UDC: 004.7

The concept of development of the intelligent tutoring system sensitive to emotions

Vladimir Veljović [1*]
[1] University in Kragujevac, Faculty of Technical Sciences Čačak, Serbia
[1] veljo09@gmail.com

Abstract: *This paper represents the contribution to the development of the intelligent tutor system sensitive to emotions. The proposal of the model of emotion recognition during the learning process, having in mind the necessary sensors and the emotion classification. The use of RBF neural network is suggested because of the simple and temporally hardly changeable nonlineral object approximation provided it is possible to place the center and determine the amounts of variables RBF neurons in advance, where the studying of the network can be reduced only to the settings of crucial coefficients in output layer.*

Keywords: *intelligent tutor systems; neural network; RBF networks*

1. INTRODUCTION

The majority of theories regarding the learning process has mainly ignored the importance of the connection between the learners emotions or emotional states and learning. Nevertheless, it is claimed that cognition, motivation and emotions are three components of the learning process. The emotion is traditionally observed as the root of motivational energy, but it is often not considered as the independent learning or motivational factor. During the last two decades, more attention was paid to the correlation between the emotions and learning process [1]. Various algorithms of machine learning are used in order to detect multiple emotional states through the model of a student [2].

The JAFFE database with the face photos was used [6]. The facial expressions of those defined in [5] were used. The Facial Expression Recognition 1.0 software is used for face recognition, and it is based on the reduction of dimensionality based on neural networks and PCA methods**.**

2. SUGGESTION OF THE EMOTION RECOGNITION MODEL DURING THE LEARNING PROCESS

The aim of this model would be the development of relation between learning and emotions when the learners are interactive with the learning system (computer tutor system). The participants can take part in the training with any section studied, where tutor system asks the questions regarding the section studied. Every interaction with the tutor can be recorded. During the tutorship phase, it can be demanded that the participants express aloud their emotional states. The participants can be provided with the list of emotional states including the following: fury, boredom, confusion, scorn, curiosity, abhorrence, cognition and frustration.

Video shooting of every participant would end with their loud statement of the emotional state they are in. The emotion expression should have the tendency to be very quick and last only around three seconds.

2.1 Sensors

The automatic detection system for the level of learners interest according to the body posture can also be developed. Various algorithms for machine learning can be used in order to classify the learner's static positions which occur most often in the real time (leaning backwards, sitting upright, etc.) Therefore, the interest of the learners could be recognized (high interest, low interest and taking break).

The assessment of the learner's answers regarding their emotional state could also be performed in the model proposed. Based on the interactive session with the student, the system could perform its own assessment of the student's answers. The assessment of the student's answers could include information such as: the correctness of the answer, reactions and the response time, length of the answer and many other parameters.

2.2. Emotion recognition

Two different approaches to the classification of emotions could be applied.

The first would integrate the data from all the three sensors in the high dimensional vector before the classification attempt.

The other way would be to classify individually each entry from the sensor, and then integrate the classification of each sensor into super classifier, in order to obtain one emotion as a result.

Each approach is connected to corresponding advantages and weaknesses, and the only way to determine which one is better is empirical research.

As far as biological motivated classifiers are concerned, some could be used, for example, different types of neural networks. Our exclusive use of noninvasive sensors for emotion detection could have side effect of creating the data with background noise. Because of that, it would be necessary to remove the background noise before the classification using the appropriate techniques.

3. RBF Networks

In the field of mathematical modelling, the network of radial basis functions (Radial Basis Function-RBF networks) is artificial neural network using the radial basis as the activation functions. These neural networks are two-layered static neural networks, where the zero (input) layer forwards the inputs into the network to the entrance of the firs layer conducted of neurons with activation functions with the round basis and represents its so-called receptive field.

The second layer is the network layer, which is also its output layer, and is consisted of perceptrons with linear activation function of identity activation.

RBF network has the ability of approximation of arbitrary continual nonlinear function, and its approximate ability is determined by the position of the center of RBF neuron, variation of activation functions as well as the values of coeficients of the output network layer.

The adequate values of these parameters of RBF network are calculated using the learning algorithms. RBF neural networks are especially used in the case of the approximation of simple and temporally hardly changeable nonlinear when it is possible to place the centers and determine the values of RBF neurons, and where the learning of network can be reduced to the setting of coefficients of the output layer. The behaviour of RBF neural networks, in this case, becomes lineary dependent on parameters.

