Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 3388–3393, DOI 10.12694/scpe.v25i5.3056

LEVERAGING EMOTIONS IN STUDENT FEEDBACK TO IMPROVE COURSE CONTENT AND DELIVERY

ABID HUSSAIN WANI*

Abstract. Emotions play a vital role in almost all the activities we perform, including learning. In fact, the success of any learning system is largely dependent upon its ability to deliver the course content in such a form so as to meet the learning requirements of the target audience. Learning Systems can be tailored to effectively utilize the feedback from learners to improve the course content, and thus the feedback can prove to be a valuable asset. There is an increased demand for focusing on a learner-centric approach to content delivery. In this study we attempt at detecting different learning-relevant emotions from the feedback for a course so as to enable course designers to incorporate the type of content that matches a learners requirements. Rather than taking into account six basic emotions (sadness, happiness, fear, anger, surprise and disgust) we consider interest, engagement, confusion, frustration, disappointment, boredom, hopefulness and satisfaction emotions for the purpose of our study since they are more relevant in a learning setup. We employed a supervised algorithm, Support Vector Machine, for affect detection from the textual feedback in our experiments.

Key words: Emotion Detection, Support Vector Machine, Content delivery, Feedback Mining

1. Introduction. The rapid growth in information technology coupled with the huge benefits it offers, has increased its adoption by leaps and bounds in almost every sphere of our lives. In the field of education too, ICT-based solutions are playing a major role and providing overall value addition to the process of learning [1]. Specifically, eLearning Systems have a potential to reform the traditional teaching structure by incorporating technology-driven learning stuff and more importantly catering to the needs of individual learners who may not have a face-to-face interaction with the instructor [2]. Although much research has been conducted to understand the learning requirements of a group audience, there is a lot of scope for improving the content presented to the audience with varied emotional states. Since the target group of learners can be in different emotional states the content delivery can be tailored so as to match and satisfy their individual needs. The essence of these systems lies in providing individualized instruction by being able to cater and furnish to the varying knowledge grasping capacities and information needs of prospective learners. In a learning setup, we are more concerned about effective absorption of the content by the learner. Therefore, contemporary eLearning Systems are designed to be more and more learner-centric and close to the students self-learning needs as per the demand from the student community. A course can be tagged as learner centric only if the course is rich enough to include content for learners who can be in different emotional states. The fact that the human brain performs a blend of both cognitive and affective processing demands that the systems which are modeled after it not only process the information from a cognitive point of view but also integrate and assimilate an affect sensing and processing functionality into the system. Thus, the technology-driven learning environments should not only be intelligent enough but also take into consideration emotional aspects [3]. Having emotion processing capability will enable these systems to customize the contents of a course and its flow as per the needs of the learner so as to make the learning process more productive. True engagement of the prospective learners both intellectually and emotionally forms the hallmark of productive eLearning. The student feedback pertaining to a course can prove as an asset to structure the content of a course and decide the delivery plan by detecting the emotions expressed in the feedback. In this work, we propose a framework grounded on supervised learning that performs this affect detection from student feedback and provides generalized guidelines for the course designers to fine-tune the course contents and delivery plan for each target category of students according to their mood and learning pattern.

*South, Campus, University of Kashmir, Jammu and Kashmir, India (abid.wani@uok.edu.in,).

