Moridis, C. N. & Economides, A. A. (2009). Prediction of student's mood during an online test using formula-based and neural network-based methods. Computers & Education. Elsevier. ISSN: 0360-1315. [Impact factor SSCI=0.968 (2005), 1.085 (2006), 1.602 (2007); A* (top 5% journal) in Australian Ranking of Journals (2009); 'A' journal in ERIH list (2007); SJR= 0.057 (2007); H-index= 21]. Accepted

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Prediction of student's mood during an online test using formula-based and neural network-based method

Christos N. Moridis *, Anastasios A. Economides

Information Systems Department, University of Macedonia, 156 Egnatia Avenue, Thessaloniki 54006, Greece

ARTICLE INFO

Article history:
Received 31 July 2008
Received in revised form 21 March 2009
Accepted 5 April 2009
Available online xxxx

Keywords:
Intelligent tutoring systems
Interactive learning environments
Pedagogical issues
Architectures for educational technology system

ABSTRACT

Building computerized mechanisms that will accurately, immediately and continually recognize a learner's affective state and activate an appropriate response based on integrated pedagogical models is becoming one of the main aims of artificial intelligence in education. The goal of this paper is to demonstrate how the various kinds of evidence could be combined so as to optimize inferences about affective states during an online self-assessment test. A formula-based method has been developed for the prediction of students' mood, and it was tested using data emanated from experiments made with 153 high school students from three different regions of a European country. The same set of data is analyzed developing a neural network method. Furthermore, the formula-based method is used as an input parameter selection module for the neural network method. The results vindicate to a great degree the formula-based method's assumptions about student's mood and indicate that neural networks and conventional algorithmic methods should not be in competition but complement each other for the development of affect recognition systems. Moreover, it becomes apparent that neural networks can provide an alternative for and improvements over tutoring systems' affect recognition methods.

1. Introduction

Recent research in affective neuroscience and psychology have reported that human affect plays a significant and useful role in human learning and decision making, as it influences cognitive processes (Bechara, Damasio, Tranel, & Damasio, 1997; Goleman, 1995). However, the extension of cognitive theory to explain and exploit the role of affect in learning is in its infancy (Picard et al., 2004). Researchers of artificial intelligence in education have considered the fundamental nature of integrating emotional factors in intelligent tutoring systems. A step towards this direction is to provide computer aided learning systems with an automatic affect recognizer, in order to collect data which identify a student's emotional state. With this information, the computer could respond appropriately to the student's affective state rather than simply respond to student's commands (Lisetti & Schiano, 2000; Picard, 1997). An appropriate computer response to a student's affective state also requires evolving and integrating new pedagogical models into computerized learning environments, which assess whether or not learning is proceeding at a healthy rate and intervene appropriately (Kort, Reilly, & Picard, 2001).

Nevertheless, the risk of inappropriate interactions takes several forms. For example, if an artificial agent is overly excited about a learner's success, the learner may feel awkward which may lessen his/her motivation for continued interactions with the agent and on the task (Burleson & Picard, 2004). Providing individualized feedback according to student's cognitive and affective states has been neglected until recently when its value has now become more apparent (Economides, 2005; Economides, 2006; Mavrikis, Maciocia, & Lee, 2003). However, too much feedback may also prove detrimental if it results in information overload, unnecessary interruptions or an irrational amount of pressure (Alder, 2007).

Knowledge relative to how emotions influence learning is a fundamental part of computer-aided affective learning systems. A number of theoretical models of learning assumed that learning occurs in the presence of affective states (Craig, Graesser, Sullins, & Gholson, 2004). Henceforth, it is recognized that positive and negative emotional states trigger different types of mental states and this can have an important influence on the learning process (Moridis & Economides, 2008a). It seems that extreme emotional states, either positive or negative, are not beneficial to concentration and learning. Negative affect initially focuses the mind, leading to better concentration (Schwarz & Bless, 1991). In situations of an urgent threat, this is favourable since it concentrates processing power upon the danger. When creative problem

* Corresponding author. Tel.: +30 2310 891768; fax: +30 2310 891292.
E-mail addresses: papaphilips@gmail.com (C.N. Moridis), economid@uom.gr (A.A. Economides).

0360-1315/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compedu.2009.04.002

solving is necessary this is unfavourable, since it leads to narrow tunnel vision (Norman, 2002). Positive affect widens the thought processes, making it easier to be distracted. When the problem involves focusing, positive affect may interfere with the subject's concentration, whereas when the problem is treated through creative thinking then the results are optimal. Similarly, the proper amount of anxiety or fear can help individuals to focus, because anxiety focuses the mind reducing distractions. It is when the negative affect is too strong that learning tasks are inhibited (Bower, 1992).

Yet, this knowledge would have no use in emotional instructional technology, if these systems were not able to recognize a student's emotional state. Humans recognize emotional states in other people by a number of visible and audible cues. Facial expression is a valuable means in the communication of emotion. Moreover, there is evidence of the existence of a number of universally recognized facial expressions for emotion such as happiness, surprise, fear, sadness, anger and disgust (Ekman, 1982). In addition, the body (gesture and posture) and tone of voice are the other core channels for the communication of emotion (Argyle, 1988). There are also a number of psycho-physiological correlates of emotion, such as pulse or respiration rate, most of which cannot easily be detected by human observers, but which could be made accessible to computers given appropriate sensing equipment. From all of these channels researchers of artificial intelligence in education are attempting to infer the student's affective state.

