the analysis of existing microarray data should enable the determination of prediction models for early diagnosis of allergies. Another important aspect of the study of microarray data is that it generally carries a significant number of irrelevant features leading to miss classification. The determination and pruning of those irrelevant features tend to promote performance improvement.

In this work we investigate equine insect bite hypersensitivity (IBH). IBH is a well-characterised equine immune response (involving IgE antibodies) to ordinary salivary proteins from insects, with a known aetiology and fully determined clinical symptoms, as described in Marti *et al.* [1]. The IBH microarray data was collected and initially investigated by Marti *et al.* [1] using a Partial Least Square Discriminant Analysis (PLSDA) approach. Later, in Maciel-Guerra *et al.* [2], the same dataset was also analysed using 9 single and ensemble classifiers with a wrapper backward selection (WBS) with a regularized extreme learning machine (RELM) as the embedded model to perform feature selection. Results showed similar classification performance without the need of expert input. Furthermore, the most important feature identified by their method matched that identified in Marti *et al.* [1]. It was also observed that WBS increased the AUC in all classifiers tested. Furthermore, RELM, which had not yet been largely explored in microarray data, exhibits a very satisfactory performance for this dataset.

Recently, dynamic selection (DS) methods have been employed by researchers due to their ability to perform better, in terms of accuracy, on datasets with low sample sizes when compared to traditional ensemble techniques, such as majority voting and static selection. The IBH dataset analysed in this paper have this characteristic of having a small number of samples, 109 in this case. Therefore, DS methods should perform well on this dataset. Three research questions are the

focus of this paper: 1) how DS methods perform in comparison to other well established classification methods? and 2) is DS a better alternative to the existing approach to the IBH data? 3) does feature selection improve the performance of DS methods?

We therefore employ and compare the results of nine state-of-the-art classification methods (Logistic Regression [3], Linear and Non-linear Support Vector Machines (SVM) [4], [5], Random Forest [6], Multi-Layer Perceptron (MLP) Neural Networks [7], [8], AdaBoost [9], Naive Bayes [10], Linear Discriminant Analysis (LDA) [11], [12] and Regularized Extreme Learning Machine (RELM) [13], [14]) with 21 DS methods. In addition, a wrapper backward selection (WBS) with a RELM as the embedded model is applied to the data for feature selection (as in Maciel-Guerra *et al.* [2]). The performance before and after WBS is compared. Results show that the feature selection method promotes an increase in the overall performance in most of the methods. Before feature selection all methods showed similar performances. Nonetheless, after feature selection, DS methods did not have an increase in performance such as the other 9 single and ensemble classifiers.

This paper is organized as follows. Section 2 provides background and depicts related work associated with machine learning methods and feature selection when applied to microarrays. Section 3 introduces the methodology. Experimental results using the IBH data set are shown in Section 4. Section 5 outlines the concluding remarks and contributions of this paper, followed by perspectives of future development.

II. Background

Over the past few years, significant work has been carried out to help the analysis of protein microarrays and to assist in the diagnosis of different allergies, resulting in biomarker discoveries. In this section, we review the literature regarding classification and feature selection strategies for protein microarrays. We start by introducing the existing work on the use of machine learning and feature selection to handle microarray data. Subsequently, we introduce the overall concepts of DS methods and discuss its potential to also be employed in microarray data.

A. Related Work on Microarray Classification

In 2015, Marti *et al.* [1], studied the influence of allergen-specific IgE against insect bites in horse sera to explore the causes of IBH. The authors demonstrated that their microarray approach produces high differentiation between healthy animals and those horses with IBH and that the data could be sucessfully classified using Partial Least Square Discriminant Analysis (PLSDA) [1], [15]. To further improve the initial classification and to detect the most important features influencing IBH, Marti *et al.* [1] used the variable importance in projection (VIP) scores, which resulted in 31 features (a reduction of 84%). Results before feature selection achieved specificity of 0.733 and sensitivity of 0.867. After feature selection, sensitivity increased to 1.0 and specificity reached 0.967. The authors state that these results are in agreement with clinical knowledge [1]. In our work we employ the same dataset, which is further described in Section III-A.

