Informatics in Education, 2016, Vol. 15, No. 1, 147–161

DOI: 10.15388/infedu.2016.08

Affect Recognition through Facebook for Effective Group Profiling Towards Personalized Instruction

Christos TROUSSAS[1], Kurt Junshean ESPINOSA[2], Maria VIRVOU[1]

[1] *Department of Informatics, University of Piraeus*
Piraeus, Greece
[2] *Department of Computer Science, University of the Philippines*
Cebu City, Philippines
e-mail: ctrouss@unipi.gr, kpespinosa@up.edu.ph, mvirvou@unipi.gr

Received: July 2015

Abstract. Social networks are progressively being considered as an intense thought for learning. Particularly in the research area of Intelligent Tutoring Systems, they can create intuitive, versatile and customized e-learning systems which can advance the learning process by revealing the capacities and shortcomings of every learner and by customizing the correspondence by group profiling. In this paper, the primary idea is the affect recognition as an estimation of the group profiling process, given that the fact of knowing how individuals feel about specific points can be viewed as imperative for the improvement of the tutoring process. As a testbed for our research, we have built up a prototype system for recognizing the emotions of Facebook users. Users' emotions can be neutral, positive or negative. A feeling is frequently presented in unpretentious or complex ways in a status. On top of that, data assembled from Facebook regularly contain a considerable measure of noise. Indeed, the task of automatic affect recognition in online texts turns out to be more troublesome. Thus, a probabilistic approach of Rocchio classifier is utilized so that the learning process is assisted. Conclusively, the conducted experiments confirmed the usefulness of the described approach.

Keywords: affect recognition, facebook, intelligent tutoring systems, rocchio classifier, user classification.

1. Introduction

Social networks seem to be a popular trend in modern life and a very important means of interactivity among people of different cultures. When people interact with peers, they can take advantage of crucial characteristics of social networks, such as directness and ease. Socialization has important pedagogical implications in learning by supporting the learners' personal relationships and social interaction with their classmates (Troussas *et al.*, 2014). In this way, using social networks in instructional contexts can be considered as a potentially powerful idea simply because students spend anyway a lot of their spare time on these online networking activities (Troussas *et al.*, 2013).

Social networks can play a crucial role in education and especially is the area of Intelligent Tutoring Systems (ITSs) which can produce adaptive and individualized e-learning systems. Indeed, adaptive individualized e-learning systems could enhance the educational procedure by offering a student-centered environment of learning and by prioritizing student's needs (Troussas *et al.*, 2015). Individualization is based on a student models which are fundamental to the architecture of ITSs.

One important area of ITSs specializes in language learning which is referred to as Intelligent Computer-Assisted Language Learning (ICALL). In ICALL, students are being taught a language (e.g. English) through an ITS. When an ITS is incorporated in social networks, the need of group profiling emerges so that the collaboration among users is further promoted. One crucial value for group profiling is the affect recognition of the user.

Few studies on affect recognition in social networks have already been presented (Agrawal *et al.*, 2011). These studies are mainly targeted to Twitter, for tweet updates about a specific topic (Agrawal *et al.*, 2011). What people express through their status updates is sometimes neutral, but also some of them express a particular emotion.

On the other hand, intelligent tutoring systems in social networks can benefit from understanding the emotions of social network users. Positive emotions can facilitate learning and negative ones can be an obstacle for it. Therefore, it is helpful that by these natural avenues of emotional expression, intelligent tutoring systems can also have the facility to adapt to their users so as to help them in learning new concepts.

Given that social networks are now natural avenues where people express their thoughts and opinions about their everyday life, affect recognition emerges interest. Towards this direction, automated opinion mining can be used in such circumstances. Automated opinion mining is a type of natural language processing using machine learning for tracking the mood of users and involves collecting and examining opinions about the status. Textual emotion analysis is a sub-field of automated opinion mining that has attracted growing interest from researchers who would like to know whether a particular text expresses a positive or negative emotion.

The idea for this research work came from the need of affect recognition in education. Emotion is important in education as it drives attention, which in turn drives learning and memory. Emotion matters aim to increase understanding and awareness of the psycho-social aspects of living with a long term condition and to provide skills that will enable more holistic, collaborative and person-centered learning.