2. Related Work. The utility of emotions expressed by students in eLearning environments has gained much attention with the advancements in machine learning during the recent years. The major focus in this area has been to detect fine-grained emotions to track down the emotional response of the learners in an automated and seamless manner with least additional investment in terms of tangible and intangible resource [4, 5]. The research in this area comes under a more general area in computing known as "Affective Computing". A good amount of research has been conducted to detect learner's affective state [6, 7]. Most of the studies that have been carried out in this area principally take into account models and theories of emotions and employ a number of techniques to computationally capture the emotions. Since the focus of the present work is on recognizing emotions only from textual data, this section will trace the studies made for recognizing emotions from text. Emotion recognition involves theories and concepts from various areas including psychology, linguistics, information retrieval, text mining and recently machine and deep learning techniques have gained a lot of attention as well. Based on how emotions are modeled computationally and what techniques are employed, researchers have been able to achieve results which are close to that of humans. To detect emotions in text a number of approaches have been employed in different studies. Principally these approaches can be categorized into three categories; pure emotion keyword based approach, rule-based, (machine and deep) learning or combination of these. An emotion-aware framework for elearning systems has been presented in [8] that employs supervised learning approach. Regardless of the approach used, almost all methods have their own strengths and weaknesses when it comes to proper identification of emotional affinity [9]. Most emotion-recognition models are driven by identification of syntactic features (e.g. Parts-Of-Speech tags, N-grams etc.) as well as semantic features (e.g. synonym sets). There is a lot of scope for building syntactic and semantic resources which will serve as good resources for affect detection and a number of works have already been completed in this direction [10, 11]. Affect detection has proved to be an important tool in the deciphering of the interaction with the student for eLearning Systems in recent years [12]. An automated affect sensitive intelligent tutoring system, introduced by D'Mello and Graesser, tries to simulate dialogue patterns of tutor-pupil of real-world [13]. Emotions other than the basic ones have been considered for the purpose of affect computing in only a few studies. Liew et al. [14] constructed a text corpus comprising 15,553 tweets and annotated with twenty eight different emotions for their work on fine grained emotion recognition. Among the various emotions taken into account by them, emotions like boredom, confidence, excitement etc. are of particular interest in our study as well. In [15], Abdul-Mageed and Ungar have tried to build a large corpus using twitter hashtags thereby effectively removing the major impediment of creating instances labeled with emotion categories. For the purpose of our study, however, our corpus comprises of student feedback in textual form and moreover we take into account eight emotions which are most relevant to a learning environment. Our primary focus is on the detection of emotions relevant to learning activity from student feedback to fine-tune the course contents and delivery plan.

3. Detecting Emotions and Tailoging Course Content and Flow. Although the design of the content and its flow is dictated by various technical input factors relevant to a particular course yet the feedback received can be utilized to fine-tune the design. Fig. 1 shows the overall setup of our framework for detecting and utilizing affect information from student feedback for a particular course. The affect perception from students is then aggregated to take into account only the dominant emotions expressed to remove the outlier effect.

3.1. Detecting Emotions From Student Feedback. The task of emotion detection from student feedback has been modeled as a text classification task. After initial phase of preprocessing which consists of tokenization, POS tagging, lemmatization, stop-word removal, we train our supervised classifier Support Vector Machine. Support Vector Machines (SVM) learn to recognize emotions in data by identifying an optimal hyperplane that separates different emotion categories in a feature space derived from textual data. Initially, textual data is transformed into numerical feature vectors, incorporating linguistic cues. Through training on labeled datasets where each text sample is associated with a specific emotion label, SVM adjusts its parameters to maximize the margin between different emotion classes while minimizing classification errors. This process results in the creation of a decision boundary that delineates regions corresponding to distinct emotional categories. Subsequently, SVM utilizes this decision boundary to classify new text samples by determining which side of the boundary they fall on, effectively assigning them to the appropriate emotion category. The classification of emotion from the student feedback takes place in the following steps:

1. Feature Extraction: The process begins with preprocessing the textual feedback, which involves tasks such as tokenization, removing stopwords, and stemming or lemmatization to normalize the text. Then, features are extracted from the preprocessed text. These features include various linguistic cues such as word frequencies, sentiment scores, part-of-speech tags etc. Each feedback is represented as a numerical feature vector based on these extracted features.
2. Training Phase: In the training phase, SVMs learn to distinguish between different emotion categories by finding the optimal hyperplane that separates the feature vectors belonging to each emotion class in the feature space. The hyperplane is a decision boundary that maximizes the margin, which is the distance between the hyperplane and the nearest data points (support vectors) from each emotion class. SVM aims to find the hyperplane that minimizes classification errors while maximizing the margin.
3. Optimization: SVMs optimize their parameters during training to find the hyperplane that best separates the emotion classes. This optimization process involves solving a constrained optimization problem, typically using techniques like gradient descent or quadratic programming. The objective is to find the parameters (weights and bias) that define the hyperplane, ensuring that it separates the classes with the maximum margin.
4. Kernel Trick: SVMs employ the kernel trick to handle nonlinear relationships between features and emotions. By transforming the feature space into a higher-dimensional space, SVMs can find a hyperplane that effectively separates the data points even when the relationships are nonlinear.
5. Feedback Tagging: Once trained, SVMs classify new textual feedback samples by determining which side of the decision boundary they fall on in the feature space. The signed distance of a data point from the decision boundary is used to predict its emotion category. SVM assigns the sample to the emotion class corresponding to the side of the decision boundary it lies on.