In 2018, Maciel-Guerra *et al.* [2], also studied the same dataset used by Marti *et al.* [1]. The authors compared the results found by Marti *et al.* [1] with 9 traditional single and ensemble classifiers. In addition, they proposed a novel non-linear and scalable approach wrapper backward selection with a regularized extreme learning machine as the embedded model. The results after feature selection showed that all classifiers had an increase in performance; and AdaBoost and LDA had sensitivity and specificity results similar to the ones obtained by Marti *et al.* [1]. The proposed feature selection approach had the following advantages: (1) it produces high classification accuracy with a fast learning process; and (2) it performs automatic feature selection based on non-linear embedded models with regularization.

Lin *et al.* [16] and Prosperi *et al.* [17] studied protein microarrays to analyze peanut allergy and asthma, respectively. Lin *et al.* [16] used Decision Trees and Support Vector Machines (SVM) [16], while Prosperi *et al.* [17] applied Logistic Regression, Random Forest and Bayesian Networks coupled with feature selection approaches to classify their data. Their results suggested that machine learning methods produce satisfactory results when dealing with protein microarray data. Lin *et al.* [16] obtained an overall accuracy above 90%, while Prosperi *et al.* [17] achieved an AUC in the range of 0.76 to 0.82.

B. Microarray Feature Selection

Feature selection methods are employed in microarrays to remove irrelevant features that might lead to loss of biological information and poor classification performance [18], [19]. Jirapech-Umpai and Aitken [20] and Kumar *et al.* [19] studied the effects of different feature selection (gene ranking and t-test) approaches in microarray data analysis. Jovic *et al.* [18] showed that different studies with microarray data sets have used feature selection methods to make biomarker discovery and to reduce the size of the feature set to improve classification [18].

Maldonado and Weber [4] implemented a wrapper backward selection (WBS) using a Support Vector Machine with kernel functions as a learning model for 4 different datasets (Winsconsin Breast Cancer, Colorectal Microarray data set and two credit scoring datasets.). They showed that WBS outperforms filters and other wrapper methods for their data. Nevertheless, they also stated that this method can be computationally expensive if the number of features is large [4].

Since WBS invariably requires a learning model trained multiple times, Maciel-Guerra *et al.* [2] proposed a scalable non-linear regularized extreme learning machine to be used as the learning model. By using a RELM, the learning process can be done faster with high classification accuracy. The authors compared the results before and after feature selection for the IBH dataset investigated by Marti *et al.* [1]. We employed this feature selection method in our paper because

it was able to find the most import allergen concerning the IBH dataset and had high AUC results for this dataset.

C. Dynamic Selection

Multiple Classifier Systems (MCS) are a very active area of research in machine learning and pattern recognition. Recently, several studies published results demonstrating its advantages over a single robust classifier [21]–[23]. The idea behind MCS relies on the fact that a combination of "different" classifiers might have a strong degree of "independence" in the errors, *i.e.* make few coincident errors. Thus, the errors committed by a classifier c_i can be overridden by the correct classification of other classifiers. MCS are essentially composed of three major stages: (1) pool generation, (2) selection and (3) integration. In the pool generation stage, the main goal is to train a set of classifiers that are both accurate and diverse, *i.e.* the classifiers must have a low error rate (accurate) and two classifiers must make different errors on new samples (diverse). On the second stage, based on the pool of classifiers, the goal is to select a single or an ensemble of classifiers from the pool of classifiers. This stage can be divided into two groups: static and dynamic selection. In the first group, the classifiers are selected during the training stage and are fixed for all the unknown test samples, while the later selects a different set of classifiers for each test sample. The final stage consists of combining the outputs of the selected ensemble of classifiers according to a combination rule, i.e., majority voting, threshold, random and probability [21], [23], [24]. The basic framework of a multiple classifier system is presented on Figure 1 and the most relevant techniques used in each stage is presented on Figure 2.

Dynamic Selection (DS) is a promising approach to MCS [21], [23]. Recently, several studies have been reporting DS superior performance when compared to a single robust classifier and traditional combination methods (majority voting, bagging, boosting) [21], [23]–[27]. Differently from static classifier ensemble, the selection on dynamic classifier ensemble is done online, during the generalization stage [21], [24], [28], [29]. In other words, a single (dynamic classifier selection) or an ensemble (dynamic ensemble selection) is selected specifically to classify each new test sample. As a rule, the dynamic classifier ensemble has three basic steps (Figure 3). First, the system needs to generate the local region of competence based on the training set on an independent DSEL (dynamic selection) set. Second, the system will use a selection criterion to dynamically determine the competence level of the classifiers. The selection is called dynamic because these two steps are performed during the testing phase. However, there are some approaches in dynamic classifier ensemble that predetermine the region of competence during the training