In view of the above, this paper seeks to investigate the relationship of the usefulness of affect recognition in a Facebook intelligent language learning application as a value in group profiling. For this reason, we have developed a system that is able to classify a status using sentence-level classification whether it entails positive, negative or neutral emotions by using a more probabilistic approach of Rocchio algorithm. Opinions are in the form of status updates in Facebook. The specific objectives of our study are to properly train the system to accept inputs in the form of status updates, disregarding updates that do not contain words or face emoticons and to classify the polarity of an opinion per status update basis.

2. Related Scientific Work

Affect recognition has been handled as a Natural Language Processing task. Starting from being a document level classification task, it has been handled at the sentence level and more recently at the phrase level. In this section, we present the related scientific work, firstly related to Grouping of students and secondly to affect recognition.

2.1. *Literature on Students' Grouping in tutoring systems*

In (Basile *et al.*, 2011), the authors proposed the exploitation of machine learning techniques to improve and adapt the set of user model stereotypes by making use of user log interactions with the system. To do this, a clustering technique is exploited to create a set of user models prototypes; then, an induction module is run on these aggregated classes in order to improve a set of rules aimed as classifying new and unseen users. Their approach exploited the knowledge extracted by the analysis of log interaction data without requiring an explicit feedback from the user.

In (Nino, 2009), the author presented a snapshot of what has been investigated in terms of the relationship between machine translation (MT) and foreign language (FL) teaching and learning. Moreover, the author outlined some of the implications of the use of MT and of free online MT for FL learning.

In (Friaz-Martinez *et al.*, 2007), the authors investigated which human factors are responsible for the behavior and the stereotypes of digital libraries users so that these human factors can be justified to be considered for personalization. To achieve this aim, the authors have studied if there is a statistical significance between the stereotypes created by robust clustering and each human factor, including cognitive styles, levels of expertise and gender differences.

In (Licchelli *et al.*, 2004), the authors focused on machine learning approaches for inducing student profiles, based on Inductive Logic Programming and on methods using numeric algorithms, to be exploited in this environment. Moreover, an experimental session has been carried out from the authors, comparing the effectiveness of these methods along with an evaluation of their efficiency in order to decide how to best exploit them in the induction of student profiles.

In (Shi and Sha, 2012), the authors studied the problem of unsupervised domain adaptation, which aims to adapt classifiers trained on a labeled source domain to an unlabeled target domain, since many existing approaches first learn domain-invariant features and then construct classifiers with them. They propose a novel approach that jointly learn the both.

In (Vihn *et al.*, 2010), the authors presented an organized study of information theoretic measures for clustering comparison. They have shown that the normalized information distance (NID) and normalized variation of information (NVI) satisfy both the normalization and the metric properties. Between the two, the NID is preferable since the tighter upper bound of the MI used for normalization allows it to better use the [0,1] range. They highlighted the importance of correcting these measures for chance agreement, especially when the number of data points is relatively small compared with the number of clusters.

In (Palubinskas *et al.*, 1998), the authors proposed to embed the clustering problem into a Bayesian framework to automatically detect the number of clusters. The entropy is considered to define a prior and enables them to overcome the problem of defining a priori the number of clusters and an initialization of their centers. A deterministic algorithm derived from the standard k-means algorithm was proposed and compared with simulated annealing algorithms.

In (Troussas and Virvou, 2013), the authors proposed a novel approach of information theoretic clustering, based on entropy. Their approach generalizes the standard Euclidean distance, used in k-means clustering algorithm, by admitting arbitrary linear scaling and rotations of the feature space and models the problem in an information-theoretic setting. In this way, qualitative collaboration among students of the same cluster is achieved, so that they are capable of succeeding in multiple language learning, namely in the learning of the English and French language.

2.2. *Literature on Affect Recognition*

In (Boiy *et al.*, 2007), the authors provided a good survey of various techniques developed in online sentiment analysis. It covers concept of emotion in written text (appraisal theory), various methodologies which can be broadly divided into two groups: (i) symbolic techniques that focuses on the force and direction of individual words (the so-called “bag-of words” approach), and (ii) machine learning techniques that characterizes vocabularies in context. Based on the survey, the authors found that symbolic techniques achieves accuracy lower than 80% and are generally poorer than machine learning methods on movie review sentiment analysis.