In our study we used two datasets. One is the student feedback dataset from Menekse et al. [16]. This dataset comprises both the students' responses and the gold-standard summaries created by the teaching assistant. The other dataset again encompasses student feedback, which Oza et al [17] have analyzed using artificial neural networks. Since the proposed method will be based on supervised learning we need a training corpus consisting of student feedback where each record has been assigned a certain emotion label or marked as neutral. As discussed above we take into consideration only those emotions which are relevant for affect detection in student feedback; those include interest, engagement, confusion, frustration, disappointment, boredom, hopefulness, satisfaction. For the purpose of our study, four annotators labeled both the datasets to serve as training data for Support Vector Machine, employed for emotion classification. The job of an annotator is simply to select an appropriate emotion for each of the student responses presented to him. Since the student feedback normally will be in the form of a sentence or paragraph hence the proposed system will limit the analysis to sentence level.

The interpretation of affect in text being highly subjective [18], as such it is quite possible that the perception of one judge/annotator is different from the other one. In order to account for this subjectivity, rather than being annotated by a single judge we put each student's feedback for annotation by four annotators/judges. To get a quantitative measure of the inter-judge agreement statistic Cohen's kappa is employed. The pairwise agreement in emotional categories for student feedback is shown in Table 3.1.

As evident from the inter-annotator agreement study, the perception for an emotion is person-specific. However from the pair-wise results it is evident those human judges mostly agree on the instances of interest, confusion and satisfaction.

3.2. Tailoring Course Content and Flow. Once the dominant emotional response is known, the next step is to use this affective feedback to customize the course content. For example, if the emotion detected is frustration (learner got frustrated), there is a requirement to generate hints to help the learner in understanding a particular topic, and include certain simple and illustrative examples. If the emotion detected is "boredom", then it would be imperative to display content for the learner that will get the learner look for his/her own titles of interest, likewise if the student is confused with course contents, more elaborate and worked examples need to be incorporated into the course so that the student can understand the concept being presented [19].The emotion-inspired improvements from student feedback can be serve as a guiding factor for the inclusion of various

Table 3.1: Pair-wise Agreement in Emotional Categories

Emotion	$Judge1 \leftrightarrow Judge2$	$Judge1 \leftrightarrow Judge3$	$Judge1 \leftrightarrow Judge4$	Average
interest	0.85	0.77	0.76	0.79
engagement	0.65	0.66	0.54	0.61
confusion	0.87	0.81	0.84	0.84
frustration	0.75	0.78	0.70	0.74
disappointment	0.65	0.43	0.55	0.54
boredom	0.55	0.59	0.71	0.61
hopefulness	0.67	0.66	0.66	0.66
satisfaction	0.79	0.88	0.81	0.82

Table 3.2: Course Improvement Suggestions

Emotion Detected in the Feedback	Course Improvement Suggestion
interest	Links to more advanced topics
engagement	More detailed content
confusion	Elaborate worked examples on the topic
frustration	The more simple and precise explanation
disappointment	Elementary Video/ animations
boredom	Display lively and funny examples
hopefulness	More detailed content
satisfaction	Links to related topics

Table 4.1: Evaluation Results of the Proposed Framework

	Menekse dataset			*Oza dataset*		
Emotion Category	Precision	Recall	$F1-score$	Precision	Recall	$F1-score$
interest	58.54	62.78	60.59	54.21	57.77	55.93
engagement	47.63	52.47	49.93	56.23	62.5	59.20
confusion	53.23	58.12	55.57	48.30	51.23	49.72
frustration	59.22	55.47	57.28	52.46	67.33	58.97
disappointment	54.65	61.45	57.85	64.55	61.25	62.86
boredom	68.66	76.98	52.58	71.23	65.85	68.43
hopefulness	56.30	62.45	59.22	48.56	55.75	51.91
satisfaction	46.25	51.22	48.61	45.67	62.33	52.72

value-additions like explanation, hints, worked examples to the course content as described in Table 3.2.

4. Evaluation and Results. The real essence of any machine learning-based framework lies in not only providing good results (classification results in our case) on the data on which it is trained but also on new and unseen data and under-fitting and overfitting of data is avoided. We validated our model using Leave-one-out cross validation, by training our model on all the instances of student feedback except one, “n” number of times and predicting Output emotion label for that one instance using Support Vector Machine. Table 4.1 depicts the classification performance results obtained for the above two datasets.