Emotional recognition frameworks using personal preference information are based on the assumption that people do not necessarily recognize emotions just by signals seen or heard; they also use a high level of knowledge and reason, to be able to process the goals, situations, and preferences of the user. A person's emotions could be predictable if his/her goals and perception of relevant events were known (Ortony, Clore, & Collins, 1988). Implemented in a computational model this can be achieved by using agents, artificial intelligence techniques, reasoning on goals, situations, and preferences (Conati, 2002). For example, if the system can reason about the reactions of a user from the input that the system receives, (assumption made derived from the time of day, speed of reading, provided personal information, etc.) appropriate content could be displayed in a way adapted to the emotion or the mood of the user.

Emotion recognition systems are generally based on a rule base system, or on a system that has learnt to solve the problem through extensive training (Caridakis, Karpouzis, & Kollias, 2008). The richer the information provided by the interaction is, the more parameters can be derived for extracting the interaction environment and for achieving a better emotion recognition performance. The goal of this paper is to indicate how various kinds of evidence could be combined to optimize inferences about affective states during an online self-assessment test.

With regard to learning, there have been very few approaches for the purpose of affect recognition. The adoption of affect recognition methods using personal preference information and questionnaires would probably be more preferable for certain affective learning systems (e.g. web-based for distance learning). These methods do not require special equipment, such as video cameras, microphones, sensors, etc., rendering the affective learning system more user-friendly. However, the usual techniques of getting self-report about a student's emotional state, through questionnaires and/or textboxes, have obvious disadvantages (Moridis & Economides, 2008a). Questionnaires are easy to manage but have been criticized for being static and thus incapable of recognizing changes in affective states. In addition, the way questions are framed and demonstrated, the order in which questions are presented and the terms employed in questions are all known to influence the subject's responses (Anderson, 1982; Lindgaard, 1995).

The mood recognition methods developed in this paper are personal preference information based. So as to exceed some of the limitations of personal preference information frameworks, a slidebar was used to acquire self-report about student's mood, during the experiments process, aiming to evaluate the mood recognition effectiveness of the proposed methods. The students where asked to move the slidebar according to their mood from −100 (extremely negative mood) to +100 (extremely positive mood). Using a slidebar to evaluate the mood recognition methods it was intended to avoid verbalization, so that students would give feedback concerning their mood without their activity being as disrupted as if they had to plainly verbalize their current mood. In addition, by avoiding verbalization and using a slidebar instead, it was possible for students to declare their current mood through a less culturally dependent process (Isbister, Höök, Sharp, & Laaksolahti, 2006). Students during a test may experience negative emotions such as anger, sorrow, despair etc., or positive emotions such as joy, hope, pride etc. Because of the difficulty to distinguish between each negative and positive emotion in an emotional recognition framework using personal preference information, we decided to group negative and positive emotions under negative and positive mood respectively.

Mood and emotion have common features, but also have distinctions (Larsen, 2000). Two basic distinct characteristics of mood could carry valuable information concerning a student's learning experience: (1) Duration and intensity: The duration is a distinguishing of the mood, while the intensity is a distinguishing of emotion; (2) Information: Emotion carries information concerning the environment, e.g. information about a threat in our environment, while mood carries information concerning our capacity to face the threat of the environment. That is that during a test mood could carry information concerning a student's self-evaluation about his/her capacity to successfully undertake the test. Moreover, duration as a characteristic of mood could serve long-term learning goals. The student should have a positive approach towards learning, both during and after interaction with the tutoring system (Moridis & Economides, 2009b). Consequently, it may be more useful in the long term for affective tutoring systems to produce a positive mood concerning learning, than to be focused on the emotion of the moment.

Furthermore, affective states are created through the interaction of social and personal factors (Boehner, DePaula, Dourish, & Sengers, 2007). That is to say that emotion is produced culturally and through personal interaction. Consequently, when we attempt to model affective issues during students' interaction with the on-line self-assessment test environment, these factors play a crucial role and therefore should be taken into account. Moreover, these factors should be integrated into an affective student's model when trying to provide the student with adequate feedback, so as to introduce him/her into an emotional state beneficial to learning (Economides, 2005; Economides, 2006). Hence, the proposed methods are built in such a way so that information about student's interaction with the system can be incorporated during the self-assessment test. This kind of information will make it possible for the system to provide adequate feedback on student's current mood based on student's interaction with the system. Social and personal traits could provide a basis about which factors and in which way could be used to influence student's mood favorably to learning (Economides, in press).

For example, in the case of embodied agents it has been shown that displaying a female character to reduce user's frustration could be more effective than displaying a male character (Hone, 2005). The female character proved more effective than the male character since the female gender is usually more related with qualities such as empathy. Thus, gender stereotypes coming from the real world can apply to human–computer interaction (Reeves & Nass, 1996). Similarly, other social traits could be employed to alter student's mood.

Please cite this article in press as: Moridis, C. N., & Economides, A. A. Prediction of student's mood during an online test using formula-based and neural network-based method. *Computers & Education* (2009), doi:10.1016/j.compedu.2009.04.002

The methods developed in this paper take advantage of students' personal goal which depends on personal and social factors, such as parents' expectations, personal ambition etc. (Zimmerman, Bandura, & Martinez-Pons, 1992). Therefore, students were asked to set a personal performance goal at the beginning of the test, and students' success or failure to reach this goal was monitored by the models during the test.

