Proc. of the 2nd International Conference on Electrical, Communication and Computer Engineering (ICECCE) 14-15 April 2020, Istanbul, Turkey

Detecting Education level using Facial Expressions in E-learning Systems

HabebaElRahman Hesham
Faculty Of Computer Science
Misr International University
Cairo, Egypt
habeba16085135@miuegypt.edu.eg

Mourad Nabawy
Faculty Of Computer Science
Misr International University
Cairo, Egypt
mourad1608721@miuegypt.edu.eg

Ola Safwat
Faculty Of Computer Science
Misr International University
Cairo, Egypt
Ola1606876@miuegypt.edu.eg

Youssef Khalifa
Faculty Of Computer Science
Misr International University
Cairo, Egypt
Youssef1603624@miuegypt.edu.eg

Haytham Metawie
Faculty Of Computer Science
Misr International University
Cairo, Egypt
haytham.metawie@miuegypt.edu.eg

Ammar Mohammed
Department of Computer Science
Misr International University
on leave from Cairo University
Cairo, Egypt
ammar.ammar@miuegypt.edu.eg

***Abstract*—With the growth rate of modern technologies, Computer-based learning environment receives attention for academic goals. In this environment, a computer provides learners with a set of learning contents divided into learning levels. Usually, Computer-based learning environment research efforts detect the next level of the learner automatically based on the correct responses of the learner on a test at the end of every learning level. Different efforts use fuzzy approaches to handle the uncertainly in the learning environment. In this paper, a machine learning approach is proposed to detect the current education level of the learner based on a recorded facial expressions of the learners as well as important features of the learning environment. Several classifiers are employed to recognize the education level. The evaluation of the proposed approach on a real dataset shows that Support Vector Machine (SVM) outperforms the other classifiers and achieves accuracy of 87%. The paper also presents a regression method to detect the learning level as a continuous value. The evaluation of the regression methods shows that the Linear Regression with mean squared error of 0.0048 outperform SVR.**

Index Terms*— *E-learning, Facial expressions, SVM, SVR, Decision tree, Linear regression.

I. INTRODUCTION

A computer environment becomes popular for academic purposes with the pace of growth in technologies. A programming system provides a learner in this environment with a collection of learning resources separated into learning stages. The computer-based learning system typically determines the next stage of the learner based on his responses to a placement test at the completion of the current phase of learning correctly. This way of detection does not usually obtain an accurate solution and it may affects the learner's educational level as a decision from the e-learning system because we ignored the psychological part. That why we need to take into consideration the facial expression of the learner during the learning activity. Many research papers were proposed to predict the learner's next stage. During the course test at the end of the study sessions, most efforts used the validity of the answer solely. Other measures in the learning process used fuzzy logic to manage the ambiguity of multiple parameters. The type of research includes the facial expressions of the learner to make the results of the placement test more accurate. That is the weight of the facial expressions through the learning process as a psychological point [1].

The environment of e-learning may be exposed a problem such as showing accurate decisions in upgrading or downgrading the learner in his learning process. Such a problem can be avoided by taking into consideration the psychological part of the learner through facial expressions. Machine learning is one of the popular approaches that learn from data and design a model that automatically classifies the proper actions. Machine learning can help in detecting the level of education of the learner and therefore can generate a model for the detection of specific decisions in this field. Therefore the main contribution of this paper is to propose a supervised approach of machine learning to identify a learner's level of education. The proposed method evaluated on a real data found in [2]. A machine learning technique is used to predict the learner's current level of education considering both the learner's facial expression and critical learning environment features such as the answer validity, current level, elapsed time and his facial expressions through the end of the current learning level. We evaluated the proposed model based on multiple metrics such as precision, accuracy, and recall. Using these types of metrics, we can provide a simple indicator of model generalization over training data. The experimental result shows that when compared with the other approaches the proposed machine learning model is efficient.

The content of the paper is arranged according to this. Section II presents many related works. Section III explains the overall approach. Section IV discusses the experimental

978-1-7281-7116-6/20/$31.00 © 2020 IEEE

results. Finally, the paper is concluded in section V.

II. RELATED WORKS

This section discusses the efforts of the computer based learning the environment. We classify the research effort into those research efforts that take the facial expression to detect the next level and those research effort that take fuzzy logic to detect the facial expressions.