To measure the performance of the proposed framework, we compute the classification metrics including Precision, Recall and F1-score on Menekse’s dataset and Oza’s dataset. For Menekse’s dataset, best results were obtained for ‘interest’ emotion and that for Oza’s dataset ‘boredom’ emotion category was detected with highest F1-score. Classification metrics such as Precision, Recall, and F1-score are essential for evaluating the performance of machine learning models, particularly in tasks like emotion detection from text. Precision quantifies the proportion of emotions detected that are correctly reported as positives (true positives) from the whole number of examples which are tagged as positive (true positives + false positives) by the model. For the purpose of task of emotion classification, precision presents ability of a model to accurately identify a particular emotion category without misclassifying unrelated emotions. A model with high precision score implies that when the it suggests an emotion, there is high probability for it to be correct, minimizing false positives and ensuring the relevance and specificity of the predictions. Recall measures the proportion of correctly predicted positive cases (true positives) out of all actual positive cases in the dataset (true positives + false negatives). Overall, precision, recall, and F1-score are critical metrics for evaluating the performance of emotion detection models, as they offer insights into different aspects of the model's performance, such as accuracy, relevance, and comprehensiveness.

5. Conclusion and Future Scope. This paper proposes a supervised learning based emotion detection mechanism for detecting emotions from student feedback to tailor and enrich the course contents. The classification results obtained on Menekse's dataset and Oza's dataset does not seem too good which is due to small size of the text corpus. In future, we aim to build a larger annotated text corpus so as to achieve better classification performance. The limitation of our work lies in the fact that we only take into account the textual feedback received from the learner to decide on the content. The course content and delivery can further be improvised by taking into account other modalities including facial expressions, speech and voice modulations and not just written textual utterances. We limited our study to text as we take into account only those systems which receive just the textual feedback from the learner. In future, this work can be extended to have multi-modal feedback so as to guide the content design, presentation and delivery. In our experiments, though we have not made appropriate substitution of slangs (e.g. "dunno" with "don't know") and orthographic features (e.g. "sooconfuuusing"), we intend to take these issues into account as well to better gauge the emotion expressed in the student feedback., so as to enable course designers get a better idea about what all improvements need to be made to the course content and flow. Moreover, it must be noted here that detecting learners' current mood is not the only solution but one factor to be considered; there is more research needed to identify other situational and technical input factors that should guide the design, content and delivery of a course.

REFERENCES

[1] K. Meteshkin, O. Sokolov, O. Morozova, and N. Teplova, Integration of Higher and Secondary Education: Problems and ways of their solution on the basis of Information Technologies, *Journal of Education, Health and Sport* **6** (2016), no. 7, pp. 375–390.

[2] I. Ibrahim and K. Al, Smart Learner-Centric Learning Systems Smart Learner-Centric Learning Systems, *Proceedings of the International Conference on Information Science and Applications ICISA 2017*, Barcelona, Spain, March 20, 2017,Springer, Singapore (2017), 577-584

[3] M. Feidakis, T. Daradoumis, and S. Caballé, Endowing e-learning systems with emotion awareness, *Proceedings of the 3rd IEEE International Conference on Intelligent Networking and Collaborative Systems, INCoS 2011*, Fukuoka, Japan, Nov 30 –Dec 2, 2011,IEEE, Japan (2011), 68–75

[4] M. Wong, Emotion Assessment in Evaluation of Affective Interfaces, *Neuron* **65** (2006), no. 3, p. 293.

[5] G. Zimmermann, Beyond Usability – Measuring Aspects of User Experience, Ph.D. thesis, Swiss Federal Institute of Technology Zurich (2008).

[6] S. K. D. Mello, S. D. Craig, A. Witherspoon, and B. Mcdaniel, Automatic detection of learner ' s affect from conversational cues, *User Modeling and User-Adapted Interaction* (2008), no. 1, pp. 45–80.

[7] Wang, Yang. "Affective state analysis during online learning based on learning behavior data." Technology, Knowledge and Learning 28.3 (2023): 1063-1078.

[8] AH.Wani, R. Hashmy, An Emotion-aware Framework for eLearning Systems, *Proceedings of 1st International Interactive Information Retrieval Conference*, Kish, Iran, Feb 22–24, 2017,Tehran University, Tehran (2017), 25–30

[9] Alslaity, Alaa, and Rita Orji. "Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions." Behaviour & Information Technology 43.1 (2024): 139–164.