Based on assumptions, a formula-based method has been developed (Moridis & Economides, 2008b) in order to provide a measurement for the estimation of student's mood while undertaking an online self-assessment multiple choice questions test. The method was tested using data emanated from experiments made with 153 high school students from 3 different regions of a European country.

In this paper, a neural network-based method is also used to predict student's mood. Neural networks contain no preconceptions of what the model shape will be, so they are ideal for cases with low system knowledge. They are useful for functional prediction and system modelling where the physical processes are not understood or are highly complex (Zhang, Patuwo, & Hu, 1998). The proposed neural network is initialized randomly and must undergo "supervised learning" before use as a mood recognizer. This requires knowledge of the desired output for each input vector. The neural network in this experiment was trained using the same dataset emanated from experiments made with high school students. Initially, the neural network was trained using eight input parameters that intuitively seemed to have an influence on student's mood. Then, it was assumed that using the formula-based method as an input parameter selection module would help to apply logic on the available data before importing them to the neural network. This step reduced the input parameters from 8 to 5, saving the neural network from unnecessary compute load and increasing its effectiveness. Indeed, in this case the neural network mood recognizer performed better, indicating that hybrid structures would be preferable for the construction of affect recognition systems.

The results verified our assumptions and proved both the formula-based and the neural network method's ability to approximate student's mood at a satisfactory level. The neural network's success to recognize student's mood based on the same assumptions as the formula-based method is an extra verification that the chosen input parameters have an influence on student's mood while undertaking a test. Furthermore, the results imply that neural networks can provide a high-quality solution for tutoring systems' affect recognition methods using personal preference information.

2. Formula-based method and experimental process

Several research questions in the context of an online multiple choice questions self-assessment test have been explored (Moridis & Economides, 2008b; Moridis & Economides, 2008c; Moridis & Economides, 2009a), providing a measurement for evaluating the students' mood during the test. To do so, two formula-based models were formulated and evaluated on the basis of assumptions that could have an impact on a student's mood during the test. The more effective of the two is used here to be combined with the neural network-based method, introduced later in this paper. For the sake of completeness the formula-based mood recognition model and the experiments that took place are briefly described in the following paragraphs.

A basic assumption was that student's goal does influence the student's mood during the test in relation to the remaining questions and his/her record. That is to say, that if a student knows that he has already failed to reach his/her goal during the test, because the remaining questions are less than the questions he has to answer correctly in order to reach his/her goal, then there is a high possibility to be in a negative mood. In addition to that it was assumed that student's mood is also influenced by his/her success or failure to answer the questions just before the current one. For instance, if a student has failed to provide a correct answer to all of the 5 previous questions, there is a high possibility that his/her mood would be negatively influenced, but if a student has managed to provide a correct answer to all of the 5 previous questions, there is a high possibility that his/her mood would be positively influenced. With the purpose of checking these assumptions we formulated this method:

$$R(q) = N - q, \quad R(q) \in (0, N) \tag{1}$$

where R is the number of questions remaining before the end of the test, N is the total number of questions, and q is the number of the current question.

$$D(q) = I - r(q), \tag{2}$$

where $D(q)$ is the number of questions that the student still needs to answer in order to reach his/her goal, I is the student's goal, and $r(q)$ is the number of student's correct answers up to the current point.

$$H(q) = R(q) - D(q), \tag{3}$$

where $H(q)$ is a number that shows whether the remaining questions are enough for the student to reach his/her goal. For example, $H(q) = -3$ would mean that the student has already failed to reach his/her goal for 3 questions.

$$M(q) = \begin{cases} H(q) + rr(q) \\ H(q) - wr(q) \end{cases} \tag{4}$$

where $M(q)$ is the student's mood, $rr(q)$ is the number of correct answers one after the other just before the current question, and $wr(q)$ is the number of wrong answers one after the other just before the current question. So, if there are one or more correct answers one after the other just before the current question, we add them to $H(q)$, while if there are one or more wrong answers one after the other just before the current question, we subtract them from $H(q)$.

2.1. Experiment process

2.1.1. System architecture

We build the online multiple choice questions test in a Windows XP machine using JavaScript with Perl CGI on Apache web server with MySQL. Common Gateway Interface (CGI) is used to provide dynamic web pages to students and the Perl programming language is a common choice for various reasons including simple and powerful string manipulation, Web server integration (i.e. Apache web server), and

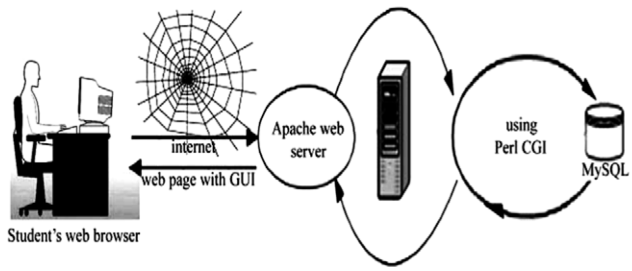

Fig. 1. System architecture.

Table 1
Linear correlation coefficient (r) for each variable.

R	D	H	rr	wr	r	w
0.23	−0.39	0.64	0.32	−0.25	0.02	−0.29

data manipulation (Guelich, Gundavaram, & Birznieks, 2000). Many students took the online test simultaneously, thus MySQL which is designed for multiple users accessing the files through a single server process was a very convenient choice for this kind of work (Kofler, 2005).