The properties of RBF network are significantly determined by the position of the center of RBF neurons. RBF functions are traditionally used for the interpolation of many nonlinear variable functions, where the number of centers is equal to the number of the data, so that one center is placed in each input data. The approximation of the arbitrary nonlinear continuous function is also possible to achieve with the smaller number of the well placed centers.

Broomhead and Lowe [Broomhead i Lowe, 1988] suggested in their papers to put the centers in accidentally chosen input data. There is also the possibility of monotonous placement of centers in the space of input entries. The variances of the activation functions influence the network behaviour less and are usually chosen as the second root product of the neuron distance from the two nearest adjacent neurons Moody and Darken, 1989. year [3]. These networks can approximate the arbitrary nonlinear continuous function with accidental equal placement of the centers of RBF neurons, but necessary number of RBF neurons can be very high. We can also achieve the reduce of the RBF neurons number by expansion of the network learning method and by adjusting the center position. The RBF network behaviour, in this case, becomes nonlinear dependent on parameters, but also with comparable approximative features [4].

4. RBF NETWORK TRAINING

Optimal architecture of RBF network is usually determined experimentally, but some practical directives also exist. Problem solving procedure using neural networks consists of: collecting and preparation of the data, network training, network testing, and determining the optimal parameters of the network and training experimentally (number of neurons, number of neuron layers, learning algorithm parameters and training data).

The data preparation for RBF networks includes: Filtering, normalization and dimensionality reduction. The success of the solving depends completely on the data used for network training. It is necessary to be careful of theoretical adequacy – the representativeness of the data used for certain problem. This is very specific depending on the problem which is being solved. The RBF training includes: the determining of optimal network parameters and training algorithm, determining the number of hidden layers and the number of neurons in every layer (more doesn't mean better, the goal is to have less), dynamic parameter setting, parameter validation (with trial set), determining the training and the testing of the data and solving of overtraining problem and generalization.

Output weights training is simple when the output neurons use linear activation. There are three kinds of parameters in the RBF network necessary to be determined for adjustment of the network for certain task: middle vectors C_i, output weights ω_i and RBF width parameters β_i.

As far as the sequential training is concerned, the weights are updated in every temporal step. For some tasks, there is sense in defining the aim function and choosing the parameter's value which minimize its value. The most present aim function is the smallest square function, which explicitly includes the dependence on the weight.

$$K(\omega) \overset{def}{=} \sum_{t=1}^{\infty} K_t(\omega)$$

where

$$K_t(\omega) \overset{def}{=} \left[y(t) - \varphi(x(t),\omega) \right]^2$$

The minimization of the smallest square aim function assisted by optimal choice of weight choice is optimizing accuracy.

There are some situations where many aims, such as the smoothness and accuracy, has to be optimized. In that case, it is useful to optimize the regulated aim function as

$$H(\omega) \overset{def}{=} K(\omega) + \lambda S(\omega) \overset{def}{=} \sum_{t=1}^{\infty} H_t(\omega)$$

where

$$S(\omega) \overset{def}{=} \sum_{t=1}^{\omega} S_t(\omega)$$

and

$$H_t(\omega) \overset{def}{=} K_t(\omega) + \lambda S_t(\omega)$$

where optimization *S* maximizes the smoothness and λ .

5. EXPERIMENTAL RESEARCH

Based on the JAFFE database with face photos (available at http://www.kasrl.org/jaffe.html) where ten people set three or four examples of each of the six basic facial expressions (happiness, sadness, surprise, anger, abhorrence, fear) [5] and neutral face with the total of 219 photos of facial expressions. Because of the experimental research's simplicity, only the Japanese female persons were used. The photo examples are shown in picture 1.

Based on the research defined in [7] and [8] for emotion recognition, the software Facial Expression Recognition 1.0 is used, and it is based on the neural networks and PCA method. This algorithm for facial expression recognition classifies the given picture into one of the seven main facial expression categories (happiness, sadness, surprise, anger, abhorrence, fear and neutral). The PCA method is used for the output data dimension reduction. The method keeps those characteristics of the data set which contribute to its variance the most, keeping the main components of the low order and ignoring those of the higher order. That is why the components consisting of the low order contain "the most important" data aspects. The drawn function vectors in the reduced space is used for the neural network classifiers training for classification. The suggested method is quick and can be used for applications in the real time.

