Proceedings of the 2013 Federated Conference on
Computer Science and Information Systems pp. 137–141

An Emotional Learning-inspired Ensemble Classifier (ELiEC)

Mahboobeh Pasrapoor, Urban Bilstrup
School of Information Science, Computer and Electrical Engineering (IDE), Halmstad University

Abstract— In this paper, we suggest an inspired architecture by brain emotional processing for classification applications. The architecture is a type of ensemble classifier and is referred to as 'emotional learning-inspired ensemble classifier' (ELiEC). In this paper, we suggest the weighted k-nearest neighbor classifier as the basic classifier of ELiEC. We evaluate the ELiEC's performance by classifying some benchmark datasets.

I. Introduction

Classification methods have been widely used in the area of science, engineering, industry, business and medicine; they can be used for classification problems e.g., anomaly detection, handwriting recognition, speech recognition and medical diagnosis. Among them, the data driven classification approaches e.g., neural network-based models and neuro-fuzzy-based methods are the most popular methods due to the self-adaptive and high generalization capability. However, they have some significant issues: over fitting, model complexity and the curse of dimensionality, etc. [1]-[5]. Thus, developing new classification models to increase the classification's accuracy while resolving the mentioned issues are an open research topic in data mining. In this paper, a new classification model is suggested that can be considered as an ensemble classification with a different integration mechanism and combination algorithm. The model is an emotionally inspired model and is named 'brain emotional learning-inspired ensemble classifier' (ELiEC).

The rest of this paper is organized as follows: Section II reviews some works in classification and emotional learning–based models. Section III explains the ELiEC's structure. In Section IV, the benchmarks classification data sets are examined by ELiEC and the obtained results compared with other methods. Finally in Section V, we conclude and recommendsome possible future improvements to ElIEC.

II. A BRIEF REVIEW

A. Related works to Classification methods

Numerous artificial intelligence-based methods have been proposed for classification problems. They can be categorized as: inductive or transductive, statistical-based or non-statistical–based, supervised or unsupervised methods. One popular group is supervised classification methods that include statistical methods (e.g., Naïve Bayes), non-statistical methods (e.g., neural network), instance based learning and support vector machine. Given a set of instances, these algorithms can assign an appropriate label to an unlabeled instance. Among the supervised learning methods, the support vector machine has the best performance in terms of classification accuracy; however, it has high time complexity that is a big issue for online classification applications [2] and [5].

Numerous efforts have been put into developing regularization methods to increase the generalization of supervised classification algorithms and reduce the time complexity of the learning procedure. A good example of the enhanced classification methods is the NFI model (Neuro-Fuzzy Inference Method for Transductive Reasoning) that provides a local model for each instance using a transductive reasoning system. The NFI model outperforms the neural network model in terms of accuracy and time complexity; nevertheless this model is not suitable for high dimensional classification applications [6]. Developing an ensemble-based [4] classifier is a major progress in addressing the misclassification and time complexity issues. The idea of ensemble-based classifiers is inspired by the human decision making process. The main components of an ensemble method are diversity generator and combiner. The former selects appropriate classifiers while the latter combines the classifiers' outputs. The combination mechanisms that have been developed can be divided into two subgroups: meta-combination and weighting methods. Choosing suitable diversity classifiers and combination procedures, the classification accuracy of ensemble-based classifiers outperforms the accuracy of each of the classifiers; however there is no adaptive procedure for choosing the classifiers and the combiner using information of classification problems.

Recently, a novel classification framework with the ability to adaptively tune the classifier's structure has been developed. This model, called meta-cognitive neural network (McNN), encompasses two components: a cognitive component and a meta-cognitive component [1],[5]. In McNN, the first component is a Radial Basis Neural Network and it is responsible for the change and optimization of the structure. The second component plays a role in choosing samples and the effective structure of the learning algorithm, obtaining knowledge from the training data. The McNN model and an extended version of that PBL-McRBFN [5] have shown excellent performance for classification problems.

The ELiEC model is a general purpose classification method that aims to reduce the misclassification rate and time complexity in classification applications. The ELiEC model can be categorized as a type of ensemble classifiers that uses Wk-NN as the classifiers.