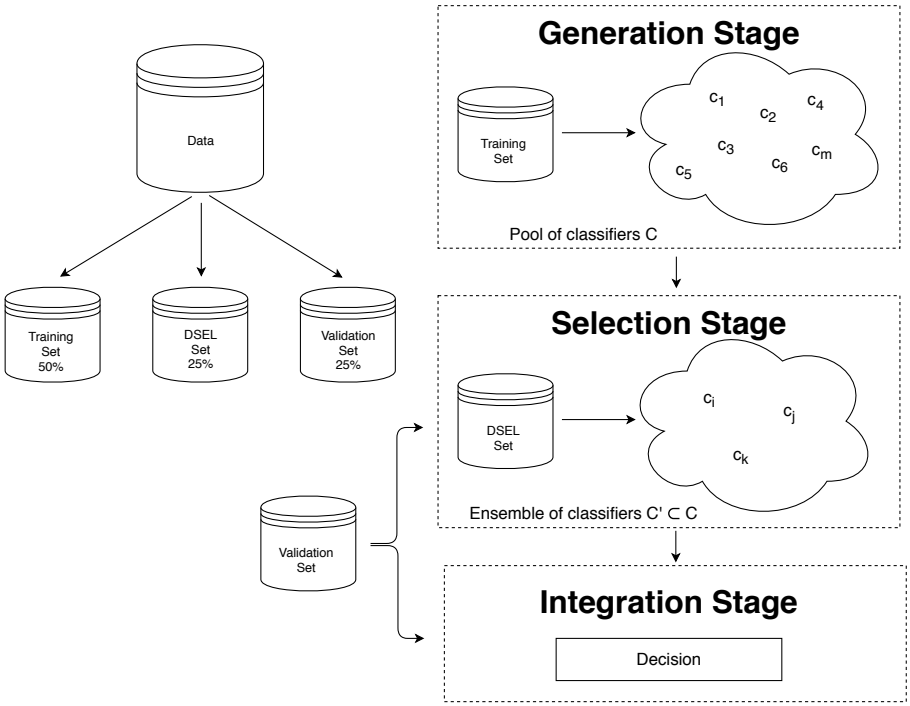

Fig. 1. Basic DS framework showing the three stages: generation, selection and integration

phase and only select the best classifiers during the testing phase [21], [24], [29].

DS techniques can select either a single classifier (*Dynamic Classifier Selection* (DCS)) or an ensemble of classifiers (*Dynamic Ensemble Selection* (DES)). These techniques are used to select classifiers based on their competence level to predict the label of a test sample. The competence is estimated considering only the samples of a local region of the feature space where the test sample is located (region of competence). The majority of DS techniques relies on k-Nearest Neighbours (k-NN) algorithms and the quality of the neighbourhood can have a huge impact on the performance of DS methods [21], [23], [24].

One of the main differences between the DS methods available in the literature is the selection criteria. This step is one of the most important tasks of ensemble of classifiers. The classifiers are selected based on their competence in the whole or a local region of the feature space (region of competence). The basic source of information is related to the classifier accuracy [30]–[33]. However, other metrics were created: ranking [30], [34], probabilistic [25], [35]–[38], oracle [26], behaviour [39], [40], meta-learning [27], complexity [41] and diversity [32], [33].

To the best of our knowledge, DS has not yet been employed to microarray data problems. We believe this could potentially be a useful tool for those datasets due to their low number of

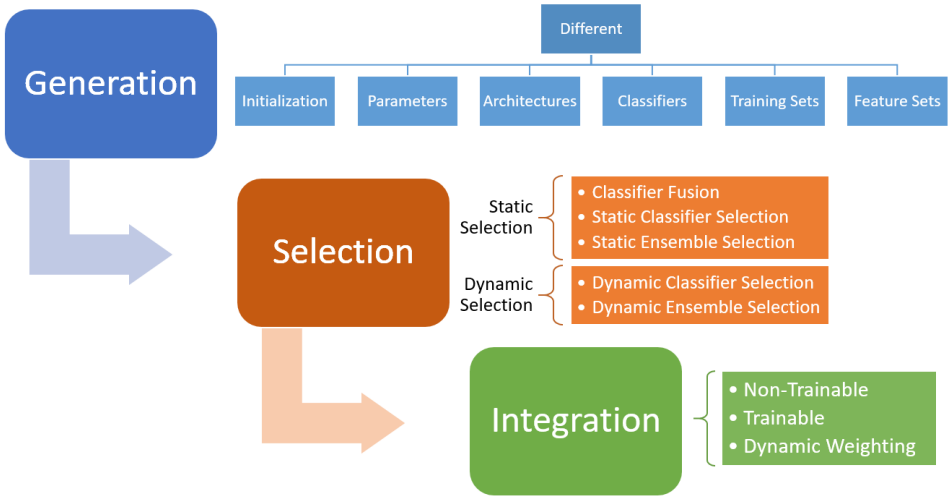

Fig. 2. Three stages of a multiple classifier system. Generation in blue, Selection in orange and Integration in green (adapted from [21]).