Another significant effort for sentiment classification on Twitter data is conducted by (Barbosa and Feng, 2010). The authors use polarity predictions from three websites as noisy labels to train a model and use 1000 manually labeled tweets for tuning and another 1000 manually labeled tweets for testing. They however do not mention how they collect their test data. They propose the use of syntax features of tweets like retweet, hash tags, link, punctuation and exclamation marks in conjunction with features like prior polarity of words and POS of words.

In (Gamon, 2004), the authors perform sentiment analysis on feedback data from Global Support Services survey. One aim of their study is to analyze the role of linguistic features like POS tags. They perform extensive feature analysis and feature selection and demonstrate that abstract linguistic analysis features contributes to the classifier accuracy.

In (Go *et al.*, 2009), the authors use distant learning to acquire sentiment data. They use tweets ending in positive emoticons like “:)” “:-)” as positive and negative emoticons like “:(” “:-(” as negative. They build models using Naive Bayes, MaxEnt and Support Vector Machines (SVM). In terms of feature space, they try a Unigram, Bigram model in conjunction with parts-of-speech (POS) features. They note that the unigram model outperforms all other models. Specifically, bigrams and POS features do not help.

In (Pak and Paroubak, 2010), the authors take a naive approach to collect and classify 300000 tweets into three categories: (i) tweets queried with emoticon queries such as ":-)", ":)", "=)" indicate happiness and positive emotion (ii) tweets with ":-(", ":(", "=(", ";(" implies dislike or negative opinions, and (iii) tweets posted by newspaper accounts such as "New York Times" are considered objective or neutral.

In (Pang and Lee, 2004), the authors applied minimum cuts in graphs to extract the subjective portion of texts they were studying and used machine learning methods to perform sentiment analysis on those snippets of texts only.

In (Mullen and Collier, 2004), the authors discussed the application of support vector machines in sentiment analysis with diverse information source.

In (Godbole *et al.*, 2007), the authors developed techniques that algorithmically identify large number (hundreds) of adjectives, each with an assigned score of polarity, from around a dozen of seed adjectives. Their methods expand two clusters of adjectives (positive and negative word groups) by recursively querying the synonyms and antonyms from WordNet. Since recursive search quickly connects words from the two clusters, they implemented several precaution measures such as assigning weights which decrease exponentially as the number of hops increases. This confirms that the algorithm-generated adjectives are highly accurate by comparing them to the results of manually picked word lists. It is worth pointing out that this work uses Lydia as the backbone to process large amount of news and blogs.

In (Wilson *et al.*, 2005), the authors discussed categorizing texts into polar and neutral first before determining whether a positive or negative sentiment is expressed through the text. However, in (Godbole *et al.*, 2007), the authors operate on the premise that little neutrality exists in online texts.

However, after a thorough investigation in the related scientific literature, we came up with the result that there is not any research describing affect recognition for the amelioration of an intelligent language learning system in Facebook using the Rocchio classifier. Moreover, the data used for training and testing are collected by search queries and is therefore biased. In contrast, we present features achieving a significant gain over a unigram baseline. Our data are a random sample of streaming Facebook status unlike data collected by using specific queries. The size of our hand-labeled data allows us to perform cross validation experiments and check for the variance in performance of the classifier across folds.

3. Methodology And Architecture

The main methodology for Affect Recognition is the Classifier method specifically the Rocchio Classifier where in a status update is being classified as positive or negative. Fig. 1 shows the overview of affect recognition using Rocchio Classifier.

In this section, we will present an analysis of the Rocchio classifier, which gives theoretical insight into the heuristics used in it, and particularly the word weighting scheme and the similarity metric. We also suggest improvements which lead to a probabilistic variant of the Rocchio classifier.

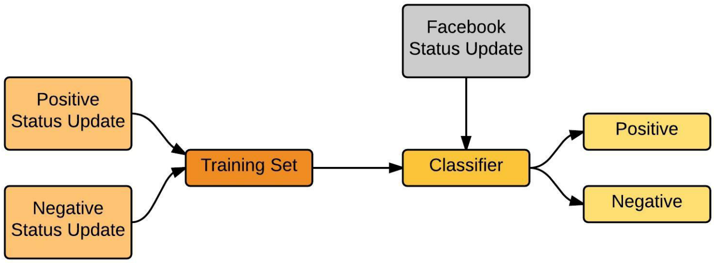

Fig. 1. Main methodology of Rocchio Classifier.