A. Emotion recognition in E-learning

Many papers detect face and emotion with many algorithms: According to Krithika L.B and Lakshmi Priya GG [3], by pointing to realize the finest features in learning by improving the way of concentration throw understanding and adding the enthusiastic state of the learner during a learning engagement. They recognize and classify the feelings of the learners and give real-time feedback to improve and support the education level. Chasing the eyes and head movement is the best way to detect the face and emotion recognition. Therefore they classify learner's involvement within the subject by extracting three levels of concentration which are: High, Medium, Low. The authors utilized frequently of algorithms within the system such as Viola-Jones is used for solving the problems of features detection and emotion recognition to extract the specifics of the face region. Local Binary Pattern (LBP) is used for quick computation and is efficient to name the facial features. Ada Boost is an algorithm that takes visual characteristics and performs different cycles. Neural Network points to improve performance over a single network between several networks, in this manner avoiding complexity in selecting the non-facial object. By checking them on the understudy at that point;they get tests.

By utilizing the Viola-Jones algorithm to identify facial expression, the paper in [4] takes the expression of the students for improving the way of education for getting the finest results. By getting the expression of the student and provide the educators' feedback around the mood and his weighted enthusiastic state that he felt in this circumstance. The Viola-Jones algorithm is more likely in facial detection in terms of real-time operation and quick detection. There are six fundamental categorized feelings (Joy, Fear, Pity, Anger, Shock, and Disgust) and that's what the authors worked on. There are many several classification algorithms that offer assistance in recognizing the facial expression, they utilized the finest precision rates that were applied utilizing the K-nearest neighbors (kNN) algorithm. The authors utilized a lot of strategies within the proposed system, Student Emotion Recognition System (SERS), such as for image processing (Open CV Libraries) and for machine learning (sci-kit-learn, NumPy, pandas, etc.) processes. The authors catch one facial around the student and provide feedback to the instructor about the caught facial. This will cause inaccurate feedback around the teachers because of what could be back from expressions. Finally, we utilized another algorithm that can classify the feelings of the learners it was CNN that is different than they compared to other algorithms.

Chen et al [5] aimed to enhance the e-learning by capturing images of learners to detect their facial expressions. The authors employ many processes for detecting facial expressions such as Neural Networks (NN), SVM and Hidden Markov Models (HMM) along with Active Shape Model (ASM) [6] to discover face and position with the VOSM [7] method and Gabor wavelets to get facial appearance information. They identify facial expression through two classes: first one features extract from inactive pictures and the second one use pictures series in a current video. Another paper focuses on facial expression in e-learning which is Chih-Hung Wu [8], who implemented a facial expression recognition system (FER) to improve the performance of our e-learning systems. And facial expressions are able to supply critical proves about feelings [9]. Such system is similar to the previous system [5] by using SVM but authors used decision trees along with SVM. They managed a database called JAFFE to save faces that are detected by using FaceSDK.

In [10], the authors' opinion is that facail expression and emotion recognition have a very vital part in intelligent e-learning systems. The authors made recent studies about the related work [11] [12] [13] [14] have shown that emotions and expressions have an essential role in the learner learning processes. This paper uses a Convolutional Neural Network (CNN) to create a system that detects facial expression recognition with the target to utilize it in the e-learning system. They used two types of datasets CK+ [15] [16] and KDEF [17] to train and test the system and in addition to these datasets, they use another dataset similar to paper [5] and paper [8] and this dataset called Japanese female facial expression. They use Open CV as a preprocessing step to detect the human face and edit it. At that time they implement the CNN algorithm as a second step to extract the function. Finally, they classify the network's full-linked facial emotion.

B. Fuzzy interference system

Similar work is proposed in [18], the authors presented an adaptive learning environment using the facial expression and fuzzy logic. In their work, the CNN is used to detect the learner facial expression and is trained using FER2013 [19] and used OpenCV [10] for feature extraction. In their proposed work, the learnt features of the learner is fed to a fuzzy inference system for detecting the next learning level. However, the evaluation of the systems did not take machine learning into consideration, the authors used a fuzzy inference system using domain expert to detect the next level.