The student's terminal sent a standard HTTP request encoded with special variables (commonly called CGI variables) and their values. The Apache web server received the HTTP request and determined that the request was destined for a CGI program and not for a HTML file. The Perl interpreter accepted the input from the server, parsed the variables, contacted the MySQL database for reading and recording data, applied some programming logic to the data, and returned a document back to the web server (Fig. 1).

2.1.2. Participants

High schools students ($N = 153$) were recruited from three different regions of a European country (60 students from the capital, 50 students from a large city, and 43 students from a town). Respondents consisted of 56% females and 44% males. The average age of participants was 16.8 (SD = 1.98) with 90% of the sample ranging from 15 to 19 years.

2.1.3. Materials

The multiple choice questions were focused on basic computer knowledge and skills, based on material taught in lectures. The context of questions was pre-specified by the teachers prior to the study. The test was constituted from 45 questions. The order of questions presented was randomly altered among students.

2.1.4. Procedure and data collection methodology

The duration of the experiment was approximately 45 min and took place during the regular schedule of laboratorial classes. Students were told that this is a general education test concerning computer knowledge that would help them assess their knowledge about computers. At the beginning of the test, the system asked students how many correct answers would make them feel satisfied by the level of their knowledge, making them set their personal goal. Throughout the test, a student selected his/her answer among four possible answers and confirmed his/her choice by clicking the "submit" button. After each question the system informed the student whether his/her answer was right or wrong and presented his/her score. Then the student could proceed to the next question by clicking the "next" button.

During these 45 questions a slide bar appeared asking the student to move it according to his/her mood concerning the test, from −100(extremely negative mood) to +100(extremely positive mood). The slide bar appeared five times during the test, once every 9 questions, at a different instant for each student, as shown in Table 1. Accordingly, it took 9 students for the slide bar to appear once after every question of the test. So, the 153 participants gave us the chance to check students' mood after every question 17 times. Thus, the data set consisted of 765 instances (5 instances of each student). The slide bar appeared once every 9 questions during the test because of the danger of irritating the student.

Each time the student declared his/her mood by moving the slide bar, 10 parameters were calculated and recorded: (1) The number of the current question, (2) the number of questions remaining before the end of the test, (3) the number of questions that the student still needs to answer in order to reach his/her goal, (4) the number that shows whether the remaining questions are more or less than the number of questions that the student has to answer so as to reach his/her goal, i.e., student's hope to reach his/her goal, (5) the number of correct answers one after another before the current question, (6) the number of wrong answers one after another before the current question, (7) the number of correct answers up to the current question, (8) the number of wrong answers up to the current question, (9) the score, and (10) the mood that the student indicated by moving the slide bar.

3. Neural network-based method

Affect recognition systems using conventional algorithmic methods follow a set of instructions in order to recognize emotional states. However, unless the specific steps that need to be followed are known, these methods can not function effectively. Besides, the problem solving capability of conventional algorithmic methods is restricted to problems that we already understand and know how to solve. But methods for affect recognition would be so much more useful if they were efficient in domains, such as affective learning, that we do not exactly know how to deal with. To solve this inefficiency, neural networks, with their significant ability to derive meaning from complicated or vague data, can be used to extract patterns and identify trends that are too complex to be noticed by either humans or other computer techniques.

Therefore, neural networks are being useful to an increasing variety of applications besides the traditional areas such as pattern recognition and control. Non-linear learning capabilities provide neural networks an advantage over conventional methods for solving certain problems, in which generalization capabilities are indispensable. Neural networks, being able to model any arbitrarily complex nonlinear relationship, are not restricted by a predefined mathematical relationship between dependent and independent variables (White, 1989). A data analysis between the variables, which where recorded and calculated during the experiments, used from the formula, the neural network and the hybrid method to predict student's mood (input), reveals that there is no significant linear relationship between them and the student's declared mood on the slidebar (output to be predicted). Table 1 shows the linear correlation coefficient (r) for each variable:

Thus, in this case neural networks may represent an attractive alternative to conventional techniques. Techniques like neural networks, which are inherently suited to addressing complex relationships, are likely to have a key place in describing the patterns that exist concerning affect recognition (Cowie, Douglas-Cowie, & Cox, 2005). This paper confirms points of theory which state that neural networks could be useful for the task of affect recognition. An automatic mood recognition system that employs neural networks, such as the one developed in this paper, requires sufficient training and testing material. This material should contain two streams: an input stream and an output stream. As far as student's mood is concerned while undertaking a test, the input stream would consist of the extracted relevant parameters (student's goal, correct or wrong answers, remaining questions, etc.) and the output would include the emotional class or category or more generally the student's mood for which the input parameters were extracted.

3.1. A brief introduction to neural networks

Neural networks are inspired by biology, and attempt to imitate learning processes in the brain. As in nature, the network function is determined largely by the connections between elements. As the knowledge on nervous systems is not highly developed, we have to define different functionalities and connection structures apart from a biological perspective. We can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements (Fausett, 1994).

An artificial neuron is an element with inputs, output and memory that may be implemented with software or hardware. It has inputs that are weighted, added, and compared with a threshold, although there can be various other types of neural network architectures. The way neurons connect to each other is generally referred to as the architecture of the neural network. Whenever the neural network makes an error, some weights and thresholds have to be changed to compensate for this error. The rules, which govern how exactly these changes are to take place, are called as the learning algorithm. Different types of neural networks may have different learning algorithms. Thus, when we refer to various types of architecture, it means the set of possible interconnections and the learning algorithm defined for it.

