Affective Computing: A Topic-Based SER Approach on Collaborative Discussions in Academic Setting

Conference Paper · October 2023

DOI: 10.1109/FIE58773.2023.10342963

Nasrin Dehbozorgi
Kennesaw State University

Affective Computing: a Topic-Based SER Approach on Collaborative Discussions in Academic Setting

Nasrin Dehbozorgi
Department of Software Engineering
Kennesaw State University
Marietta, GA, USA
dnasrin@kennesaw.edu

Mourya Teja Kunuku
Department of Computer Science
Kennesaw State University
Marietta, GA, USA
mkunuku@students.kennesaw.edu

2023 IEEE Frontiers in Education Conference (FIE) | 979-8-3503-3642-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/FIE58773.2023.10342963

***Abstract*—One of the biggest concerns in the modern day especially in the educational domain centers on the student's mental health. High rates of anxiety and depression have especially brought the attention of researchers in engineering education to apply affective computing to help with students' academic performance. It is known that a person's emotional states cause physiological and physical changes in the body. Emotions may impact facial expression, tone of speech, blood pressure, pulse, etc. Since visual and auditory signals are two variables that can be measured without the need to attach any physical device to the individuals, they are most studied in this field. Speech in particular has been known as a means that transfers much information about the mental and emotional states of the person. Speech Emotion Recognition (SER) is a growing field that has been applied in several domains including engineering education. Recent advancements in AI, Natural Language Understanding (NLU), and Large Language Models (LLM) have significantly streamlined this line of research. In this work which is a continuation of our prior work, we propose a speech analysis model that extracts both the emotions and topics from verbal discussions in a computer science classroom to understand if the expressed emotions were mostly about the course related topics or not. The goal of this research is to develop a tool that helps educators gain insights into the students' emotional states in teamwork and also understand the context of their conversations. We further analyze if the expressed emotions in the verbal class discussions are mostly about the course content or other subjects outside class setting. To expand the emotion analysis module we added a new layer to our developed pipeline by passing the speech data into the ChatGPT API to generate summarized scripts and extract additional classes of emotion. The preliminary results from this study are promising, indicating the potential value of this research direction and its prospects for further development. Application of this model in the educational domain can greatly benefit both educators and students and allows the instructors to make necessary interventions needed to maximize students' positive experiences in team settings while considering their emotional states.**

***Index Terms*—Speech Emotion Recognition (SER), Large Language Models (LLM), NLP, Topic modeling, Affective computing, ChatGPT API, Teamwork, Engineering education**

I. Introduction

Affective computing is an area that focuses on developing intelligent systems that are capable of recognizing and interpreting human emotions and has applications in multiple fields including education. Historically students' performances have been evaluated based on their knowledge of the course independent of their affective states. Research suggests that emotions play a critical role in the mental functions of the individual and can help in determining a person's behavior or performance in a given context [7]. Affective computing as a multidisciplinary field integrates different aspects of computer science, cognitive science, psychology, and artificial intelligence to develop systems that capture emotions via body language, facial recognition, speech recognition, and other behavioral patterns. The application of affective computing in the educational domain can greatly enhance learning experiences and improve students' educational outcomes. Some examples of application of affective computing in education are, intelligent tutoring systems and recommender systems that personalize the instructions and recommend tailored learning material based on learner's emotional needs and preferences. One other important application of affective computing in higher education is emotion-based learning analytics and the development of social and soft skills in students.