According to Ramón Zatarain-Cabada in his research paper [20] discussing the usage of the fuzzy logic engine, emotional facial and text. The system starts with testing the student for his placement test to define the starting level. Exercises variables (exercise validation, amount of compilations and time required) are the main core input for taking a decision with emotional state variables (facial and text emotion). By taking only one image about the emotional state of the student and text emotion with the exercise variables. The system uses a fuzzy logic engine to achieve a new user-level by merging the present level, variables of exercise and variables of emotional state. Every time the student completes an exercise. The process to get a new level is performed. The emotion recognition method works with two recognizers, where the first is a method for recognizing emotion through facial recognition. Throughout this process, when he/she solves the exercise the system capture a photo of the student. For the face, authors use a feature extractor which was implemented with OpenCV library along with neural network implemented to identify emotions. The emotions which the face recognizer identifies are joy, surprise, sadness, anger, and neutral emotion. The fuzzy rules output is the next level of the student's level (beginner, basic, intermediate, and advanced) and by using the java library named JFuzzyLogic for implementing the fuzzy system and semantic algorithm for the text emotion. The facial expression is able to distinguish emotions with a realization rate of 80 percentages as good results. The emotion recognizer based on text dialogues achieved a realization rate of 85 percentages.

In [21], it was proposed that e-learning applications could benefit from such emotion detection tools for more realistic interactions, as they capture learner's data constantly and hidden [22]. It was difficult to detect face and facial expressions correctly while distracting light shines directly into the learner's face [23]. The method is based on fuzzy logic, using the induction of unordered fuzzy rule (FURIA algorithm [24]). The fuzzy logic method uses the supervised machine learning method to supply more desirable performance because fuzzy logic rules can easily be created from a dataset of recorded emotions, whereas alternative machine learning methods such as Neural Networks, Bayesian Network, and Decision Tree will need wider implementation [25]. Started from an existing database (CK+). Based on this, they next created a database of emotions including the rotated images, they then created cosine values of facial landmarks for training and testing purposes then used fuzzy rules. They developed a small app that used Dilb [26] (open source library aims to supply a similarly rich environment for the development of python language machine learning applications and methods for classification, regression, clustering, detection of anomalies). They extracted 68 facial landmarks from each image and made 54 vertices in their database for 18 related triangles, using every three significant landmarks. Then, they stored all the cosine values in the form of a WEKA attribute-relation file format (arff) [27] along with the relevant emotion labels of each image of the CK+. The arff file is a textual database that specifies a list of instances share a set of attributes: each instance is defined with 55 attributes, respectively called Cosine0, Cosine1, Cosine53 and Emotion. By loading the database into WEKA 37 they could create so-called FURIA fuzzy rules, allowing them to automatically detect and identify emotions from facial expressions. This study achieves an average accuracy of 83.2 percentages which is equivalent to human performance [28] [29]. In fact, it can identify multiple faces in an image simultaneously. A number of game engines and e-learning environments can be quickly ported.

III. OVERALL APPROACH

We present a machine learning approach to classify the learner features and eventually, integrate the classifiers into a proposed systems. The proposed system that comes out of this paper, is shown in Fig. 1. The proposed system adopts the idea and learning modules of the work found in [18]. In the adopted work, a facial expression module, based on a pre-trained CNN model, is used to detect the facial expressions. However, different to the work of[18], our proposed system system detects the learning level after feeding the learning parameters to a machine learning classifiers.

In the proposed system, the prediction level module of the machine learning technique based on four fundamental features based on the trained data included in [2]. These important features are listed as follows:

-Answer Validity : which represents the ratio of the total right answers entered the exam checker to get it as a feature for the classifier.

-Current level : which represents the current level of the learner through placement test for detecting his level for setting the right materials.

-Test Elapsed Time : which system counts the time of the test to get the accurate time that he/she completes the test.

-Facial Expressions : which system captures through the whole exam.

Also, the dataset contains the aggregated emotional states that have been detected from the learner's facial expressions during exams.

By merging the previous features,the system extracts and adding them to the dataset to be classified by the SVM classifier. The reason behind choosing SVM classifier is that SVM achieved the best classification results through experiments as we show in the next section.