[10] S. Aman and S. Szpakowicz, Identifying Expressions of Emotion in Text, *Proceedings of the International Conference on Text, Speech and Dialogue*, Kuala Lumpur, Malaysia, April 3–5, 2009,IEEE, Malaysia (2009), 195–205

[11] S. Afzal and P. Robinson, Designing for Automatic Affect Inference in Learning Environments, *Journal of Educational Technology & Society* **14** (2011), no. 4, pp. 21–34.

[12] E. Bevacqua et al., Interacting with emotional virtual agents, *Proceedings of the International Conference on Intelligent Technologies for Interactive Entertainment*, Genova, Italy, May 24–27, 2011,Springer, Italy (2011), 243–245

[13] S. K. D'Mello, B. Lehman, and A. Graesser, A Motivationally Supportive Affect-Sensitive AutoTuto. In *S. K. D'Mello (ed) New perspectives on affect and learning technologies*, Springer, New York (2011), 113–126.

[14] J. Liew, S. Yan, H. R. Turtle, and E. D. Liddy, EmoTweet-28: A Fine-Grained Emotion Corpus for Sentiment Analysis,

Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Portoroz, Slovenia, May 23–28, 2016,Springer, Slovenia (2016), 1149–1156

[15] M. Abdul-Mageed and L. Ungar, EmoNet: Fine-Grained Emotion Detection with Gated Recurrent Neural Networks, *Proceedings of the 55th annual meeting of the association for computational linguistics*, Vancouver, Canada, July 03–Aug 4, 2017,Association for Computational Linguistics, Canada (2017), 718–728

[16] M. Menekse, G. Stump, S. Krause, and M. Chi, The effectiveness of students' daily reflections on learning in engineering context, *Proceedings of the ASEE Annual Conference and Exposition*, Vancouver, Canada, June 26–29, 2011, Canada (2011), 1149–1156

[17] K. S. Oza, R. K. Kamat, and P. G. Naik, Student Feedback Analysis: A Neural Network Approach, *Proceedings of the Tenth International Conference on Information and Communication Technology for Intelligent Systems*, Ahmedabad, India, March 25–26, 2017,Springer, Cham (2017), 342–348

[18] H. Binali, C. Wu, and V. Potdar, Computational approaches for emotion detection in text, *Proceedings of 4th IEEE International Conference on Digital Ecosystems and Technologies*, Dubai, United Arab Emirates, April 13–16, 2010,IEEE, Cham (2010), 172–177

[19] S. D'Mello and R. A. Calvo, Beyond the basic emotions: what should affective computing compute?, *Proceedings of CHI '13: CHI Conference on Human Factors in Computing Systems*, Paris, France, April 27– May 2, 2013,Association for Computing Machinery,United States (2013), 2287–2294

Edited by: Mudasir Mohd
Special issue on: Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions
Received: Jan 1, 2024
Accepted: Apr 12, 2024

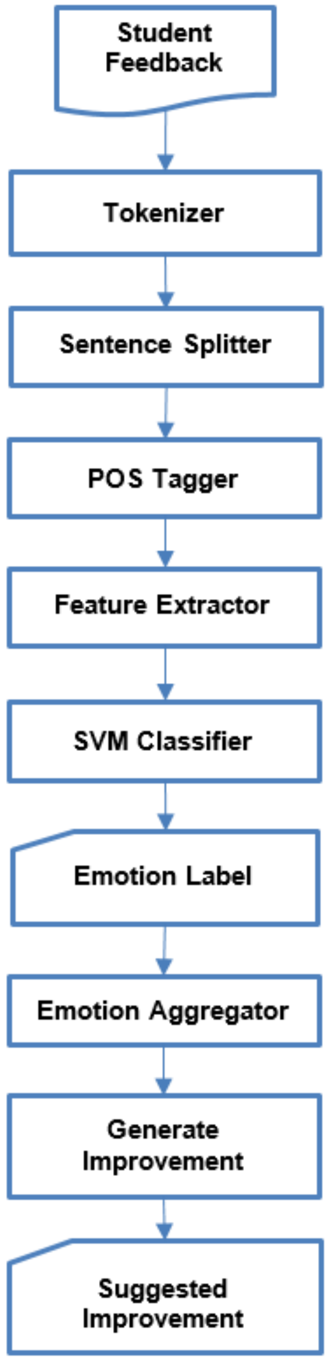

Fig. 3.1: Overall Framework for Course Improvement Using Student Feedback