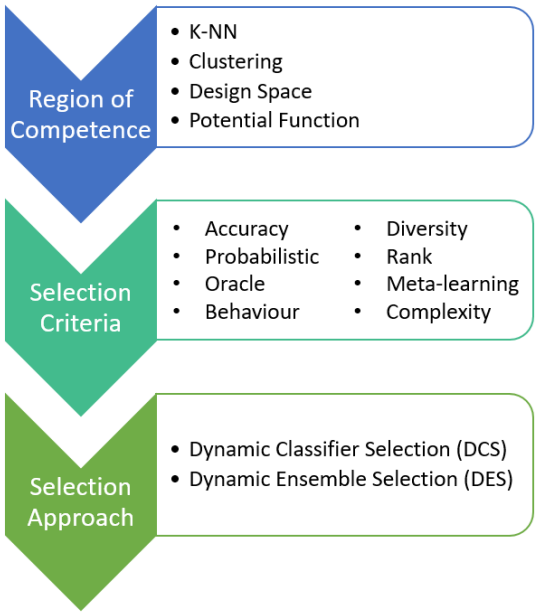

Fig. 3. Main approaches considering the selection stage of DS methods (adapted from [21]).

samples. Therefore, to investigate this potential we employed 21 DS methods to the IBH dataset problem and compared them with traditional single and ensemble techniques.

III. Methodology

In this section, we present further details of the dataset investigated, the methodology applied for pre-processing and analysis, and the experimental design.

A. The Insect Bite Hypersensitivity Dataset

The same dataset employed by Marti *et al.* [1] to study equine insect bite hypersensitivity (IBH) is adopted in this study. The data set contains 109 observations (66 healthy controls and 43 IBH diseased animals) described by 193 features. The minimum value of this data set is 0 and the maximum value is 874.91. The data set does not contain missing values. The data set is pre-processed according to the scheme adopted by Marti *et al.* [1], in which the negative control microarray data (consisting of all reagents except the animal serum) was subtracted from the sample slide to eliminate non-specific binding and inherent autofluorescence of some proteins; after the occurrence of this process, the slides received a second normalization (Equation 1, where n_x is the norm of a 1 by *N* vector **x**), involving the sum of absolute values of all expressions (associated with each individual), in order to reduce technical variability. Finally, the data is mean centered and scaled to unit standard deviation for each feature.

$$n_x = \sum_{i=1}^{N} |x_i| \tag{1}$$

B. Experimental Design

For evaluation of the results we employ classification area under the receiver operating characteristic curve (*AUC*) of each classifier.

The performance of the classifiers Nave Bayes [10], Linear and Non-Linear (RBF kernel) Support Vector Machines (SVM) [4], [5], Random Forest [6], Multi-Layer Perceptron Neural Networks (MLP) [7], [8], AdaBoost [9], Logistic Regression and Linear Discriminant Analysis (LDA) [11], [12] was investigated using the *scikit-learn* library in Python [42]. The Regularized Extreme Learning Machine classifier was implemented in MATLAB R2016a (The MathWorks, Inc., Natick, Massachusetts, United States), using the proposal of Kulaif and Von Zuben [43]. A computer with processor Intel(R) Core(TM) i3-2310M, master frequency 2.10GHz and RAM memory 8Gb was used to run the experiments.