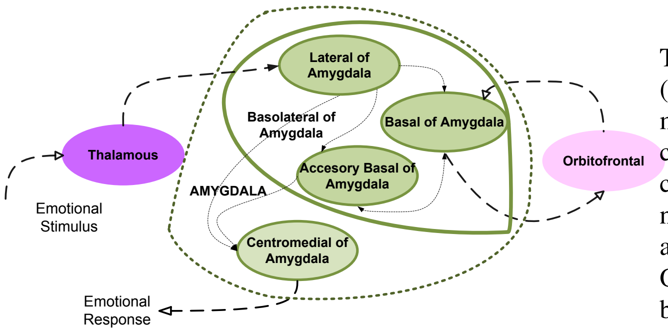

Figure1. The components of the amygdala and their connections.

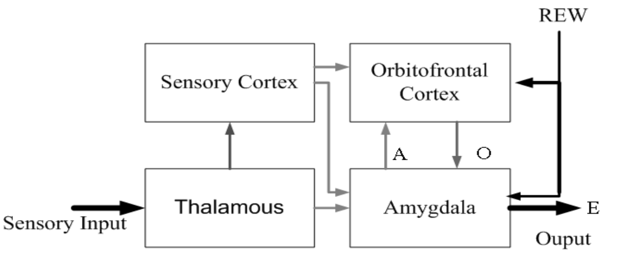

Fig. 2.The graphical description of amygdala-orbitofrontal subsystems

B. Brain Emotional Learning-based Models

Modeling the brain emotional processing and the memory to develop the artificial intelligence (AI-based) based tools has been an interesting topic in the machine learning research area. The emotionally-inspired AI tools are often referred to as the computational models of emotional learning. A well-known computational model is the amygdala-orbitofrontal system that defines emotional processing on the basis of the limbic system. The assumption of this model is that the limbic system, which consists of thalamus, sensory cortex, amygdala and orbitofrontal cortex, is mainly responsible for emotional learning in mammalian brains [8]. Fig. 1 depicts the connection between amygdala and its components. Fig. 2 describes the amygdala-orbitofrontal subsystem model that consists of four parts that interact with each other to mimic the emotional learning processing [8]. Due to its simple structure, the fundamental framework has been used for developing AI tools for control applications and and nonlinear system prediction. [12]-[17].

One popular emotionally inspired controller is BELBIC[9] that has been proven to overcome uncertainty and complexity issues of other intelligent controllers. Studies [9]-[11] have proved that BELBIC outperforms many other controllers such as PID controller and linear controllers in terms of simplicity, reliability and stability.

The emotionally based prediction models have often been applied for chaotic time series prediction and have also shown improvements in prediction accuracy [12]-[17].

III. Emotional Learning-Inspired Ensemble Classifier (ELiEC)

The brain emotional learning-based ensemble classifier (ELiEC) model has a similar architecture to our two previous modelsBELRFS and BELFIS [13], [15]. The ELiEC model consists of four main parts and imitates the internal connection of the emotional system. The parts of ELiEC are named as : TH, CX, AMYG and ORBI that **are** referred to as THalamous, sensory CorteX, AMYGdala and ORBItofrontal cortex**.** The ELiEC model and the connection between these parts are described in Fig. 3.

For a classification problem, we define the set of training data as $(\mathbf{x}_1, c^1), \ldots, (\mathbf{x}_i, c^i) \ldots, (\mathbf{x}_n, c^n)$, where $\mathbf{x}_i$ is an instance with m features and $\mathbf{c}^i$ determines the label class of $\mathbf{x}_i$. In a multi-class classification problem, we have n classes and the corresponding class of $\mathbf{x}_i$ which can be encoded as $\mathbf{y}^i = y_1^i, \ldots, y_2^i, \ldots, y_n^i$. If $\mathbf{c}^i$ is equal with j^{th} class, the value of y_j^i will be equal to one and other values will be zero. Using the following steps we explain how ELiEC classifies each instance in order to minimize the misclassification. The TH part evaluates the features of $\mathbf{x}_i$ and adds several extra features to $\mathbf{x}_i$. The extra features are calculated according to equation (1).

$$\mathbf{th}_i = [\max(\mathbf{x}_i), \mathrm{mean}(\mathbf{x}_i), \min(\mathbf{x}_i)] \qquad (1)$$

The CX evaluates the features of $\mathbf{x}_i$ and eliminates redundant features. The CX has a role to select the most informative features and eliminate the redundant features. Thus, the CX receives $\mathbf{x}_i$ with m features and provides $\mathbf{s}_i$ with l features $(l \leq m)$.