Teamwork has always been a challenge in the educational domain, posing difficulties from different perspectives, including student engagement, individual assessment and conflict management. The development of an AI-based system that can provide tailored feedback to students to promote emotional regulation, empathy, and communication can greatly improve the quality of teamwork. Furthermore by application of such systems educators can gain insights into student's engagement and team's climate to provide timely intervention and support. The analysis of students' speeches in teamwork can provide valuable insights into their collaboration dynamics and the overall effectiveness of teamwork. Speech-emotion recognition (SER) is a common practice that involves analyzing various acoustic features and linguistic cues to extract multiple emotional features. Some of the common techniques used in SER are analysis of communication patterns, languages, lexicon analysis, and the assessment of non-verbal cues such as tone, pauses, or pitches in speech. There are multiple machine learning algorithms that are trained on labeled emotional speech data to build models that can classify or predict emotions. These models can incorporate various features extracted from speech and learn to recognize emotional patterns. In earlier research, we applied SER to study the impact of different emotions on students' performance in an introductory programming course. The multi-class sentiment analysis on students' dialogues in teamwork showed that teams that expressed more positive emotions in their speech earned higher final grades in the course. Although we didn't study the causal relationship between positive emotions and performance, given the limited available data, the early result showed promising future research. The findings of the prior study showed high accuracy in identifying different emotions from speech, however, the context in which the emotions were expressed was not analyzed. For example, we didn't study if the students' positive or negative expressed emotions were about teamwork experience, the course content, or other external factors. Analyzing emotions along with context analysis will provide additional signals and give educators better insight into the aspects that students are challenged with.

II. Related Work

Affective computing remains a prominent and highly researched topic in the field of education, attracting significant attention and ongoing exploration by researchers [1]. In [2] the authors emphasize the significance of SER and the challenges it poses due to the intricate nature of emotional expressions in speech. They explore the application of deep learning algorithms to address these challenges, that allow direct extraction of meaningful features from raw audio data. The proposed method involves an end-to-end deep learning model that applies a deep neural network directly to the raw audio waveform. Performance evaluation on pre-trained datasets demonstrates promising results, outperforming traditional machine learning approaches and deep learning techniques utilizing spectrogram representations. The findings of this paper highlight the effectiveness of the proposed approach in SER, which offers valuable potential for providing meaningful feedback and interventions in the engineering classes.

The application of SER in the educational domain is not a new concept and educational researchers have applied different AI algorithms that assist in SER to provide better learning experiences for students based on their affective states. One of the areas of SER application in education is called affective e-learning which improves the online learning experience for the learners based on the input emotional data. As an example in [3] the authors focus on the application of SER in an e-learning system and explore different ML models and classifiers to recognize emotions from input speech. They achieved 50% accuracy by application of neural networks. In another work [4] on affective e-learning the authors develop a real-time SER model to enhance students' learning experiences by incorporating real-time analysis of their emotions through speech. The highlighted finding of this work is the real-time analysis of the emotion to provide prompt feedback and assistance to students. However, due to resource availability issues, real-time analysis may not be a feasible option in all settings or at any level. That is the reason that most of the existing work on the emotional analysis of the student's speech is done either in a non-real-time manner or on their written posts in discussions, group chats, reflection, social media, etc.

The authors in [5] also acknowledge the significance of emotions as an important area of research within the education field. They address the challenge of efficiently and precisely classifying learners' emotions in specific aspects within a large-scale online learning environment using students' text/comments. To tackle this, the researchers developed a dimensional classification system for emotion aspects and proposed an aspect-oriented emotion recognition method. The method combines an Aspect-oriented Convolutional Neural Network (A-CNN) and an academic emotion classification algorithm based on long short-term memory with an attention mechanism (LSTM-ATT). Experimental results demonstrated the model's effectiveness, achieving an accuracy of 89% for the A-CNN model and 71% for the LSTM-ATT model on the test dataset. The study contributes a novel approach for measuring large-scale online academic emotions, supporting further research on students' well-being in online learning environments. One of the applications of SER in the educational domain is the development of models that can predict students' performance earlier in the semester based on their expressed emotions in teamwork [14], [15], [8]. In [6] The authors propose an aspect-based approach to identify specific aspects of emotions in speech and examine their relationship with performance outcomes. In this study, the researchers extend the analysis of students' speech in teams by conducting multi-class aspect-based emotion analysis. The process involves two main steps. Firstly, different classes of sentiment such as Happy, Angry, Sad Surprise and Fear are extracted from collaborative speech in an introductory programming course (CS1), and their correlation with students' performance is identified. Secondly, an Aspect-Based Emotion Analysis (ABEA) approach is employed. The researchers utilize supervised machine learning methods and rule-based models on speech datasets. After preprocessing the text, multiple classes of sentiments are identified. Aspect extraction is achieved through Part-of-speech (POS) tagging, and patterns are extracted from the identified aspects. Finally, a combination of emotion classes and aspect patterns is used as feature vectors to train the K-Nearest Neighbor (KNN) algorithm for predicting students' performance.