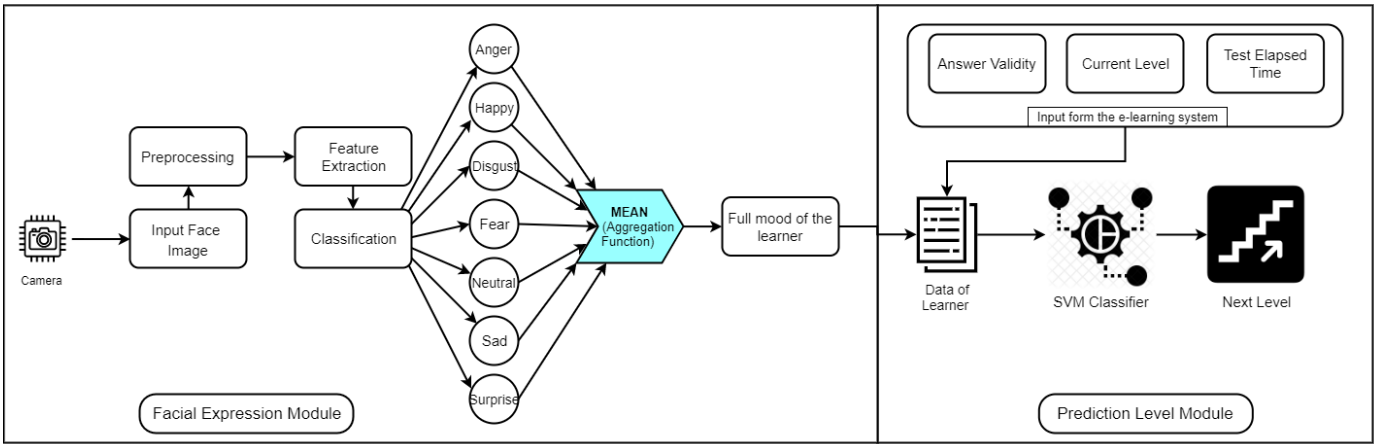

Fig. 1. The process of proposed system

In general, classifier's evaluation is performed using multiple classification matrices. Accuracy, precision and recall are among these requirements. They are determined according to the equations below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Where TP is the number of true positive, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

IV. EXPERIMENTAL RESULTS

This section describes the experimental results of the proposed approach. By evaluating the proposed approach on the dataset in [2] to classify the next desired learning levels of the learner. Firstly we describe the dataset used in the training of the machine learning approach. Then, we discuss the evaluation of the results on two different machine learning techniques namely classification and regression.

A. Data Description

The dataset is taken from [2] and is formed of 1753 instance data of different emotional states of several learners of different skills. The data contains 6 different levels of difficulties called Starter, Elementary, Pre-intermediate, Intermediate, Upper-intermediate, and Advanced with corresponding ranges [0, 0.15], [0.15, 0.35], [0.35, 0.55], [0.55, 0.75], [0.75, 0.95], [0.95, 1]. Table I summarizes the dataset. This dataset in addition contains some features as desired learning levels as shown in Table II. This dataset contains in general 71 records of learning levels and contains features including Answer Validity, Test Elapsed Time, Seven Facial Expressions, Current Level, and labeled with Desired Level. The labeled class is an unbalanced distributed of Desired Level which has 6 different levels of difficulties it was continuous values then we used the discretization method to manipulate those values with interval numerical ranges, also removed from the dataset the unused features.

TABLE I: Summary of the data [2]

The Trained Data				
Individuals	Instances of learning activities	Minimum learning activities per each individual	Maximum learning activities per each individual	Instances of emotional states
12	72	4	9	1735

TABLE II: Statistics of the Desired Level [2]

Total Numbers of each levels						
Desired Level	Starter	Elem.	Pre-INTER	INTER	Upper-INTER	Advanced
71	16	25	27	3	0	0

B. Results

To experiment the data, we perform two different experiments. The first one is to apply classification technique to predicate the class of next learning activities. The second experiment is to apply regression approach to predict a continuous value representing the intended learning level.

After preprocessing the dataset, as step presented in Section IV to extract the features, the dataset is spilt into training and testing using the ratios of 0.8 to 0.2 respectively for the classification purpose. The first experiment goal is to choose a classifier to develop a machine learning model to classify the next learning level. The Accuracy, recall, and precision is taken as a performance measure to evaluate the classifiers. By applying SVM as well as DecisionTreeClassifier and KNN as they are among the best performance classifiers in the literature. Figure 2 illustrates the accuracy of both KNN, DecisionTreeClassifier, and SVM. The SVM classifier achieved the highest accuracy percentage of **87%** while DecisionTreeClassifier 73% and KNN 53%. In addition to accuracy, precision-recall is a effective measure of success of prediction when the classes are very imbalanced, Figure 2 shows the evaluations of those classifiers in terms of precision and recall respectively for each classification model. Also, Figure 2 represents the precision of KNN, DecisionTreeClassifier, and SVM. The SVM classifier achieved the highest percentage of **87%** while DecisionTreeClassifier 78% and KNN 50%. Finally, Figure 2 represents the recall of KNN, DecisionTreeClassifier, and SVM. The SVM classifier achieved the highest percentage of **87%** while DecisionTreeClassifier 73% and KNN 53%.