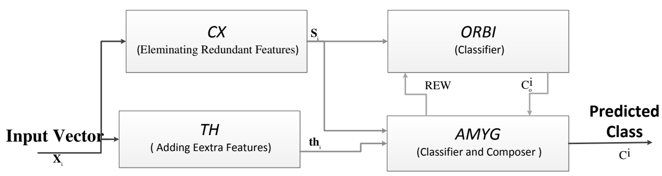

Fig. 3.The architecture of ELiEC.

The AMYG consists of a classifier and a combiner. The classifier that is represented as equation (4) predicts an appropriate class for $\mathbf{x}_i^a$ which is determined as equation (3). The combiner of the AMYG combines the outputs of the AMYG and ORBI to provide the final class as equation (5).

$$\mathbf{x}_i^a = [\mathbf{th}_i, \mathbf{s}_i] \qquad (3)$$

$$C_a^i = \text{Class}(\mathbf{x}_i^a) \quad (4)$$

The combiner strategy depends on the type of classification methods. In this paper, the Wk-NN method has been utilized as the classifiers of AMYG and ORBI. The combiner is another Wk-NN and determines the final class of the input vector $\mathbf{x}_i^c$ which is a vector as $\mathbf{x}_i^c = [\mathbf{x}_i^a, \mathbf{x}_i^o, C_a^i, C_o^i]$.

$$C^i = \text{Class}(\mathbf{x}_i^c) \quad (5)$$

The ORBI is another classifier that can be a dependent classifier or independent classifier. For a dependent classifier the ORBI classifies the input vector $\mathbf{x}_i^o = [\mathbf{s}_i, REW_i]$; while for an independent classifier the input vector is $\mathbf{x}_i^o = \mathbf{s}_i$. Finally, it forwards the classification result $C_o^i = \text{Class}(\mathbf{x}_i^o)$ to AMYG. For the examples of this study an independent classifier is assigned to ORBI.

It should be noted that the classifiers of AMYG and ORBI can be defined on the basis of any supervised classification method, e.g., decision tree, single or multilayer perceptron, and support vector machine, etc. We can also form a meta-ensemble classifier by choosing an ensemble-based classifier for the AMYG and the ORBI.

IV. WEIGHTED K- NEAREST NEIGHBOR : A BASIC CLASSIFIER OF ELiEC

Weighted k-nearest neighbor (Wk-NN) is a type of instance-based algorithm that has been widely used as a classification and regression method. For a given training set as: $(\mathbf{x}_1, c^1), ..., (\mathbf{x}_i, c^i)..., (\mathbf{x}_{N_t}, c^{N_t})$, the Wk-NN determines the class of a test vector, $\mathbf{x}_{test}$,using the following steps [18]:

1) The Euclidian distance between $\mathbf{x}_{test}$ and $\mathbf{x}_i$ is calculated, $d_i = \|\mathbf{x}_{test} - \mathbf{x}_i\|_2$, in which each $\mathbf{x}_i$ is a member of $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{N_t}$, where N_t denotes the number of samples of the training data set.

2) The k minimum values of $\mathbf{d} = \{d_1, d_2, ..., d_{N_t}\}$ are selected as $\mathbf{d}_{min}$. The $\mathbf{x}_i$ s that are corresponding to $\mathbf{d}_{min}$ are the k nearest neighbors to $\mathbf{x}_{test}$ and define the local neighborhoods of the test vector, $\mathbf{x}_{test}$ [18].