Towards the application of SER in analysis of students performance in another study [7] the authors investigate the relationship between students' positive sentiments expressed during teamwork and their individual performance in the course. This study introduces the use of verbal conversations as a means to measure affective states by focusing on low-stake teams, which have less emphasis in performance evaluation compared to capstone teams. By capturing students' affective states in low-stake teams, this research argues that individual's emotion states can impact their performance. The study develops a text-mining algorithm to operationalize positive sentiment and subjectivity levels from student's speech in teams. The results demonstrate a strong correlation between performance and positive sentiment, as well as the level of subjectivity. This method enables the evaluation of students' engagement in low-stake teams, allowing for early identification of lower performers and the provision of additional learning opportunities to support their progress.

In [8] the authors present a model for predicting students' performance based on the emotions they express in their team conversations, along with their formative assessment scores. The researchers utilized SpeechBrain to transcribe recorded speech and employed a transformer-based emotion recognizer (T5) to extract emotion classes. The results indicate the promising performance of the model, as the predicted values closely align with students' actual grades.

Finally, the authors in [9] examine the relationship between self-efficacy, positive emotion, and academic performance in the context of collaborative learning. Self-efficacy is a crucial component of social cognitive theory and plays a significant role in understanding human thoughts, motivations, and actions. The study highlights the impact of self-efficacy on students' performance and accomplishments, emphasizing its importance in engineering education as a metric for tracking improvement. Building upon previous work that demonstrated a positive correlation between students' positive emotions during team interactions and their course performance, the study investigates the correlation between self-efficacy, emotions, and performance. Self-efficacy is measured using the "Student Attitudes Toward STEM (S-STEM) Survey," which consists of 20 Likert-scale questions grouped into three categories of computer science, learning abilities and social skills. Students' emotions are extracted from their team speeches using NLP techniques. The data analysis reveals a statistically significant correlation between overall self-efficacy, course performance, and positive emotions during teamwork. Additionally, the study examines which specific self-efficacy question categories have the greatest impact on student's performance. Data shows that self-efficacy in interpersonal skills and learning ability significantly influence performance outcomes.

Overall SER techniques can be used to assess student engagement and emotional states during classroom activities. This information can help educators gauge the effectiveness of instructional strategies, identify areas for improvement, and tailor teaching approaches accordingly. By analyzing students' speech in teamwork, educators can gain a deeper understanding of team dynamics, communication effectiveness, and collaborative skills, and identify areas for improvement. This analysis can inform targeted interventions, facilitate constructive feedback, and promote more effective teamwork and collaboration among students.

As the AI algorithms advance there is a need to keep pace with the state of the art in developing SER pipeline. Towards this goal, in this research, we developed a model that analyses the context of students' speech in teams by topic and emotion extraction. To perform a more in-depth emotion analysis we add another layer to the pipeline that processes the input data by the use of Large Language Models (ChatGPT-API). In the following section details of the developed model are presented.

III. Research Method

The primary goal of this research is to extract topics (course related and non-related) as well as sentiments from students' verbal conversations during classroom activities. By utilizing NLP methods, we provide insightful information about student engagement, preferences, and overall performance thorough the identification of topics addressed during class activities and the analysis of associated sentiments. The developed research methodology consists of several steps in order to achieve the research goal. This process includes five steps: 1) Data collection, 2) Audio transcription, 3) Data cleaning, 4) Emotion detection and 5) Topic extraction and classification.