Text categorization is the procedure of clustering documents (and hence Facebook statuses) into different categories or classes. With the amount of online educational systems in Facebook growing rapidly, the need for reliable text categorization of users' statuses has increased.

One of the most widely applied learning algorithms for text categorization is the Rocchio algorithm. Although the algorithm is intuitive, it has a number of problems which lead to comparably low classification accuracy (Joachims, 1997):

a. The objective of the Rocchio algorithm is to maximize a particular functional. Nevertheless, Rocchio does not show why maximizing this functional should lead to a high classification accuracy.
b. Heuristic components of the algorithm offer many design choices and there is little guidance when applying this algorithm to a new domain.
c. The algorithm was developed and optimized for relevance feedback in information retrieval; it is not clear which heuristics will work best for text categorization.

The major heuristic component of the Rocchio algorithm is the TFIDF (term frequency / inverse document frequency) word weighting scheme (Joachims, 1997). Different flavors of this heuristic lead to a multitude of different algorithmic approaches. If Rocchio uses probabilistic models for classification, it can allow the explicit statement of simplifying assumptions.

Because of its heuristic components, there is a number of characteristics promoting probability which are the word weighting method, the document length normalization using Euclidian vector length and the similarity measure (cosine similarity) (Joachims, 1997).

The algorithm returns a ranking of documents to define a decision rule for class membership and therefore the algorithm has to be adapted to be used for text categorization. The variant seems to be the most straightforward adaptation of the Rocchio algorithm to text categorization and domains with more than two categories. The algorithm builds on the following representation of text. Each text d is represented as a vector so that texts with similar content have similar vectors (according to a fixed similarity metric) and each element represents a distinct for a document (Joachims, 1997). The term frequency is the number of times a word is found in document and the document frequency is the number of documents in which word is found at least once. The inverse document frequency is proportionate to the document frequency. Intuitively, the inverse document frequency of a word is high if the words occurs in only one document and lower if it occurs in many documents. A word is an important indexing term for document if it is found frequently in it (the term frequency is high).

On the other hand, words which are found in many documents are rated less important indexing terms due to their low inverse document frequency. Learning is achieved by combining document vectors into a prototype vector for each class. First, both the normalized document vectors of the positive and negative examples for a class are summed up. The prototype vector is then calculated as a weighted difference of each. Using the cosine as a similarity metric, Rocchio shows that each prototype vector maximizes the mean similarity of the positive training examples with the prototype vector minus the mean similarity of the negative training examples with the prototype vector. The resulting set of prototype vectors, one vector for each class, represents the learned model. This model can be used to classify a new document. Again the document is represented as a vector using the scheme described above (Joachims, 1997).

In this way, we are working with conditional probabilities that allow us to flip the condition around conveniently. A conditional probability is a probability that event X will occur, given the evidence Y. That is normally written $P(X \mid Y)$. Thus, we can determine this probability when all we have is the probability of the opposite result and of the two components individually: $P(X \mid Y) = P(X) P(Y \mid X) / P(Y)$.

In this case, we are estimating the probability that a text is positive or negative, given its contents. We can restate that, so that is in terms of the probability of that text occurring if it has been predetermined to be positive or negative. This is convenient, because we have examples of positive and negative opinions from our data set.

The underlying idea is that we make a large assumption about how we can calculate the probability of the document occurring. We can estimate the probability of a word occurring, given a positive or negative emotion by looking through a series of examples of positive and negative emotions and counting how often it occurs in each class. This is what makes this supervised learning, the requirement for pre-classified examples to train on.

3.1. *Creation of Corpus*

Corpus consists of the collection of writings or recorded remarks used for linguistic analysis. In this Facebook application, recorded remarks are classified into groups of negative and positive feelings in Facebook users' status. Range of 5000 – 10000 status updates will be the targeted number for corpus. It will be divided for two classes, negative and positive. Corpus should be large in number and for this reason the number of 5000 data appears to provide very satisfactory results.

Data will be collected from Facebook users based on the records in it. The system will be trained on the emotions of users, to whom our Facebook language learning application is addressed. The collected data will be manually identified whether they are positive or negative. Positive and negative status updates will then be stored in a class.