For the classifiers, the following set of values are employed for the hyper-parameters, before and after feature selection [2]:

- **Logistic Regression:** inverse of regularization strength *C* = [0.001, 0.01, 0.1, 1, 10, 100, 1000].
- **Linear SVM:** penalty parameter of the hinge loss error *C* = [0.001, 0.01, 0.1, 1, 10, 100, 1000].
- **Random Forest and Adaboost:** Number of estimators = [2, 4, 8, 16, 32, 64, 128, 256, 512, 1024].
- **MLP Neural Network:** α (L2 penalty parameter) = [0.001, 0.01, 0.1, 1, 10, 100], learning rate (initial learning rate used to control the step size in updating the weights with adam solver) = [0.001, 0.01, 0.1, 1] and hidden layer sizes = [10, 20, 40, 100, 200, 300, 400, 500].
- **Non-linear SVM with RBF kernel:** γ (RBF kernel coefficient) = [0.0001, 0.001, 0.01, 0.1] and *C* (L2 penalty parameter) = [0.001, 0.01, 0.1, 1, 10, 100, 1000].
- **RELM:** λ (L2 penalty parameter) = [0.0001, 0.001, 0.01, 0.1, 0, 1, 10, 100, 1000], hidden layer sizes = [200, 201, 202, ..., 698, 699, 700], and random weights of the hidden layer in the range [-0.5, 0.5].
- **Naive Bayes and LDA:** do not have hyper-parameters.

For the WBS approach, an RELM with λ (L2 penalty parameter) = [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000] and hidden layer size of 500 was adopted [2].

Table I shows the different DS methods found in the literature and used in this paper. More information about each one can be found on their respective reference or on the recent reviews done by [21], [24]

IV. Results

Table II shows the AUC results for all techniques mentioned on Section III-B, before and after feature selection. Each experiment is conducted thirty times. The numbers after the "$\pm$" symbol are standard deviation.

According to the results produced in [2], which were also replicated in this work, the WBS with RELM as the embedded model produced 36 features which are considered the most relevant. The third column of Table II shows that the feature selection is able to remove redundant and non-important proteins for the classification of IBH for most classifiers, since their performance is improved after feature selection or remains statistically the same. After feature selection, the AUC was close to one for most of the single classifiers. However, the DS methods did not produce a statistically significant increase in their performance, maintaining the accuracy within one standard deviation from the average before feature selection.

Comparing the AUC before and after feature selection, LDA has the highest increase (31.36%), followed by Naive Bayes and RELM with an increase of 21.11% and 18.34%, respectively. Random Forest and Adaboost are the only two classifiers with an AUC over 0.9 before feature selection; however, they produce the lowest performance increase after WBS is applied, i.e., 5.99% and 0.05%, respectively. Linear SVM, Logistic Regression, MLP and SVM with RBF kernel have improvements in the range of 11% to 15%.

From the obtained results it is also relevant to observe that RELM is an effective classifier for the IBH allergy dataset and the wrapper backward selection is an effective way of extracting the most important features. Random Forest and Adaboost performed better before selection and all 9 static classifiers performed better after feature selection. Although the literature shows that DS methods usually have a higher performance when compared with single classifiers or other

ensemble techniques, for this specific dataset DS underperformed.

These results can potentially be explained by the fact that the IBH microarray data has far more features than samples. As the DS methods employ a bagging algorithm to generate the pool of classifiers in different regions of the feature space, fewer instances available do not allow the methods to create effective regions of competence for classification.

We were unable therefore to demonstrate the potential benefits of employing this type of classification for our microarray data. In addition, for those classifiers, the feature selection stage also did not produce any significant improvement. On the other hand, Table II shows that in most of the times dynamic ensemble selection methods perform better than dynamic classifier selection methods, which is in accordance with the literature. This is because selecting an ensemble of classifiers rather than a single one tends to produce less classification errors. It is necessary in the future to test the methodology in other microarray data sets to evaluate whether DS is not suitable in general or if there are cases and specific datasets where the approach could produce better classification results.

V. Conclusions and Future Work

Accurate diagnosis of a disease is vital for a successful therapy. Protein microarrays are a powerful tool employed in allergy diagnostics in order to monitor interactions of antibodies with allergens. In this paper, we investigated the potential of DS methods to microarray data. We used an insect bite hypersensitivity dataset as our case study and compared the DS results with traditional machine learning methods. We also compared the results of DS with single and static classifiers before and after feature selection.

Machine learning classifiers along with WBE for feature selection were investigated. A Regularized Extreme Learning Machine with WBS was used as suggested by Maciel-Guerra [2]. We compared the classification results before and after WBS. The DS methods did not have a higher increase in performance. In addition, most of the outputs of the 21 different DS models produced statistically similar results.

Most of single and static classifiers however had a higher increase in performance. These results may be explained by the fact that this dataset has more features than samples, and the DS methods use a bagging algorithm to generate the pool of classifiers in different regions of the feature space. With few samples in a high dimensional space, the k-Nearest Neighbour algorithm may not find the correct regions of competence for each new test sample.