1) Data collection: The target population in this study is freshmen college-level students in the CS program. We collected their verbal conversations in the class as they worked on the class activities with their peers in low-stake teams. Data were collected in every session of the class during the semester from a total of 24 students.

2) Audio transcription : The AssemblyAI API is applied to convert the raw audio files into text format. The script uses the requests library to send HTTP requests to the API to achieve this. Three functions involved in this conversion process are: 1) the uploadMyFile() function that receives an audio file as its first argument. This procedure gets an upload URL and uploads the audio file to the AssemblyAI API server, 2) the AssemblyAI API server's startTranscription() function uses the upload URL that was obtained from uploadMyFile() to perform the transcription process. The authentication key is also needed by this function to approve the API request, 3) finally when the transcription is finished, the getTranscription() function uses the transcription ID to locate the text transcript on the AssemblyAI API server. The text transcript of the audio file is the final outcome of the procedure (Figure 1).

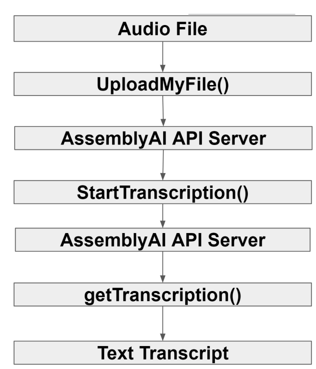

Fig. 1. AssemblyAI API Audio to Text Conversion

3) Data cleaning : We applied the Python package Neattext as our preferred tool for data cleaning as it significantly streamlines our text pre-processing workflow. This robust library provides thorough solutions to a range of problems that arise when working with unstructured text data. We can efficiently handle problems like the elimination of special characters, numbers, and URLs, which are frequently present in unstructured text data, by utilizing Neattext. Neattext also offers useful functionality for reducing noise in text data by removing extraneous elements like stop-words.

4) Emotion detection : In this module the Python Text2emotion package was employed to extract emotional classes. This package comprises several features that align with the objectives of this study.

For the emotion extraction we analyzed the pre-processed text and determined the frequency of each emotion in order to conduct additional analysis. This package recognizes words that are suggestive of expressing feelings or sentiments, establishes the emotion category connected to each sentence, and keeps track of the number of relevant emotions found. The extracted emotion classes are, happy, angry, surprised, sad, and fear. The package generates an output in the form of a dictionary, where the keys are emotion categories and the values are corresponding emotion scores. A message is categorized under a specific emotion category if it receives a higher score in that category. Utilizing the Text2emotion package's capabilities allowed us to analyze and interpret the emotions within the transcriptions. These results provide a profound understanding of the emotional dynamics shared among students.

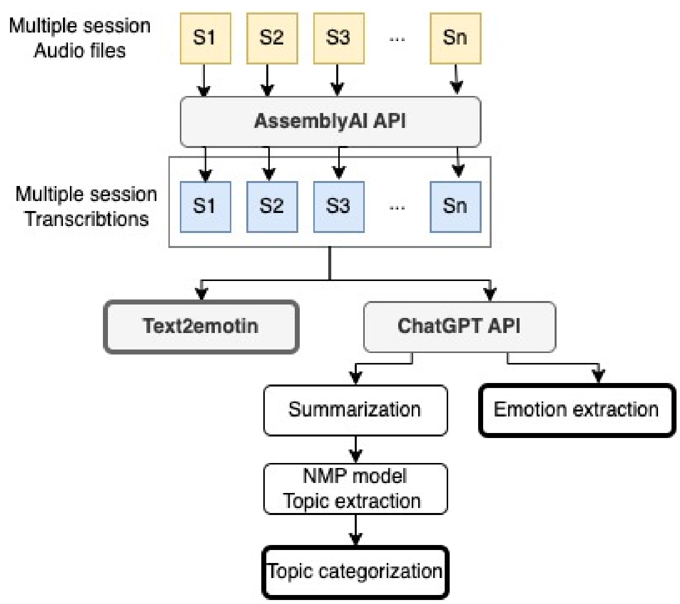

Fig. 2. Process of Emotion and Topic Extraction

5) Topic Extraction: The topic extraction phase of the process is divided into two distinct steps.