3.2. *States Classification*

A conditional probability is a probability that event X will occur, given the evidence. Hence, our initial formula has the following rationale:

$$P(\text{emotion} \mid \text{sentence}) = P(\text{emotion})\ P(\text{sentence} \mid \text{emotion}) \,/\, P(\text{sentence}) \quad (1)$$

We can drop the dividing P(line), as it's the same for both classes. The need is to rank them rather than calculate a precise probability. We can use the independence assumption to let us treat P(sentence | emotion) as the product of P(token | emotion) across all the tokens in the sentence. So, we estimate P(token | emotion) as:

$$\text{count(this token in class)} + 1 \,/\, \text{count(all tokens in class)} + \text{count(all tokens)} \quad (2)$$

The extra 1 and count of all tokens stops a zero finding its way into the multiplications. If there was not any sentence with an unseen token in, it would score zero.

The classify function starts by calculating the prior probability (the chance of it being one or the other before any tokens are looked at) based on the number of positive and negative examples; in our case, that will always be 0.5, as for each observation (positive / negative status update), there are the same amount of data. We then tokenize the incoming document and for each class multiply together the likelihood of each word being seen in that class. We sort the final result and return the highest scoring class.

Our research classifies the polarity of the status update in a sentence level. Sentence level, in most cases, is more accurate than the phrase level because every status update has its own style in addressing users' emotion.

Fig. 2 illustrates two screenshots of the Facebook educational application. At the left side, there is the log-in page of the Facebook learning application and at the right side there is the recommendation for student collaboration based on the group profiling using their characteristics (including emotional state) which are presented at Section IV.

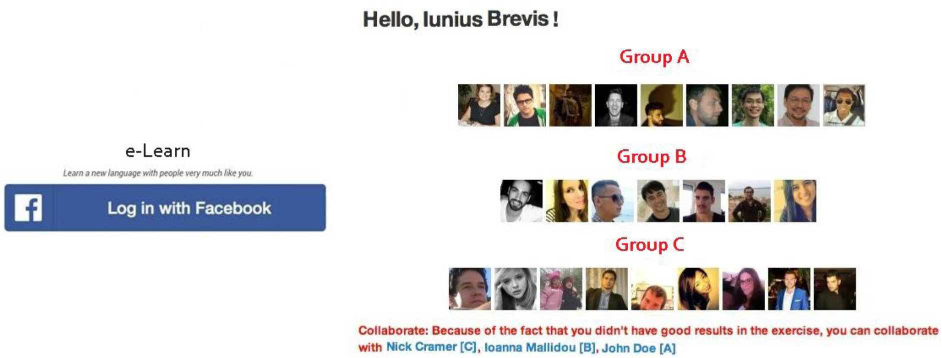

Fig. 2. Screenshots of the application.

4. Affect Recognition In Intelligent Language Tutoring

Emotions are complex states of mind and body. Cognitively, individuals interpret an event as one that may be sad or happy. Behaviorally, a student may seek comfort when s/he is sad and seek help when s/he faces danger. Our emotional state has the potential to influence our thinking (Darling-Hammond *et al.*, 2003). For instance, students learn and perform more successfully when they feel secure and happy about the subject matter (Oatley and Nundy, 1996). Although emotions have the potential to energize students' thinking, emotional states also have the potential to interfere with learning. If students are overly excited or enthusiastic, they might work carelessly or quickly rather than working methodically or carefully (Darling-Hammond *et al.*, 2003).

Moreover, negative emotions have the potential to distract students' learning efforts by interfering with their ability to involve in the educational process successfully. Emotions can interfere with students' learning in several ways, including limiting the capacity to balance emotional issues with tutoring. Some students might need one-on-one time with their peers, which can be achieved by instant or asynchronous text messaging in Facebook, in order to help the process of their feelings or the resolution of a problem.