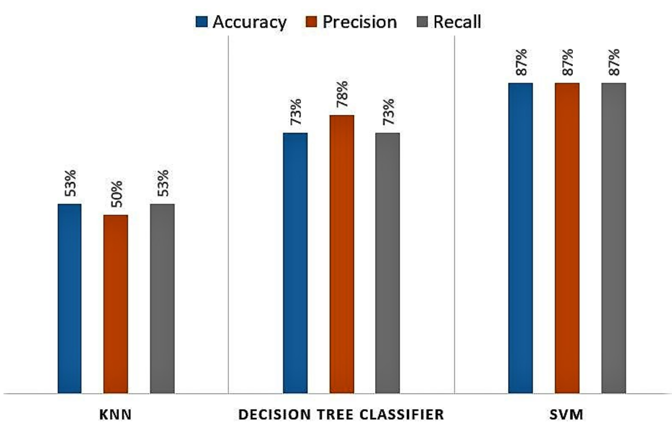

Fig. 2. Describe the Accuracy, Precision, Recall of KNN, Decision Tree and SVM in different Classification Model

The second experiment goal is to choose a regression model to create a machine learning model to predict the next learning level of the learner without discretization of the desired level column. We applied two regression models namely Linear Regression, and Support Vector Regression to compare between them in terms of mean squared error. The mean squared error of the two regression models is as shown in Table III. The Linear Regression achieved the best mean squared error = **0.0047480174** while Support Vector Regression (SVR) = 0.0103258427.

TABLE III: Mean squared error of KNN, DecisionTree and SVM

Regression	Mean Squared Error
Support Vector Regression (SVR)	0.0103258427
Linear Regression	**0.0047480174**

V. Conclusion

Nowadays, E-learning faces a very important problem that is showing accurate decisions in upgrading or downgrading the learner in his learning process. To prevent this problem we proposed a machine learning approach to detect the next level of the learner in e-learning environment. In particular, the paper proposed two different machine learning techniques namely classification and regression. First, several classifiers have been tested using a real dataset of previous learners. The classification accuracy has shown that SVM with accuracy **87%** accuracy outperformed the other classifiers.The regression approach has shown that the linear regression outperforms SVR in terms of mean-square error. The Classifier is inter grated into a system that is used for future predication of the learner.

References

[1] Alice M Isen. "Some ways in which positive affect influences decision making and problem solving". *Handbook of emotions* 3 (2008), pp. 548–573.

[2] Mohammed Megahed, Ahmed Asad, and Ammar Mohammed. "Data on learners emotional states, mental responses and fuzzy learning flows during interaction with learning environment". *Data in brief* 25 (2019), p. 104378.

[3] LB Krithika and Lakshmi Priya GG. "Student emotion recognition system (SERS) for e-learning improvement based on learner concentration metric". *Procedia Computer Science* 85 (2016), pp. 767–776.

[4] Paul Viola and Michael Jones. "Rapid object detection using a boosted cascade of simple features". *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*. Vol. 1. IEEE. 2001, pp. I–I.

[5] Liyuan Chen, Changjun Zhou, and Liping Shen. "Facial expression recognition based on SVM in E-learning". *Ieri Procedia* 2 (2012), pp. 781–787.

[6] Timothy F Cootes, Cristopher J Taylor, et al. *Statistical models of appearance for computer vision*. 2004.

[7] P Jia. "2D statistical models". *Technical Report of Vision Open Working Group* (2010).

[8] Chih-Hung Wu. "New technology for developing facial expression recognition in e-learning". *2016 Portland International Conference on Management of Engineering and Technology (PICMET)*. IEEE. 2016, pp. 1719–1722.