- Step 1: We first passed the input text to the ChatGPT API to get a summary of the conversation in the transcript. The main ideas and discussions covered in the text are summarized effectively through this API.
- Step 2: We fed the summarized output text into a topic modeling algorithm that we developed for topic extraction. This algorithm examines the text and then generates topics that make sense given the content. The algorithm recognizes recurrent patterns, relationships, and overarching themes present in the summarized text by using topic modeling techniques. This makes it possible to pick out important and illustrative topics that capture the essence of the conversation (Figure 2).

For topic modeling, we used the Non-negative Matrix Factorization (NMF) model, a widely used method in unsupervised learning. NMF has proven successful in performing a variety of NLP tasks, including document clustering, classification, and sentiment analysis. These tasks include extracting latent topics from text corpora and performing various NLP tasks. Following the transcript summarization via the ChatGPT API, we employed the NMF algorithm to extract more comprehensive and easily understandable topics from the text corpus.

The fundamental idea behind NMF is to split a non-negative matrix into two smaller, non-negative matrices. The matrices W (m x k matrix) and H (k x n matrix) serve as a representation for this decomposition. The objective is to keep all entries in W and H non-negative while minimizing the difference between the original matrix A an'd the product of W and H. The desired level of topic granularity determines the value of the hyperparameter k, which denotes the number of columns in W (Equation 1).

$$\frac{1}{2}\|\mathbf{A}-\mathbf{WH}\|_{\mathrm{F}}^{2}=\sum_{i=1}^{n}\sum_{j=1}^{m}(A_{ij}-(WH)_{ij})^{2} \tag{1}$$

In NLP applications, matrix A is typically represented as a document-term matrix, where each row corresponds to a document in the corpus, and each column represents a term in the vocabulary. The values within matrix A typically denote the frequency or some other measure of the term's occurrence in the corresponding document.

The primary objective of NMF in topic modeling is to uncover the latent topics inherent in the corpus. This is achieved by decomposing the matrix A into two separate matrices, W and H. The rows of the matrix (W) represent the topics and provide a distribution of the terms in the vocabulary. The values within each row of W indicate the relevance or contribution of the corresponding terms to that particular topic. On the other hand, the columns of the matrix (H) contain the weights or importance of the terms within each topic, reflecting the extent to which each term contributes to the overall representation of that document.

By factorizing the document-term matrix A into matrices W and H through NMF, we uncovered the underlying topics in the corpus and obtained their corresponding term distributions. This enabled us to interpret and analyze the topics by examining the most important terms associated with each topic, facilitating a deeper understanding of the content and structure of the documents within the corpus.

NMF generates non-negative weights for the terms in each topic, making them easier to interpret, which is one advantage it has over other topic modeling techniques. The topics that are produced by NMF also have a sparsity constraint, which limits the number of pertinent terms in each topic. As a result, the subjects are simpler to understand and apply in subsequent tasks. In topic extraction process we take advantage of the

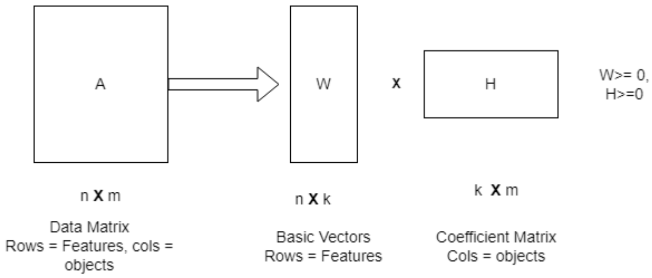

Fig. 3. NMF Equation Illustration

ChatGPT API's power to obtain a conversational summary that serves as the basis for the subsequent topic modeling algorithm. With the help of this approach, we were able to extract insightful information and identify relevant topics from the audio transcript, which helped us comprehend the conversation's subject matter better.