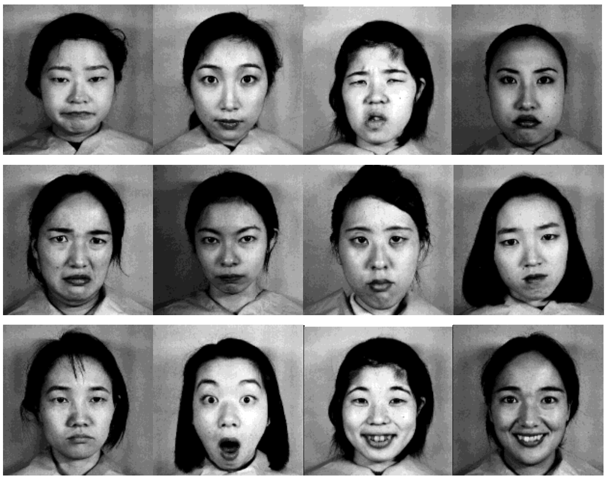

Figure 1. *Sample images JAFFE database [6]*

The face classifier was primarily tested by using the photos with the face expressions of the seven Japanese ladies with following initials: KA, KL, KM, KR, MK, NA i NM. Each of these persons has set three or four copies of each of the six basic facial expressions, as well as their neutral face. The set of images is divided into seven sets, each of the sets corresponds to one person. The system is trained on seven segments, and then tested on a new data set, which is not used in training. In testing, it was noticed that the system does not recognize all the facial expressions equally true.

Table 1 represents the confusion matrix showing the wrong classifications for the data set of the woman with the initial YM. The results of our simulation experiments show that neural networks are effective in recognizing emotions using facial expressions, when we tested six emotions, we achieved a recognition rate of about 50%.

Table 1: *Rates of the wrong classification for the person with the initial YM. HAP - happiness, USA - sadness, SUR - surprise, ANG - anger, DIS - depression, FEA - fear.*

ANG	DIS	FEA	HAP	SAD	SUR	I/O
3	-	-	-	-	-	ANG
-	-	-	-	-	-	DIS
-	-	-	-	-	-	FEA
-	-	-	-	-	-	HAP
-	3	2	1	3	-	SAD
-	-	-	-	-	3	SUR
-	-	1	2	-	-	NEU

It should be noted that it is not easy to compare classification performance, because facial expressions are not always pure specimens of one category of expressions. It is important to understand that expressions are never a pure expression of an emotion, but they always represent different emotions. In each of the images used in the research, only the dominant expression in this figure is presented - the expression that the person was asked to pose.

6. CONCLUSION

This research could help in future papers, such as capturing non-stationary images in real time, and simultaneously analyzing these images in accordance with the techniques of affective computing. In the future, improvements in image selection methods are also possible. Algorithms and data sets will be selected according to precise criteria: classify algorithms and several data sets. Also, these conclusions and recommendations can be tested on larger sets of data using different classification algorithms in the near future. The broader impact of the proposed research could contribute to the improvement of education, intelligent learning environments and interfaces between people and computers.

The main computing effort should be focused on achieving new methods of estimation and classification of affective emotional states. Some of the existing emotional classifiers can be used, and which would possibly be updated. A common classifier for these needs are neural networks.

Neural networks are an acceptable tool when problem solving approaches are standardized in such a way that for each situation the network architecture, learning rules, the number of hidden layers, and the number of neurons in them, as well as the portable functions that need to be used in solving, are exactly defined.

REFERENCES

[1] B. Kort, R. Reilly, R. Picard. An affective model of interplay between emotions and learning: Reengineering educational pedagogy—building a learning companion. In T. Okamoto, R. Hartley, Kinshuk, & J. P. Klus (Eds.), Proceedings IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges (pp.43-48). Madison, Wisconsin: IEEE Computer Society, 2001.

[2] D. Milošević. Ontološko inženjerstvo u inteligentnim tutorskim sistemima, monografija, Tehnički fakultet Čačak, 2008.

[3] J. Moody, C. Darken. Fast learning in networks of locally-tuned processing units, Neural Computation, 1:281-294, 1989.

[4] I. Petrović. Osnove inteligentnog upravljanja (sustavi upravljanja zasnovani na umjetnim neuronskim mrežama), Zagreb, 2009.

[5] P. Ekman & W. V. Friesen. Unmasking the Face. A guide to recognizing emotions from facial clues, Palo Alto: Consulting Psychologists Press, 1975.

[6] http://www.kasrl.org/jaffe.html

[7] G. Mandler. Emotion. In B. M. Bly & D. E. Rumelhart (Eds.), Cognitive science. Handbook of perception and cognition 2nd ed. (pp. 367-384). San Diego, CA: Academic Press, 1999.

[8] D. K. Meyer, J. C. Turner. Discovering emotion in classroom motivation research. Educational Psychologist, 37, 107-114, 2002.