Towards the efficient creation of user clusters, we incorporate algorithmic approaches into the resulting Facebook intelligent multi-language learning application which receive as input, pre-stored data or data from empirical studies, either directly by asking Facebook users or indirectly by alleging them from users' profile. In our system, we have used several fundamental characteristics which in accordance with the authors' expertise in the domain and with past experiments conducted by them (Troussas *et al.*, 2013 and Troussas *et al.*, 2015) tend to influence the educational procedure:

- Emotional state: Emotions can affect the educational process by promoting or downgrading the willingness of users in learning.
- Age: This characteristic provides significant information about the efficiency of users to conceive new information. It is widely accepted that age can play a very crucial role in the understanding of new concepts and ideas.
- Score: This characteristic shows information about the prior-existent knowledge of students in the curriculum being taught and may come of preliminary tests or preparatory lessons.
- Gender: This characteristic is used to check the likelihood of various differences between the sexes. This characteristic shows the degree of differentiation in learning between male and female students.
- Number of languages spoken: This characteristic can answer the question "Do you think that you have a flair for languages?". It is widely accepted that the more languages the user knows, the more apt s/he is in learning a new one.
- Educational levels: This characteristic provides information concerning the levels of education of the user. The underlying reasoning is that the language learning ability is proportional to the educational qualifications.

- Work experience: This characteristic can show the responsibility of users and can imply how experienced a user is in learning new concepts.
- Duration of computer use: This characteristic reveals information about users' tendency in computers. Then computer-based approaches in learning may have better results in the educational process.

Using the prototype application, the aforementioned characteristics were extracted from each user. Basically, as mentioned before, all of them except score and duration of computer use were gathered from their Facebook profile. Concerning the emotional state of the user, it is drawn and analyzed from his/her status in Facebook by using Rocchio classifier. Based on the aforementioned characteristics, the system creates clusters of the already existing students.

In view of the above, in this paper we focused on "measuring" the emotional sate of each user and then according to this state and his/her personal user model, we provide him/her with advice concerning the ability to start or proceed with the language learning application and we propose him/her other users for collaboration.

Fig. 3 illustrates how affect recognition can be involved in the educational process.

5. Experimental Results And Discussion

In this study, we used the Rocchio classifier in order to compare its performance in predicting whether a Facebook status update is positive or negative with the emotional status of Facebook where a user can directly state his/her emotions as in the following figure. We collected around 7000 status updates from 90 users. The status updates were then manually labeled as positive or negative. The Table 1 contains sample of status updates in each class.

Since there were a lot fewer negative samples, we based the distribution of the final dataset from it. We used the following data distribution (Table 2) for training and testing set (50%–50%):

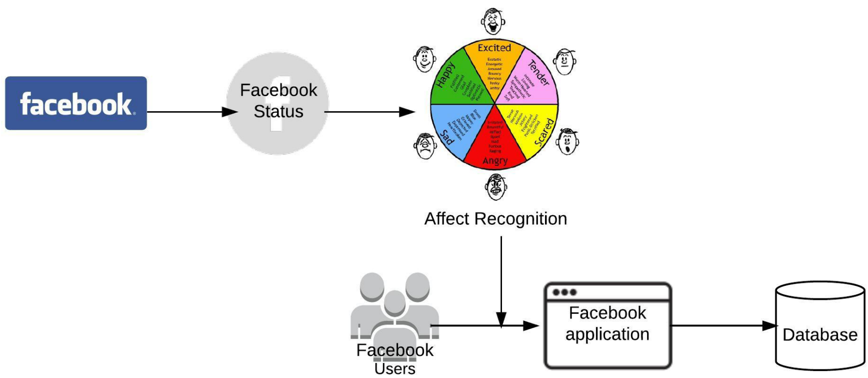

Fig. 3. General Architecture.

Table 1
Sample of status updates

Sample of negative status updates:	Sample of positive status updates:
you & me won't be happy anymore ☹	waiting for holidays!!!
i wish you were here…	thanks my friends! love you all ☺
tired and sick... ☹	celebrating a year together <3

Table 2
Data distribution

	Training	Testing
Positive	1135	1135
Negative	1135	1135

The dataset for each partition was selected randomly. The classifier was compared in terms of precision, recall and F-score performance using the computations shown below:

$$\text{Precision} = \frac{t_p}{t_p + f_p} \tag{3}$$

$$\text{Recall} = \frac{t_p}{t_p + f_n} \tag{4}$$

$$\text{F} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Precision and recall are the basic measures used in evaluating search strategies. These measures assume that there is a set of records in the database which is relevant to the search topic Records are assumed to be either relevant or irrelevant (these measures do not allow for degrees of relevancy). The actual retrieval set may not perfectly match the set of relevant records. Precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. The Table 3 summarizes the results.