A. Evaluation Metrics

The accuracy of the transcripts produced by AssemblyAI for each of the audio files was assessed manually. The evaluation found that transcript accuracy was close to 93%, which is commendable. This shows how well the AssemblyAI performs in accurately transcribing audios recorded in a classroom setting. It's important to note that the transcription process has certain limitations. For instance, it is difficult to accurately capture sounds of laughter or other expressed emotions. Despite this obstacle, AssemblyAI handled ambient noise that is typically present in actual classroom environments well, demonstrating its high performance in noise handling.

In unsupervised learning algorithms, various metrics can be utilized to evaluate the clustering quality. The silhouette score, Calinski-Harabasz index, and Davies-Bouldin index are some of the frequently used metrics. By taking into account elements like intra-cluster similarity, inter-cluster separation, and cluster compactness, these metrics demonstrate the efficacy and coherence of the clustering algorithm.

1) silhouette score: By evaluating each data point's suitability within its assigned cluster in comparison to other clusters, the silhouette score acts as a reliable indicator of the quality of clustering. A higher score indicates well-defined clusters, which are characterized by data points that are significantly closer to their own cluster than to other clusters. This metric has a range of -1 to 1, where a lower score indicates the absence of well-defined clusters.

The following steps are used to calculate the silhouette score (SC)for each individual data point, 1) determining the intra-cluster distance 'a' by calculating the average distance between the data point and all other data points that are part of the same cluster, 2) calculating the nearest-cluster distance (abbreviated as 'b'), which is the average distance between the data point and all other data points that are contained within the closest neighboring cluster, 3) calculating the data point's silhouette coefficient (SC) using Equation 2:

$$SC = (b - a)/max(a, b) \tag{2}$$

SC measures how closely the data point aligns with its assigned cluster in comparison to the closest neighboring cluster. The final step is calculating the average silhouette score by averaging the silhouette coefficients across all data points after iterating through steps 1 through 3 for each dataset.

The resulting silhouette score is a measure to assess the quality of clustering, and shows the effectiveness and coherence of clustering algorithms [10].

2) Calinski-Harabasz index: The Calinski-Harabasz index is a metric for clustering quality that assesses how distinct and distant the clusters are from one another. Its foundation is the proportion of within to between-cluster variance.

The Calinski-Harabasz index (CHi) is calculated as follows: 1) calculating the data set's overall mean, 2) calculating the sum of squares between clusters (SSB), which is the sum of the squared distances between the means of each cluster and the overall mean, weighted by the number of data points in each cluster, 3) adding up the squared distances between each data point and the assigned cluster mean to obtain the sum of squares within clusters (SSW), 4) calculating the Calinski-Harabasz index using Equation 3,

$$CHi = (SSB/SSW) * (n - k)/(k - 1) \tag{3}$$

where n is the total number of data points, and k is the number of clusters.

A higher Calinski-Harabasz index suggests better-defined clusters with greater separation between them. The peak of the index, which represents the ideal number of clusters for a dataset, can be used to determine the index's usefulness. To fully comprehend the clustering results, it should be used in conjunction with other measures and visualizations since it is also sensitive to the size and dimensionality of the data. [11]

3) Davies-Bouldin index: The Davies-Bouldin index assesses both the compactness within clusters and the separation between clusters as a measure of clustering quality. It contrasts the average separation between points in various clusters with the average distance between the centers of each cluster. Better clustering with well-separated and compact clusters is indicated by a lower score.