Future work will be conducted by verifying if similar results occur with other high dimensional datasets and how it is possible to improve the way DS methods find the regions of competence. In addition, we intend to investigate different methods of feature selection that can select adequate features taking into consideration the way DS regions of competence are formed.

References

[1] E. Marti, X. Wang, N. Jambari, C. Rhyner, J. Olzhausen, J. J. Prez-Barea, G. P. Figueredo, and M. J. C. Alcocer, "Novel in vitro diagnosis of equine allergies using a protein array and mathematical modelling

TABLE I
DS methods investigated for the IBH microarray data

Name	Selection criteria	DS Method	Reference	Year
Classifier Rank (CR)	Ranking	DCS	[34]	1993
Modified Classifier Rank (MCR)	Ranking	DCS	[30]	1997
Overall Local Accuracy (OLA)	Accuracy	DCS	[30]	1997
Local Class Accuracy (LCA)	Accuracy	DCS	[30]	1997
A Priori	Probabilistic	DCS	[35]	1999
A Posteriori	Probabilistic	DCS	[35]	1999
Multiple Classifier Behaviour (MCB)	Behaviour	DCS	[39]	2002
Modified Local Accuracy (MLA)	Accuracy	DCS	[31]	2002
DES - k-Means (DES-kMeans)	Accuracy & Diversity	DES	[32], [33]	2006
DES - k-Nearest Neighbour (DES-kNN)	Accuracy & Diveristy	DES	[32], [33]	2006
KNORA - Eliminate (KNORA-E)	Oracle	DES	[26]	2008
KNORA - Union (KNORA-U)	Oracle	DES	[26]	2008
DES - Exponential (DES-EXP)	Probabilistic	DES	[36]	2009
DES - Randomised Reference Classifier (DES-RRC)	Probabilistic	DES	[25]	2011
DES - Minimal Difference (DES-MD)	Probabilistic	DES	[38]	2011
DES - Kullback-Leibler Divergence (DES-KL)	Probabilistic	DES	[37]	2012
DES - Performance (DES-P)	Probabilistic	DES	[37]	2012
KNOP - Eliminate (KNOP-E)	Behaviour	DES	[40]	2013
KNOP - Union (KNOP-U)	Behaviour	DES	[40]	2013
Meta-Learning - DES (Meta-DES)	Meta-learning	DES	[27]	2015
Dynamic Selection on Complexity (DSOC)	Accuracy & Complexity	DCS	[41]	2016

TABLE II
AUC RESULTS IN INSECT BITE HYPERSENSITIVITY DATA SET

Classifiers	Before Feature Selection	After Feature Selection
Logistic Regression	0.8561 ± 0.0245	0.9760 ± 0.0096
Linear SVM	0.8455 ± 0.0308	0.9702 ± 0.0118
SVM - RBF kernel	0.8698 ± 0.0228	0.9709 ± 0.0097
Random Forest	0.9147 ± 0.0230	0.9695 ± 0.0119
MLP	0.8620 ± 0.0241	0.9738 ± 0.0115
AdaBoost	$\mathbf{0.9339 \pm 0.0265}$	0.9344 ± 0.0279
Naive Bayes	0.7345 ± 0.0215	0.8896 ± 0.0236
LDA	0.7565 ± 0.0383	$\mathbf{0.9937 \pm 0.0051}$
RELM	0.8346 ± 0.1097	0.9877 ± 0.0235
CR	0.8155 ± 0.0650	0.7733 ± 0.0588
MCR	0.7733 ± 0.0607	0.7858 ± 0.0633
OLA	0.8324 ± 0.0618	0.8342 ± 0.0513
LCA	0.7725 ± 0.0520	0.7930 ± 0.0615
A Priori	0.8163 ± 0.0692	0.8072 ± 0.0515
A Posteriori	0.7447 ± 0.0608	0.7789 ± 0.0853
MCB	0.8217 ± 0.0455	0.8377 ± 0.0266
MLA	0.7947 ± 0.0796	0.7976 ± 0.0367
DES-kMeans	0.8444 ± 0.0755	0.8134 ± 0.0543
DES-kNN	0.8471 ± 0.0678	0.8594 ± 0.0486
KNORA-E	0.8198 ± 0.0794	0.8064 ± 0.0629
KNORA-U	0.8765 ± 0.0585	0.8500 ± 0.0348
DES-EXP	0.8040 ± 0.0644	0.8000 ± 0.0531
DES-RRC	0.8642 ± 0.0567	0.8406 ± 0.0619
DES-MD	0.8452 ± 0.0562	0.8037 ± 0.0678
DES-KL	0.8345 ± 0.0939	0.8441 ± 0.0470
DES-P	0.8703 ± 0.0639	0.8422 ± 0.0583
KNOP-E	0.8618 ± 0.0577	0.8471 ± 0.0573
KNOP-U	0.8070 ± 0.0606	0.8324 ± 0.0422
Meta-DES	0.8618 ± 0.0577	0.8471 ± 0.0573
DSOC	0.8781 ± 0.0580	0.8409 ± 0.0567