[9] Carlos Busso et al. "Analysis of emotion recognition using facial expressions, speech and multimodal information". *Proceedings of the 6th international conference on Multimodal interfaces*. 2004, pp. 205–211.

[10] Oussama El Hammoumi et al. “Emotion Recognition in E-learning Systems”. *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE. 2018, pp. 1–6.

[11] Michalis Feidakis, Thanasis Daradoumis, and Santi Caballé. “Endowing e-learning systems with emotion awareness”. *2011 Third International Conference on Intelligent Networking and Collaborative Systems*. IEEE. 2011, pp. 68–75.

[12] Jason M Harley et al. “An integrated emotion-aware framework for intelligent tutoring systems”. *International Conference on Artificial Intelligence in Education*. Springer. 2015, pp. 616–619.

[13] FatimaEzzahra Benmarrakchi et al. “Exploring the use of the ICT in supporting dyslexic students’ preferred learning styles: A preliminary evaluation”. *Education and Information Technologies* 22.6 (2017), pp. 2939–2957.

[14] GH Bower. “Mood and Memory American Psychologist 36: 129–148”. *CrossRef Google Sch.* (1981).

[15] Takeo Kanade, Jeffrey F Cohn, and Yingli Tian. “Comprehensive database for facial expression analysis”. *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*. IEEE. 2000, pp. 46–53.

[16] Patrick Lucey et al. “The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression”. *2010 ieee computer society conference on computer vision and pattern recognition-workshops*. IEEE. 2010, pp. 94–101.

[17] Daniel Lundqvist, Anders Flykt, and Arne Öhman. “The Karolinska directed emotional faces (KDEF)”. *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet* 91.630 (1998), pp. 2–2.

[18] Mohammed Megahed and Ammar Mohammed. “Modeling Adaptive E-Learning Environment using Facial Expressions and Fuzzy Logic”. *Expert Systems with Applications* (2020), p. 113460. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2020.113460.

[19] Panagiotis Giannopoulos, Isidoros Perikos, and Ioannis Hatzilygeroudis. “Deep learning approaches for facial emotion recognition: A case study on fer-2013”. *Advances in Hybridization of Intelligent Methods*. Springer, 2018, pp. 1–16.

[20] Ramón Zatarain-Cabada et al. “Java Tutoring System with Facial and Text Emotion Recognition.” *Research in Computing Science* 106 (2015), pp. 49–58.

[21] Hasimah Ali et al. “Facial emotion recognition using empirical mode decomposition”. *Expert Systems with Applications* 42.3 (2015), pp. 1261–1277.

[22] Ivon Arroyo et al. “Emotion sensors go to school.” *AIED*. Vol. 200. Citeseer. 2009, pp. 17–24.

[23] Carmen Bachiller, Carlos Hernández, Jorge Sastre, et al. “Collaborative learning, research and science promotion in a multidisciplinary scenario: information and communications technology and music”. *Proceedings of the International Conference on Engineering Education*. 2010, pp. 1–8.

[24] Kiavash Bahreini and Atilla Elci. “A New Software Architecture for J2EE Enterprise Environments via Semantic Access to Web Sources for Web Mining by Distributed Intelligent Software Agents”. *2008 32nd Annual IEEE International Computer Software and Applications Conference*. IEEE. 2008, pp. 902–907.

[25] Kiavash Bahreini and Atilla Elci. “SDISSASA: A multiagent-Based web mining via semantic access to Web resources in Enterprise Architecture”. *2008 32nd Annual IEEE International Computer Software and Applications Conference*. IEEE. 2008, pp. 553–558.

[26] Mohamed Ben Ammar et al. “The affective tutoring system”. *Expert Systems with Applications* 37.4 (2010), pp. 3013–3023.

[27] Kiavash Bahreini, Rob Nadolski, and Wim Westera. “Towards multimodal emotion recognition in e-learning environments”. *Interactive Learning Environments* 24.3 (2016), pp. 590–605.

[28] Kiavash Bahreini, Rob Nadolski, and Wim Westera. “Data fusion for real-time multimodal emotion recognition through webcams and microphones in e-learning”. *International Journal of Human–Computer Interaction* 32.5 (2016), pp. 415–430.

[29] Kiavash Bahreini, Rob Nadolski, and Wim Westera. “Communication skills training exploiting multimodal emotion recognition”. *Interactive Learning Environments* 25.8 (2017), pp. 1065–1082.