3) The class label of $\mathbf{x}_{test}$ is chosen from the class labels of the local neighborhoods. Using the weighted k-nearest neighbor (W-kNN), a weight is assigned to each neighbor; the assigned weight is defined according the kernel function $K(.)$.

Any arbitrary function that holds the properties below can be considered as the kernel function [18].

1) For all d , $K(d) \geq 0$.
2) If $d = 0$ then $K(d)$ gets the maximum value.
3) If $d \to \pm\infty$ then $K(d)$ gets the minimum value.

In this paper, the kernel function is defined as (6). Thus, the closer neighbors to $\mathbf{x}_{test}$ have higher weights on estimating the class label of $\mathbf{x}_{test}$.

$$K(d) = \frac{\max(\mathbf{d}) - (d_j - \min(\mathbf{d}))}{\max(\mathbf{d})} \quad (6)$$

V. BENCHMARK DATASETS

In this section, we test the ELiEC model to classify several data sets that have obtained from the University of California, Irvine (UCI) machine learning repository [20]. It should be noted that, each data instance has been normalized and the values of the attributes are scaled to the range [0,1]. For each benchmark, the cross validation set has the same size as the training data set. We test the classification model 100 times; however, we present the best results. To present an efficient comparison, we test the ELiEC model for both multiclass and binary class benchmark data sets. We also consider both balanced and imbalanced classification benchmark data sets. The performance of classification models are calculated using the average per class classification accuracy as (7) which is represented as η_a . The parameter n determines the number of classes and η_j indicates the classification accuracy of j class. The test error determines the number of misclassification samples of the test set; the samples of the test data set is determined by N_{test}

$$\eta_a = \frac{1}{N_{test}} \sum_{j=1}^{n} \eta_j \quad (7)$$

A. Multiclass and Well Balanced Benchmark Data Sets

Iris data set and wine determination are two well-known classification benchmarks that have been categorized under the well balanced and multiclass data sets. Iris data set waobtained from the University of California, Irvine (UCI) machine learning repository [20] and has been examined by several classification methods e.g., NFI, PBL-McRBFN, etc [1],[5],[21],[22]. This data set consists of 150 samples; each sample has 4 attributes named as: sepal length, sepal width, petal length and petal width. Iris data set is a set with multiclassand consists of three classes: iris setosa, iris versicolor and iris virginica [20].

As the first case study, 50 percent of iris data sets are randomly chosen as the training data set; while the rest of the samples are considered as the test data set. Table I compares the number of misclassifications using ELiEC, NFI and MLP. The results indicate that using ELIEC the average number of test errors is equal to 2.2, that is less than the two other methods.

Table I. Comparison of performance classification of BELBEC and other methods for Iris data set with 50% as the training data set and 50% test data set.

Classification Model	Specification of results		
	Structure	Test error	The number of training data
BELBEC	15 neighbor	2.2	75
NFI[6]	6.3 neighbor	3.3	75
MLP[6]	12 neuron	4.6	75

Table II. The classification accuracy of BELBEC for the Iris data set with 45 samples as training data and 105 samples test data.

Classification Model	Specification of results		
	Structure	The average per class accuracy	The number of training data
BELBEC without normalized	16 neighbor	99.02 %	45
McNN[1]	5 neuron	97.14 %	22
PBL-McRBFN[5]	6 neuron	98.10 %	20
BELBEC	18 neighbor	%96.24	60
McNN[1]	9 neuron	98.49 %	27
PBL-McRBFN[5]	11 neuron	98.69 %	29
ensemple of (OC-SVM) [22]	----	92.00%	----

As the second case, the ELiEC classifies iris set using 45 samples as the training data set..

The wine data set is the second data set that is used to evaluate ELiEC. This data set that is a multiclass and a well-balanced data set; it is obtained from the chemical analysis of some wines that are produced from various cultivars in the same region of Italy. The sample of the data set consists of 13 features and can be categorized into three classes. Table II compares the results of ELIEC and three methods: McNN, PBL-McRBFN and another method types of ensemple classifiers; this model was named OC-SVM that is referred to as one-class support vector machines[22]. In this experiment, we use 60 samples as the training data samples and 118 samples as the test data. It is observable that ELiEC has good results in classification of this well-balanced data set

As was mentioned, the ELiEC is based on W-kNN; thus, we compare the results of ELiEC and W-kNN for classifying the wine data set (see Table II). We also compare the ELiEC and W-kNN to classify wine data with two different structures. As Table III shows, the classification accuracy of ELiEC has no noticeable difference when the total number of neighbors of ELiEC's classifiers is only eight neighbors. However, when we define a lower number of neighbors for Wk-NN, its accuracy has a significant decrease.