One way to train a neural network is to initialize it with random weights and then provide a series of inputs. An outstanding feature of neural networks is their learning ability. They learn by adaptively updating the synaptic weights that characterize the strength of the connections. The weights are updated according to the information extracted from new training patterns.

Artificial neural networks are often classified into two distinctive training types, supervised or unsupervised (Leea, Boothb, & Alamc, 2005). Supervised learning networks have been the mainstream of neural model development. The training data consists of many pairs of input/output training patterns. For an unsupervised learning rule, the training set consists of input training patterns only. The network learns to adapt based on the experiences collected through the previous training patterns. There are two phases in neural information processing. They are the learning phase and the retrieving phase. In the training phase, a training data set is used to determine the weight parameters that define the neural model. This trained neural model will be used later in the retrieving phase to process real test patterns and yield classification results.

Of the many neural network architectures proposed, the multilayer perceptron (MLP) with back-propagation learning algorithm is found to be effective for solving a number of real-world problems (Li & Yu, 2002; Murphey & Luo, 2002). Thus, the neural network created in this paper is of a MLP network structure. The MLP network structure consists of an input layer, one or more hidden layers, and one output layer. Each layer can have one or more neurons (Fig. 2).

The traditional training method is the standard back-propagation (Rumelhart, Hinton, & Williams, 1986), although numerous training schemes are available to impart better training with the same set of data. According to Demuth and Beale (1995), the Levenberg–Marquardt updating rule is more effective than the delta rule, which is used in regular back-propagation. Trainlm is a MATLAB network training function that updates weight and bias values according to Levenberg–Marquardt optimization. Trainlm is often the fastest back-propagation algorithm in the MATLAB neural network toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more Random Access Memory (RAM) than other algorithms (Hadi, 2003).

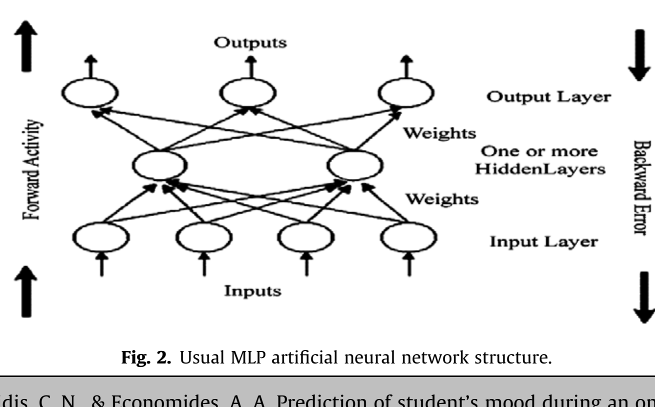

Fig. 2. Usual MLP artificial neural network structure.

3.2. Software

In order to develop the neural network method the neural network toolbox 4.0.6 and MATLAB 7.1 were used. A MATLAB script was written which loaded the data file, trained and validated the networks and saved the models' architecture and performance in a file ready for further use. The input and output data were normalised and de-normalised before and after the actual application in the network. The neural network toolbox offers a broad variety of parameters for neural network development which can be chosen flexibly. MATLAB implements numerous learning algorithms described by Demuth, Beale, and Hagan (2005).

By applying the developed MATLAB script it was possible to search for the best model in an easy manner. Several hundreds of models were trained and validated, through the process of trial and error, until a satisfactory architecture and MATLAB Neural Network Toolbox settings for the method were developed.

3.3. Proposed neural network

One of the most serious troubles that arise in learning by neural networks is overfitting of the provided training examples. This means that the learned function fits very closely the training data but it does not generalise well, that is it can not model sufficiently well unseen data from the same task (Lawrence, Giles, & Tsoi, 1997).

A three-layer feedforward neural network was trained as a mood recogniser using trainlm. The neural network had the tansig transfer function in hidden layer and the purelin transfer function in output layer. Transfer functions calculate a layer's output from its net input. The tansig and purelin transfer functions used in MATLAB are tan-sigmoid and linear transfer function respectively. Each input vector was matched to the desired output, which is the student's mood. The dataset emanated from the experiments with 153 high school students consisted of 765 vectors. The total available data were randomly divided into a training set (sample data) of 150 vectors and a test set of 615 vectors. To avoid overfitting, a smaller part of the available data was used as a training set and a bigger part as a test set. For the same reason, the learning process was limited to 100 epochs.

Building a neural network is one thing, but making it a useful application is something much more important. A key point to neural network's success to recognize a student's mood is the proper selection of the input parameters. For this purpose, the results of two stages of training were compared to decide which of the two would be more effective for selecting the input parameters to the neural network mood recognizer:

3.3.1. Stage 1

At a first stage, the neural network was initially trained using 8 input parameters that intuitively seemed to have an influence on student's mood: (1) The number of the current question, (2) the number of questions remaining before the end of the test, (3) the number of questions that the student still needs to answer in order to reach his/her goal, (4) the number that shows whether the remaining questions are more or less than the number of questions that the student has to answer so as to reach his/her goal, i.e., student's hope to reach his/her goal, (5) the number of correct answers one after the other just before the current question, (6) the number of wrong answers one after the other just before the current question, (7) the number of correct answers up to the current question, (8) the number of wrong answers up to the current question. The neural network of stage 1 had one neuron for each input parameter (8 neurones) in input layer, 16 neurones in hidden layer, and 1 neuron in output layer.