The Davies-Bouldin index is calculated as follows: 1) determining the centroid of each cluster, which is the average of all the points assigned to the cluster, 2) calculating the average distance between the centroid of each cluster and each point that belongs to it (cluster dispersion), 3) calculating the distance between the centroids of each pair of clusters, 4) finding the cluster that is most similar to each cluster, where similarity is measured as the ratio of the sum of the cluster dispersion to the inter-cluster distance, 5)the highest similarities for each cluster are averaged to create the Davies-Bouldin index.

A lower Davies-Bouldin index indicates more compact, well-separated clusters and better clustering. When comparing different clustering algorithms or different parameter configurations for the same algorithm, this index is employed [12] [13].

The analysis of our extracted topic data using the Davies-Bouldin (DB) index and SC show consistent results with 14 topic clusters, as shown in Table 2. This offers insights into how well data points with similarities are grouped together to produce meaningful topics.

TABLE I
CLUSTERING SCORES

Clusters	silhouette scores	CH index	DB index
5	0.043286	1.318950	1.877807
6	0.041429	1.304241	1.618224
7	0.035379	1.282792	1.419212
8	0.032937	1.275745	1.213067
9	0.035161	1.292700	1.041855
10	0.034256	1.314997	0.915213
11	0.035788	1.327731	0.839533
12	0.034679	1.404734	0.661426
13	0.017911	1.371115	0.597953
14	0.006441	1.238834	0.508743

It is clear from Table 2 that using 14 clusters for topic extraction produces results that are highly satisfactory. This selection of cluster numbers shows a strong alignment between the topics generated and the conversation's underlying structure as well as higher levels of internal similarity. It ensures the coherence and uniqueness of each cluster while capturing the diversity of topics.

IV. RESULTS AND OBSERVATIONS

The developed model helps in the identification of key themes and subject matter that are frequently brought up in conversation by grouping these topics together and offering insightful information. This makes it possible for us to better comprehend, analyze, and extract meaningful information from speech conversations.

Analysis of the result shows a high level of coherency and accuracy in the extracted topics. As an example, in one audio file, the students mainly discussed the challenges they had with the class activity that was about search and sort algorithms with some additional side talks about physical fitness and football. The extracted topics (top 15 words) well represent the context of the discussion as shown in Figure 4.

```
The top 15 words for Topic #2
['confusing', 'struggles', 'method', 'gain', 'mentions',
 'talk', 'techniques', 'recording', 'football', 'sorting',
 'group', 'sort', 'algorithms', 'search', 'problem']
```

Fig. 4. A sample Topic Cluster

The developed model takes a step further by incorporating the analyzed emotions into the extracted topics from student conversations in each session of the class. This integration offers a deeper contextual understanding, encompassing both the discussed subjects and the emotional states expressed during these conversations. The extracted topics were further grouped into two categories of 'course-related' and 'non-course-related' topics. A sample of the outcome using the developed model is shown in Table 1. This table summarizes what are the dominant classes of emotions in a team's discussion during class activities as well as the proportion of course-related and non-course-related topics discussed in each session.

TABLE II
SAMPLE OUTCOMES OF TOPICS AND EMOTIONS

Audio File	Course Related Topics	Non-course Related Topics	Dominant Emotion
Session 1	33%	67%	Fear 38%, Surprise 22%
Session 2	87%	13%	Fear 44%, Sad 23%
Session 3	60%	40%	Fear 37%, surprise 23%
Session 4	86%	14%	Happy 37%, Surprise 23%
Session 5	60%	40%	Happy 43%, Surprise 23%

In addition to analyzing emotions extracted using Text2emotion, we conducted a comprehensive emotion analysis by utilizing ChatGPT API. We explored various emotional categories within students' conversations and examined their correlation with academic performance within each group. The students' cumulative score in the course formative assessments through out semester including quizzes, preparation assignments, class assignment, and lecture/lab test scores are considered as the academic performance metric in this study. Below we share some of the key findings from our analysis, shedding light on the relationship between students' emotions and their academic performance.