Table III. Comparison between WKNN and BELBEC for the wine data set.

Classification Model	Specification of results		
	Structure	The average per class accuracy	The number of training data
BeLBeC	18 neighbor	%96.24	60
BeLBeC	8neighbor	%95.27	60
KNN	3 neighbor	%95.40	60
KNN	18neighbor	%91.60	60

B. Binary and Low Dimensional DatabaseLiver Disorders (LD) and PIMA Indian Diabetes are two medical data sets that have been tested by different classification methods. Liver Disorders (LD) is related to a blood test and has 345samples and 6 attributes, respectively. Pima Indians diabetes set is another medical data set and has been provided by the National Institute of Diabetes and Digestive and Kidney Diseases. It is used to diagnostic diabetes. The number of samples and number of attributes are 768 and 8 respectively. Table IV lists the results of ELIEC for the data sets and compares the results with three methods: McNN, PBL-McRBFN and a modification of adapting splitting and selection (AdaSS) [23]. It shows that the results of McNN, PBL-McRBFN and a modification of adapting splitting and selection (AdaSS)are better than the results of ELiEC; however, the ELiEC has a simpler structure than the two other methods.

Table IV. Comparison between different methods to classify the LD and PIMA. The number of training samples is equal to [5]

Classification Model	Specification of results		
	Structure	The average per class accuracy	DATA
BELBEC	48neuron	68.4	LD
McNN[1]	68 neurons	71.60	LD
PBL-McRBFN[5]	87 neurons	72.63	LD
BELBEC	100neuron	70.37	PIMA
McNN[1]	193 neurons	77.31	PIMA
PBL-McRBFN[5]	162 neurons	76.67	PIMA
MAD [23]	7 clusters and 5 classifiers	72.010	PIMA

V. CONCLUSION

This paper presents a new ensemble-based classifier that is inspired by the brain emotional network. The architecture is referred to as ELiEC and utilizes W-kNN as the basic classification method. However, the ELIEC differs from other ensemble methods in the way that the classifiers are fed (see Figure 2).

The performance of ELiEC is evaluated by classifying several benchmark data sets. The results indicate a fairly good performance of ELiEC for classification.

As future works, we replace W-kNN with other learning classification methods e.g., the support vector machine to address the time complexity and the curse of dimensionality issues. In addition, a random forest method can be used to form a meta-ensemble classifier.

REFERENCES

[1] G. S. Babu and S. Suresh, “Meta-cognitive neural network for classification problems in a sequential learning framework,” Neurocomputing, vol. 81, pp. 86–96, Apr. 2012.

[2] S.B. Kotsiantis, I.D. Zaharakis, and P.E. Pintelas, “Machine Learning: A Review of Classification and Combining Techniques,” Artificial Intelligence Review, vol.26, pp. 159-190, 2006.

[3] T.Phyu, “Survey of Classification Techniques in Data Mining”, in.Proc. International Multi Conference of Engineers and Computer Scientists, vol, I IMECS 2009, 2009.

[4] L. Rokach, “Ensemble-based classifiers,” Artificial Intelligence Review, vol. 33, no. 1, pp. 1–39, 2 2010.S.B. Kotsiantis, Supervised Machine Learning: A Review of Classification Techniques, J. Informatica(31),pp. 249-26, 2007.