3.3.2. Stage 2

At a second stage, it was assumed that an input parameter selection module (in our case the formula-based method described in Section 2) would reduce the compute load and increase the generalization ability of the neural network mood recognizer. That is because issues related to emotion may have a clear cause (e.g. happiness because of the presence of a good friend), but they can also be difficult to explain (e.g. sadness after winning the lottery). Affect seem to be on both sides of the logical, discrete representations that conventional algorithmic methods handle carefully, and the non-symbolic representations that neural networks construct. It is an advantage of neural networks that they have the potential to allow evidence to make the emergence of fitting intervening structures. However, it is a risk that they may generate weighting patterns which work mostly in a limited area but which can neither be understood nor extended. Therefore, hybrid structures would be preferable for addressing those issues (Cowie et al., 2001). For that reason, the neural network was also trained using as input parameters only the output of the functions of the formula-based method: (1) The number of questions remaining before the end of the test, (2) the number of questions that the student still needs to answer in order to reach his/her goal, (3) the number that shows whether the remaining questions are more or less than the number of questions that the student has to answer so as to reach his/her goal, i.e., student's hope to reach his/her goal, (4) the number of correct answers one after the other just before the current question, (5) the number of wrong answers one after the other just before the current question. The neural network of stage 2 had one neuron for each input parameter (5 neurones) in input layer, 16 neurones in hidden layer, and 1 neuron in output layer.

4. Results

Statistical analysis was performed using extra code written in Matlab. The calculation of the mean error was normalised from 0 to 1. Initially, the neural networks of stage 1 and of stage 2 were evaluated based on their divergence from the mood that students pointed out on the slide bar. Nevertheless, this is a quantitative way of evaluation. The linear correlation coefficient between the user declared mood and the one estimated by the neural network method trained with the input vector of stage 1 is strong too ($r = 0.7$). The mean error for the neural network of stage 1 is 0.08 with a standard deviation of 0.1 and has a normal error distribution. A confidence level of 90% for the neural network mood recognizer of stage 1 gives us a confidence interval of 0.01, which means that the range for true population mean is 0.07–0.09.

The neural network trained with the input vector of stage 2 has a stronger linear correlation coefficient ($r = 0.74$). The mean error for the neural network of stage 2 is 0.04 with a standard deviation of 0.13 and has a normal error distribution. A confidence level of 90% for the

Table 2
Quantitative comparison between the neural network of stage 1 and the neural network of stage 2.

Methods	*r*	Mean error	SD	Confidence level (%)	Confidence interval	Range for true population mean
Neural network (stage 1)	0.7	0.08	0.1	90	0.01	0.07–0.09
Neural network (stage 2)	0.74	0.04	0.13	90	0.02	0.02–0.06

Table 3
Qualitative comparison between the neural network of stage 1 and the neural network of stage 2.

Methods	Mean success recognizing whether student is in a positive or negative mood (%)	Mean success recognizing whether student is in a positive mood (%)	Mean success recognizing whether student is in a negative mood (%)
Neural network (stage 1)	82	84	80
Neural network (stage 2)	86	87	84

neural network mood recognizer of stage 2 gives us a confidence interval of 0.02, which means that the range for true population mean is 0.02–0.06.

It is obvious that the neural network of stage 2 performs better than the neural network of stage 1, which indicates that hybrid structures would be preferable for the construction of affect recognition systems (Table 2).

Trying to determine the exact percentage of student's positive or negative mood is a difficult task. However, if we try to determine just whether a student is in a positive mood or in a negative mood, things are getting much easier. So, in a qualitative evaluation of the two methods, we would judge the methods by their success in predicting whether a student is in positive or negative mood, no matter how positive or how negative this mood is. Again, the neural network of stage 2 performs better than the neural network of stage 1. Using the hybrid method we can have a safe prediction of student's mood in terms of whether this mood is positive or negative (Table 3). A problem that is difficult to be solved from a quantitative point view becomes more manageable from a qualitative point of view.

5. Discussion

Students receive emotional pressure from many different sources (teachers, schoolmates, parents, themselves etc.) to perform well in a learning task. Educational learning and success are among the most significant areas across a person's life span in our society nowadays, particularly since educational and professional careers, social relations, and the distribution of many kinds of resources are largely dependent on individual achievement. This indicates that learning and achievement are essential and therefore key sources of human emotions nowadays, instigating a variety of self-referenced, task-related, and social emotions. More important, judging from the general functions of emotions for human agency, it may be deduced that emotions influence student's cognitive processes and performance, as well as his/her psychological and physical health (Pekrun, Goetz, Titz, & Perry, 2002). Consequently, student's failure to accomplish a learning task may be due to emotional disability. The idea is to have an online self-assessment test that would help students psychologically and cognitively during their preparation for the exams.

In the long term, the aim of this paper is to provide foundations for the development of an online platform of multiple choice questions. The students, using this platform, will be able to improve their knowledge and acquire a positive attitude towards learning. The system would help students to improve their affective state and enhance their cognitive weaknesses before exams, so that they will be psychologically and cognitively equipped to deal with the final test. Whether feedback should be also incorporated into final exam tests, is a controversial matter. It could be however incorporated into the final tests of several courses in elementary or high schools, or universities. We do not expect that tests like GRE (Graduate Record Examination) or GMAT (Graduate Management Admission Test) will incorporate any emotional or cognitive feedback routine into their code. However, a feedback routine incorporated into preparation tests for practice could help the candidates formulate an affective state favourable to learning and advance their knowledge during preparation for the test.