approach: a proof concept using insect bite hypersensitivity," *Veterinary Immunology and Immunopathology*, vol. 167, pp. 171 – 177, 2015.

[2] A. Maciel-Guerra, F. J. Von Zuben, R. Veroneze, E. Marti, J. M. Garibaldi, G. P. Figuredo, and M. J. C. Alcocer, "Microarray feature selection and classification for early detection of allergies in horses," paper submited for publication.

[3] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 20, no. 2, pp. 215 – 242, 1958.

[4] S. Maldonado and R. Weber, "A wrapper method for feature selection using support vector machines," *Information Sciences*, vol. 179, p. 22082217, 2009.

[5] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273 – 297, 1995.

[6] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5 – 32, 2001.

[7] J. Moody and C. J. Darken, "Fast learning in networks of locally tuned processing units," *Neural Computing*, vol. 1, pp. 281 – 294, 1989.

[8] F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington DC, 1961.

[9] Y. Freund and R. Schapire, "A decision theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997.

[10] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann Publishers, 2011.

[11] G. J. McLachlan, *Discriminant analysis and statistical pattern recognition*. Wiley, New York, 2004.

[12] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning*. Springer, New York, 2003.

[13] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489 – 501, 2006.

[14] F. Benoit, M. van Heeswijk, Y. Miche, M. Verleysen, and A. Lendasse, "Feature selection for nonlinear models with extreme learning machines," *Neurocomputing*, vol. 102, pp. 111 – 124, 2013.

[15] S. Wold, M. Sjostrom, and L. Eriksson, "Pls-regression: a basic tool of chemometrics," *Chemometrics and Intelligent Laboratory systems*, vol. 58, pp. 109 – 130, 2001.

[16] J. Lin, F. M. Bruni, Z. Fu, J. Maloney, L. Bardina, A. L. Boner, G. Gimenez, and H. A. Sampson, "A bioinformatic approach to identify patients with symptomatic peanut allergy using peptide microarray immunoassay," *J Allergy Clin Immunol*, vol. 129, no. 5, p. 13211328.e5, 2012.

[17] M. C. F. Prosperi, D. Belgrave, I. Buchan, A. Simpson, and A. Custovic, "Challenges in interpreting allergen microarrays in relation to clinical symptoms: A machine learning approach," *Pediatr Allergy Immunol*, vol. 25, pp. 71 – 79, 2014.

[18] A. Jovic, K. Brkic, and N. Bogunovic, "A review of feature selection methods with applications," *In Information and Communication Technology, Electronics and Microelectronics (MIPRO). 2015, 38th International Convention on,*, pp. 1200 – 1205, 2015.

[19] M. Kumar, S. Singh, and S. K. Rath, "Classification of microarray data using extreme learning machine classifier," *Internation Jounal of Information Processing*, vol. 9, no. 3, pp. 1 – 16, 2015.

[20] T. Jirapech-Umpai and S. Aitken, "Feature selection and classifcation for microarray data analysis: evolutionary methods for identifying predictive genes," *BMC Bioinformatics*, vol. 6, p. 148, 2005.

[21] R. M. Cruz, R. Sabourin, and G. D. Cavalcanti, "Dynamic classifier selection: Recent advances and perspectives," *Information Fusion*, vol. 41, no. Supplement C, pp. 195 – 216, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253517304074

[22] L. I. Kuncheva, "Switching between selection and fusion in combining classifiers: an experiment," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 32, no. 2, pp. 146–156, Apr 2002.

[23] R. M. Cruz, H. H. Zakane, R. Sabourin, and G. D. Cavalcanti, "Dynamic ensemble selection vs k-nn: why and when dynamic selection obtains higher classification performance?" in *The Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Montreal, Canada, 2017.