[5] G. S. Babu and S. Suresh, “Sequential Projection-Based Metacognitive Learning in a Radial Basis Function Network for Classification Problems,” Neural Networks and Learning Systems, IEEE Transactions on , vol.24, no.2, pp.194,206, Feb. 2013

[6] Q.Song; N.K.Kasabov, "NFI: a neuro-fuzzy inference method for transductive reasoning," Fuzzy Systems, IEEE Transactions on , vol.13, no.6, pp.799,808, Dec. 2005

[7] J.Moren, C.Balkenius,“A computational model of emotional learning in the amygdala,’’, in From Animals to Animats, MIT, Cambridge, 2000.

[8] C. Lucas, D. Shahmirzadi, N. Sheikholeslami, “ Introducing BELBIC: brain emotional learning based intelligent controller,’’ J. INTELL. AUTOM. SOFT. COMPUT., vol. 10, no. 1, pp. 11-22, 2004.

[9] R. M. Milasi, C. Lucas, B. N. Araabi, "Intelligent Modeling and Control of Washing Machines Using LLNF Modeling and Modified BELBIC," in Proc. Int. Conf. Control and Automation., pp.812-817, 2005,

[10] N. Sheikholeslami, D. Shahmirzadi, E. Semsar, C. Lucas., "Applying Brain Emotional Learning Algorithm for Multivariable Control of HVAC Systems," J. INTELL. **FUZZY.** SYST.vol.16, pp. 1–12, 2005.

[11] M. Parsapoor, C. Lucas and S. Setayeshi, "Reinforcement _recurrent fuzzy rule based system based on brain emotional learning structure to predict the complexity dynamic system," in Proc. IEEE Int. Conf. ICDIM,pp.25-32, 2008.

[12] M. Parsapoor, U. Bilstrup, "Brain Emotional Learning Based Fuzzy Inference System (BELFIS) for Solar Activity Forecasting," in Proc. IEEE Int. Conf. ICTAI 2012, 2012.

[13] M. Parsapoor, Lucas.C, Setayeshi.S, “Modifying Brain Emotional Learning Model for Adaptive Prediction of Chaotic Systems with Few Data Training Samples, ’’ in Proc. Int. Conf. ICAOR. pp. 328-341,2008.

[14] M. Parsapoor, M, U. Bilstrup, "Neuro-fuzzy models, BELRFS and LoLiMoT, for prediction of chaotic time series," in Proc. IEEE Int. Conf. INISTA., pp.1-5, 2012.

[15] C. Lucas, A. Abbaspour, A. Gholipour, B. Nadjar Araabi, M. Fatourechi, “Enhancing the performance of neurofuzzy predictors by emotional learning algorithm,’’ J. Informatica (Slovenia)., vol. 27, no. 2 pp.165–174, 2003.

[16] T. Babaie, R. Karimizandi ,C. Lucas, “Learning based brain emotional intelligence as a new aspect for development of an alarm system,’’ J. Soft Computing., vol. 9, issue 9 ,pp.857-873, 2008.

[17] G. Shakhnarovich, T. Darrell,and P.Indyk, Nearest-Neighbor Methods in Learning and Vision:Theory and Practice, MIT Press, March 2006.

[18] C. Blake and C. Merz. (1998). UCI Repository of Machine Learning Databases. Dept. Information & Computer Sciences, Univ. California, Irvine [Online]. Available: http://archive.ics.uci.edu/ml/.

[19] M. Fatourechi, C. Lucas, and A.K. Sedigh, "Emotional Learning as a New Tool for Development of Agent-based System," J. Informatica (Slovenia), vol. 27, no. 2, pp.137-144., 2004.

[20] C. Blake and C. Merz. (1998). UCI Repository of Machine Learning Databases. Dept. Information & Computer Sciences, Univ. California, Irvine [Online]. Available: http://archive.ics.uci.edu/ml/.

[21] M. Wozniak and B. Krawczyk, "Combined classifier based on feature space partitioning, " J.Applied Mathematics and Computer Science vol. 22, no4, pp. 855-866., 2012.

[22] B. Cyganek, "One-Class Support Vector Ensembles for Image Segmentation and Classification, ", J. Mathematical Imaging and Vision, vol. 42, no 2-3, pp.103-117 2012.

[23] R.Burduk, "New AdaBoost Algorithm Based on Interval-Valued Fuzzy Sets, " J. IDEAL pp. 794-801, 2012.