Summarizing, we are dealing with a tutoring system capable of handling students' affective and cognitive needs during an online self-assessment test. Obviously, the presented work is not enough to construct such a system. A system of this kind would have to be able to: (1) Recognize the running affective state of the student; (2) Recognize when to intervene in order to influence the student's affective state, based on a new educational pedagogy integrating affective models in learning; (3) Produce the most optimal affective state for learning. In addition, the tutoring system's affective handling has to be successfully rooted with student's cognitive handling. Therefore, further research related to the above mentioned requirements, although they belong to various fields of research, is indispensable to the evolution of such a system.

6. Conclusions

Modelling student's mood during an online self-assessment test through a hybrid method, using a neural network combined with the formula-based method presented above, was compared with a method using a neural network alone. The neural network learned the data well enough to show a high correlation coefficient, but the hybrid model showed a higher correlation coefficient and a lower mean error. Moreover, in a qualitative evaluation of the two methods, judging the methods by their success in predicting whether a student is in positive or negative mood, no matter how positive or how negative this mood is, the hybrid model was again more successful. This shows that the hybrid method is more effective.

The results of the present study are very promising indicating that the proposed method can form an effective tool for the task of mood recognition during an online test, based on student's goal and record during the test. The information obtained by the proposed mood recognition method during the self-assessment test, should be used to provide the student with adequate feedback. The affective feedback can take place before and after the test, during the test, and before and after a student's answer to a question. In all these cases affective feedback can be provided either routinely according to the student's affective state, either by the student's or the teacher's request. Affective feedback can be applied by using constructively positive emotions, while preventing, controlling and managing negative emotions. Moreover, the affective feedback can also take place by means of negative emotions in order to increase the student's devotion and engagement. Thus, it can be implemented using humour and jokes, amusing games, expressions of sympathy, reward, pleasant surprises, encouragement, acceptance, praises but also through criticism and punishment. Nevertheless, further research is needed to produce successful methods of affective feedback.

Also, the results confirmed that goals are central to mood regulation throughout a test for the reason that they provide the comparison point used during the appraisal process. It was made obvious that student's verification or rebuttal of hope to reach his/her goal configures his/her mood during a test. Furthermore, the results showed that it may be helpful to see human–computer interaction problems from a qualitative point of view rather than from a quantitative point of view. If we try to determine just whether a student is in a positive mood or in a negative mood, no matter how positive or how negative this mood is, things are getting much easier.

The neural network of stage 2 performed better than the neural network of stage 1, indicating that hybrid structures would be preferable for the construction of affect recognition systems. Using the formula-based method as an input parameter selection module for the neural network reduces the compute load and increases the generalization ability of the neural network mood recognizer. This indicates that neural networks and conventional algorithmic methods should not be in competition but complement each other for the development of affect recognition systems.

Moreover, both neural network methods succeeded in recognizing student's mood based on the same assumptions as the formula-based method. This is an extra verification that the chosen input parameters have an influence on student's mood while undertaking a test. The proposed neural network could be trained to perform even better. What is essential is that the presented work indicates that neural networks can provide a significant prediction of student's mood using personal preference information.

Acknowledgements

We thank the anonymous reviewers for their constructive comments. The authors also wish to thank the Alexander S. Onassis Public Benefit Foundation for the grant of a doctorate scholarship to the first author.

References

Alder, G. S. (2007). Examining the relationship between feedback and performance in a monitored environment: A clarification and extension of feedback intervention theory. *Journal of High Technology Management Research, 17*(2), 157–174.

Anderson, N. H. (1982). *Methods of information integration theory*. London: Academic Press.

Argyle, M. (1988). *Bodily communication* (2nd ed.). New York: Methuen.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293–1295.

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How emotion is made and measured. *International Journal of Human–Computer Studies, 65*, 275–291.

Bower, G. (1992). How might emotions affect learning? In C. Svenake, & E. Lawrence (Eds.), *Handbook of Emotion and Memory: Research and Theory*, Hilsdale, NJ.

Burleson, W., & Picard, R. W. (2004). Affective agents: Sustaining motivation to learn through failure and a state of 'stuck. In *Workshop on social and emotional intelligence in learning environments, seventh conference on intelligent tutoring systems*, Maceió, Brasil.

Caridakis, G., Karpouzis, & K., & Kollias, S. (2008). User and context adaptive neural networks for emotion recognition. *Neurocomputing, 17*(13–15), 2553–2562.

Conati, C. (2002). Probabilistic assessment of user's emotions in education games. *Journal of Applied Artificial Intelligence, 16*(7-8), 555–575 (special issue on managing cognition and Affect in HCI).

Cowie, R., Douglas-Cowie, E., & Cox, C. (2005). Beyond emotion archetypes: Databases for emotion modelling using neural networks. *Neural networks, 18*(4), 371–388.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., et al. (2001). Emotion recognition in human–computer interaction. *Signal Processing Magazine, IEEE, 18*(1), 32–80.

Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with autotutor. *Journal of Educational Media, 29*(3), 241–250.

Demuth, H., & Beale, M. (1995). *Neural network toolbox for use with MATLAB*. MA, USA: The Mathworks, Inc.