[24] A. S. Britto, Jr., R. Sabourin, and L. E. S. Oliveira, "Dynamic

selection of classifiers - a comprehensive review," *Pattern Recogn.*, vol. 47, no. 11, pp. 3665–3680, Nov. 2014. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2014.05.003

[25] T. Woloszynski and M. Kurzynski, "A probabilistic model of classifier competence for dynamic ensemble selection," *Pattern Recogn.*, vol. 44, no. 10-11, pp. 2656–2668, Oct. 2011. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2011.03.020

[26] A. H. R. Ko, R. Sabourin, and A. S. Britto, Jr., "From dynamic classifier selection to dynamic ensemble selection," *Pattern Recogn.*, vol. 41, no. 5, pp. 1718–1731, May 2008. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2007.10.015

[27] R. M. Cruz, R. Sabourin, G. D. Cavalcanti, and T. I. Ren, "Meta-des: A dynamic ensemble selection framework using meta-learning," *Pattern Recognition*, vol. 48, no. 5, pp. 1925 – 1935, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320314004919

[28] D. Ruta and B. Gabrys, "Classifier selection for majority voting," *Information Fusion*, vol. 6, no. 1, pp. 63 – 81, 2005, diversity in Multiple Classifier Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253504000417

[29] J. Xiao, C. He, X. Jiang, and D. Liu, "A dynamic classifier ensemble selection approach for noise data," *Inf. Sci.*, vol. 180, no. 18, pp. 3402–3421, Sep. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.ins.2010.05.021

[30] K. Woods, W. P. Kegelmeyer, and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 405–410, Apr 1997.

[31] P. C. Smits, "Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 4, pp. 801–813, Apr 2002.

[32] R. G. F. Soares, A. Santana, A. M. P. Canuto, and M. C. P. de Souto, "Using accuracy and diversity to select classifiers to build ensembles," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, 2006, pp. 1310–1316.

[33] M. C. P. de Souto, R. G. F. Soares, A. Santana, and A. M. P. Canuto, "Empirical comparison of dynamic classifier selection methods based on diversity and accuracy for building ensembles," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, June 2008, pp. 1480–1487.

[34] M. Sabourin, A. Mitiche, D. Thomas, and G. Nagy, "Classifier combination for hand-printed digit recognition," in *Document Analysis and Recognition, 1993., Proceedings of the Second International Conference on*, Oct 1993, pp. 163–166.

[35] G. Giacinto and F. Roli, "Methods for dynamic classifier selection," in *Proceedings 10th International Conference on Image Analysis and Processing*, 1999, pp. 659–664.

[36] T. Woloszynski and M. Kurzynski, "On a new measure of classifier competence applied to the design of multiclassifier systems," in *Proceedings of the 15th International Conference on Image Analysis and Processing*, ser. ICIAP '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 995–1004. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-04146-4$_1$06

[37] T. Woloszynski, M. Kurzynski, P. Podsiadlo, and G. W. Stachowiak, "A measure of competence based on random classification for dynamic ensemble selection," *Information Fusion*, vol. 13, no. 3, pp. 207 – 213, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253511000297

[38] B. Antosik and M. Kurzynski, "New measures of classifier competence - heuristics and application to the design of multiple classifier systems," in *Computer Recognition Systems 4*, R. Burduk, M. Kurzyński, M. Woźniak, and A. Żołnierek, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 197–206.

[39] G. Giacinto and F. Roli, "Dynamic classifier selection based on multiple classifier behaviour," *Pattern Recognition*, vol. 34, 11 2002.

[40] P. R. Cavalin, R. Sabourin, and C. Y. Suen, "Dynamic selection approaches for multiple classifier systems," *Neural Computing and Applications*, vol. 22, no. 3, pp. 673–688, Mar 2013. [Online]. Available: https://doi.org/10.1007/s00521-011-0737-9

[41] A. L. Brun, A. S. Britto, L. S. Oliveira, F. Enembreck, and R. Sabourin, "Contribution of data complexity features on dynamic classifier selection," in *2016 International Joint Conference on Neural Networks (IJCNN)*, July 2016, pp. 4396–4403.

[42] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825 – 2830, 2011.

[43] A. C. P. Kulaif and F. J. V. Zuben, "Improved regularization in extreme learning machines," *In: 11th Brazilian Congress on Computational Intelligence, 2013, Ipojuca PE. Proceedings of the 11th Brazilian Congress on Computational Intelligence*, vol. 1, pp. 1 – 6, 2013.