Table 3
Rocchio precision and recall performance

Rocchio Classifier	Actual Positive	Actual Negative
Predicted Positive	0.74	0.23
Predicted Negative	0.26	0.77

The Table 4 compares the precision, recall, and the F-score of Rocchio classifier and the direct state of emotions of Facebook users (Fig. 4).

Based on the F-score, Rocchio classifier performed very well without significant differences to the direct state of emotions by Facebook users.

The reason why the probabilistic approach of Rocchio algorithm has been used was that it can indeed show performance improvements of reduction of error rate and noise in Facebook status.

6. Conclusions And Future Work

In this paper, we described the affect recognition for intelligent language learning using Rocchio Classifier. Furthermore, we presented important features for achieving a probabilistic approach of Rocchio classifier. The significance of using a more probabilistic approach of Rocchio algorithm for affect recognition is that the probabilistic methods are preferable from a theoretical viewpoint, since a probabilistic framework allows the clear statement and easier understanding of the simplifying assumptions made.

The used data is a random sample of streaming Facebook states and were not collected by using specific queries. The size of our hand-labeled data allows us to perform cross validation experiments and check for the variance in performance of the classifier across folds. In this way, knowing the emotional state of each user, we can use this characteristic as a value of the vector used for the group profiling, which can further ameliorate the educational experience through Facebook.

Finally, we present our experimental results, which show that the accuracy in analyzing the emotional state of Facebook users, using Rocchio Classifier, is really high.

Table 4

Precision, recall and f-score comparison of rocchio classifier and direct emotional state of facebook users

	Direct state of emotions of Facebook users	Rocchio Classifier
Precision	1.00	0.76
Recall	1.00	0.74
F-score	1.00	0.75

Fig. 4. Way of direct state of emotions of Facebook users.

The main findings of this study are the proper training of the system so that it can accept inputs in the form of Facebook status updates (disregarding updates that do not contain words or face emoticons) and classify the polarity of an opinion per status update basis. Hence, the affect recognition of students will serve as a characteristic for the group profiling to the direction of collaboration in the educational process.

Limitations of this study could be that the Rocchio algorithm cannot succeed to some extent in classifying multimodal relationships. For instance, two queries of similar emotions may appear much further apart in the vector space model. However, this does not affect the educational process at all, because the affect recognition will achieve to identify the student's emotions.

Different people can benefit from this study as follows: Students can gain knowledge from the collaboration with their peers of same or different groups and teachers can be assisted in the educational process given the grouping of their students. Moreover, the results of this study can also be used in other fields, e.g. special education needs, advertisement, user modeling and personalization, etc.

It is in our future plans to perform further study on the recognition and analysis of emotional states of Facebook users in order to further promote the language learning procedure. Furthermore, the relaxation as well as the combination of the assumptions resulting from the probabilistic framework provide promising starting points for future research.

References

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R. (2011). Sentiment analysis of Twitter data. In: *LSM '11 Proceedings of the Workshop on Languages in Social Media, Association for Computational Linguistics*, 30–38.

Barbosa, L., Feng, J. (2010). Robust sentiment detection on twitter from biased and noisy data. In: *Proceedings of the 23rd International Conference on Computational Linguistics*, 36–44.

Basile, T., Esposito, F., Ferilli, S. (2011). Improving user stereotypes through machine learning. *Communications in Computer and Information Science*, 249, 38–48.

Boiy, E., Hens, P., Deschacht, K., Moens, M.F. (2007). Automatic sentiment analysis. In: On-line Text. *Proceedings of the 11th International Conference on Electronic Publishing*, 349–360.

Darling-Hammond, L., Orcutt, S., Strobel, K., Kirsch, E., Lit, I., Martin, D. (2003). Feelings count: emotions and learning. *Stanford University School of Education*, 89–104.

Friaz-Martinez, E., Chen, S.Y., Macredie, R.D., Liu, X. (2007). The role of human factors in stereotyping behavior and perception of digital library users: a robust clustering approach. *User Modelling and User-Adapted Interaction*, 13, 305–337.

Gamon, M. (2004). Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. In: *Proceedings of the 20th international conference on Computational Linguistics*, 841–847.