Demuth, H., Beale, M., & Hagan, M. (2005). *MATLAB user's guide, v 4.0.6: Neural network toolbox*. MA, USA: The MathWorks, Inc.

Economides, A. A. (2005). Personalized feedback in CAT (computer adaptive testing). *WSEAS Transactions on Advances in Engineering Education, 3*(2), 174–181.

Economides, A. A. (2006). Emotional feedback in CAT (computer adaptive testing). *International Journal of Instructional Technology and Distance Learning 3*(2). Available from: <http://itdl.org/Journal/Feb_06/article02.htm>.

Economides, A. A. (in press). Adaptive context-aware pervasive and ubiquitous learning. *International Journal of Technology Enhanced Learning*.

Ekman, P. (1982). *Emotion in the human face* (2nd ed.). Cambridge, MA: Cambridge University Press.

Fausett, L. (1994). *Fundamentals of neural networks*. Englewood Cliffs, NJ: Prentice Hall.

Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.

Guelich, S., Gundavaram, & S., Birznieks, G. (2000). *CGI Programming with Perl* (2nd ed.). O'Reilly.

Hadi, M. N. S. (2003). Neural networks applications in concrete structures. *Computers and Structures, 81*(6), 373–381.

Hone, K. (2005). Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interacting with Computers, 18*(2), 227–245.

Isbister, K., Höök, K., Sharp, M., & Laaksolahti, J. (2006). The sensual evaluation instrument: Developing an affective evaluation tool. In *Proceedings of human factors in computing systems* (pp. 1163–1172). ACM.

Kofler, M. (2005). *The definitive guide to MySQL 5* (3rd ed.). Apress.

Kort, B., Reilly, R., & Picard, R. W. (2001). External representation of learning process and domain knowledge: Affective state as a determinate of its structure and function. In *Proceedings of the artificial intelligence in education workshops (San Antonio, TX, AIED)* (pp. 64–69).

Larsen, RandyJ. (2000). Target articles: Toward a science of mood regulation. *Psychological Inquiry, 11*(3), 129–141.

Lawrence, S., Giles, C. L., & Tsoi, A. C. (1997). Lessons in neural network training: Overfitting may be harder than expected. In *Proceedings of the fourteenth national conference on artificial intelligence AAAI-97* (pp. 540–545). AAAI Press.

Leea, K., Boothb, D., & Alamc, P. (2005). A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Systems with Applications, 29*(1), 1–16.

Li, X., & Yu, W. (2002). Dynamic system identification via recurrent multilayer perceptrons. *Information Science, 146*(5), 45–63.

Lindgaard, G. (1995). Human performance in fault diagnosis: Can expert systems help. *Interacting with Computers, 7*(3), 254–272.

Lisetti, C. L., & Schiano, D. J. (2000). Automatic facial expression interpretation: Where human–computer interaction, artificial intelligence and cognitive science intersect. *Pragmatics and Cognition, 8*(1), 185–235 (special issue on facial information processing).

Mavrikis, M., Maciocia, A., & Lee, J. (2003). Targeting the affective state of students studying mathematics on a web-based ILE. In *Young researchers track of the 11th International Conference on Artificial Intelligence in Education (AIED)*.

Moridis, C. N., & Economides, A. A. (2008a). Towards computer-aided affective learning systems: A literature review. *Journal of Educational Computing Research, 39*(4), 313–337.

Moridis, C. N., & Economides, A. A. (2008b). A computer method for giving adequate feedback to students' current mood. *IEEE Multidisciplinary Engineering Education Magazine, 3*(3), 104–107.

Moridis, C. N. & Economides, A.A. (2008c). Modelling student's mood during an online self-assessment test. In M. D. Lytras, J. M. Carroll, E. Damiani, R. D. Tennyson, D. Avison, G. Vossen, & P. O. De Pablos (Eds.), *The open knowledge society: A computer science and information systems Mani-festo, Springer Commu-nications in Computer and Information Science* (Vol. 19, pp. 334–341).

Moridis, C. N., & Economides, A. A. (2009a). Classification models of students' moods during an online self-assessment test. *International Journal of Knowledge and Learning, 5*(1), 50–61.

Moridis, C. N., & Economides, A. A. (2009b). Mood recognition during online self-assessment test. *IEEE Transactions on Learning Technologies, 2*(1), 50–61.

Murphey, Y. L., & Luo, Y. (2002). Feature extraction for a multiple pattern classification neural network system. *IEEE International Conference on Pattern Recognition, 2*, 220–223.

Norman, D. A. (2002). Emotion and design: Attractive things work better. *Interactions Magazine, 9*(4), 36–42.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105.

Picard, R. (1997). *Affective computing*. MIT Press.

Picard, R., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., et al. (2004). Affective learning – A manifesto. *BT Technology Journal, 22*(4), 253–269.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533–536.

Schwarz, N., & Bless, H. (1991). Happy and mindless, but sad and smart? The impact of affective states on analytic reasoning. In J. P. Forgas (Ed.), *Emotion and social judgments* (pp. 55–720). New York: Pergamon.

White, H. (1989). Learning in artificial neural networks: A statistical perspective. *Neural Computation, 1*, 425–464.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*(1), 35–62.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663–676.

Please cite this article in press as: Moridis, C. N., & Economides, A. A. Prediction of student's mood during an online test using formula-based and neural network-based method. *Computers & Education* (2009), doi:10.1016/j.compedu.2009.04.002