Go, A., Bhayani, R., Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, 1(12).

Godbole, N., Srinivasaiah, M., Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. In: *Proceedings of International Conference on Weblogs and Social Media*, 7, 1–21.

Joachims, T. (1997). A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization. In: *Proceedings of the International Conference on Machine Learning*, 143–151.

Licchelli, O., Basile, T.M.A., Di Mauro, N., Esposito, F., Semeraro, G., Ferilli, S. (2004). Machine learning approaches for inducing student models. In: *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 3029, 935–944.

Mullen T., Collier N. (2004). Sentiment analysis using support vector machines with diverse information sources. In: *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 412–418.

Nino, A. (2009). Machine translation in foreign language learning: language learners and tutors perceptions of its advantages and disadvantages. ***ReCALL*,** 21, 241–258.

Oatley, K., Nundy, S. (1996). Rethinking the role of emotions in education. In: D.R. Olson & N. Torrance (Eds.), *The Handbook of Education and Human Development: New Models of Learning, Teaching and Schooling*, 257–274.

Pak, A., Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation*, 1320–1326.

Palubinskas, G., Descombes, X., Kruggel, F. (1998). An unsupervised clustering method using the entropy minimization. In: *ICPR'98*, 2, 1816–1818.

Pang, B., Lee, L. (2004). A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In: *ACL '04 Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, 271–278.

Shi, Y., Sha, F. (2012). Information-Theoretical Learning of Discriminative Clusters for Unsupervised Domain Adaptation. In: *Proceedings of the 29 th International Conference on Machine Learning*, 1079–1086.

Troussas, C., Virvou, M., Alepis, E. (2014). Collaborative learning: group interaction in an intelligent mobile-assisted multiple language learning system. *Informatics in Education*, 13(2), 279–292.

Troussas, C., Virvou, M., Caro, J., Espinosa, K.J. (2013). Multivariate clustering for group language learning in Facebook. In: *Frontiers in Artificial Intelligence and Applications:Proceedings of the 6th International Conference on Intelligent Interactive Multimedia Systems and Services*, 254, 80–88.

Troussas, C., Virvou, M. (2013). Information theoretic clustering for an intelligent multilingual tutoring system. *International Journal of Emerging Technologies in Learning*, 8 (6), 55–61.

Troussas, C., Virvou, M., Espinosa, K.J. (2015). Cooperative learning assisted by automatic classification within social networking services. In: G.A. Tsihrintzis *et al.* (eds.), *Intelligent Interactive Multimedia Systems and Services in Practice, Smart Innovation, Systems and Technologies.* Switzerland, Springer International Publishing, 36, 35–47.

Vinh, N.X., Epps, J., Bailey, J. (2010). Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, 11, 2837–2854.

Wilson, T., Wiebe, J., Hoffman, P. (2005). Recognizing contextual polarity in phrase level sentiment analysis. In: *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 347–354.

C. Troussas is currently a Ph.D. Candidate in the Department of Informatics at the University of Piraeus in Greece. He received a M.Sc. degree in "Advanced Computing and Informatics Systems" (2010) and a B.Sc. degree in Informatics (2008) from the same Department. He has published over 30 articles in international conferences, books and journals. His current research interests are in the areas of user modeling, social networking services, artificial intelligence in education, mobile learning and adaptive systems.

K.J. Espinosa is an Associate Professor and Head of the Department of Computer Science, University of the Philippines Cebu in Philippines. His research interests are in the areas of machine learning for big data specifically on event and pattern detection, sentiment analysis and recommender systems.

M. Virvou is a Full Professor at the Department of Informatics, University of Piraeus in Greece. She received a Ph.D. Degree in Computer Science and Artificial Intelligence from the University of Sussex, UK (1993), a M.Sc. degree in Computer Science from the University College London, UK (1987) and a B.Sc. degree in Mathematics from the University of Athens, Greece (1986). She is Head of the Department of Informatics, University of Piraeus (Greece), Director of the Software Engineering Laboratory and also Director of postgraduate studies in the same Department. She has published over 300 articles in international conferences, books and journals. Her research interests are in the areas of user modeling, human-computer interaction, knowledge-based software engineering, artificial intelligence in education, adaptive systems and affective computing.