Teams that experience a higher percentage of emotions like confusion, uncertainty and frustration in course related topics have lower final grades. These emotions can indicate challenges in understanding the class activities or difficulty in applying the learned concepts to solve those problems. Emotions like curiosity/interest, determination/motivation, and relief/satisfaction for some teams, were linked to higher final grades. Teams that exhibit emotions such as collaboration/support, support/encouragement, and pride/accomplishment had on average higher final grades. These emotions reflect a positive and supportive team environment, which will contribute to better quiz performance. Emotions like playfulness/excitement and relaxation/enjoyment, while positive, did not correlate strongly with final grades. These emotions were mentioned alongside varying final grades for different teams, suggesting that other factors play a more significant role. The proportion of course-related topics discussed during team interactions have an impact on student performance. Teams with a higher percentage of course-related topics demonstrated a better understanding of the subject matter, leading to higher final grades. Collaboration/support and support/encouragement among team members may contribute to higher final grades. Teams that experience a higher percentage of relaxation/enjoyment have varying final grades, suggesting that it may not strongly correlate with student performance. The proportion of Non-Course Related Topics discussed during team interactions does not have a clear correlation with final grades.

V. Conclusion

In this study, we introduced an NLP pipeline designed to extract both the topics and the emotions expressed by students in a computer science classroom. The primary objective of this research is to develop a model that empowers educators to gain insights into students' emotional states, particularly in team setting, and to get insights about the topics discussed during their interactions.

In earlier research, we analyzed students' emotions in verbal communications and identified a correlation between positive emotions and students' performance. However, we did not investigate the specific topics discussed in these conversations. In this current work, we took a step further to determine whether the emotions expressed during verbal class discussions mainly revolve around course-related content or extend to other subjects beyond the classroom context. To achieve this, we enhanced our existing model by integrating the ChatGPT API, allowing us to generate summarized scripts from input data and subsequently applying our topic extraction model.

Preliminary findings from our analysis, based on a limited dataset, shows promising results. Overall, these findings underscore the significance of our research direction and its potential applications in educational settings. In future work, we plan to analyze larger datasets using our developed model to draw more robust conclusions. Additionally, we aim to expand the model's capabilities to analyze speech data in real-time. This enhancement will better assist instructors in promptly intervening or providing timely support and feedback to students who may require it.

References

[1] A review of affective computing: From unimodal analysis to multimodal fusion

[2] S. Suganya and E. Y. A. Charles, "Speech Emotion Recognition Using Deep Learning on audio recordings," 2019 19th International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2019, pp. 1-6, doi: 10.1109/ICTer48817.2019.9023737.

[3] W. Li, Y. Zhang and Y. Fu, "Speech Emotion Recognition in E-learning System Based on Affective Computing," Third International Conference on Natural Computation (ICNC 2007), Haikou, China, 2007, pp. 809-813, doi: 10.1109/ICNC.2007.677.

[4] Bahreini, K., Nadolski, R., & Westera, W. (2016). Towards real-time speech emotion recognition for affective e-learning. Education and information technologies, 21, 1367-1386.

[5] Feng, X., Wei, Y., Pan, X., Qiu, L., & Ma, Y. (2020). Academic emotion classification and recognition method for large-scale online learning environment—Based on A-CNN and LSTM-ATT deep learning pipeline method. International journal of environmental research and public health, 17(6), 1941.

[6] Dehbozorgi, N., & Mohandoss, D. P. (2021, October). Aspect-based emotion analysis on speech for predicting performance in collaborative learning. In 2021 IEEE Frontiers in Education Conference (FIE) (pp. 1-7). IEEE.

[7] Dehbozorgi, N., Maher, M. L., & Dorodchi, M. (2020, October). Sentiment analysis on conversations in collaborative active learning as an early predictor of performance. In 2020 IEEE Frontiers in Education Conference (FIE) (pp. 1-9). IEEE.

[8] Attota, D. C., & Dehbozorgi, N. (2022, October). Towards Application of Speech Analysis in Predicting Learners' Performance. In 2022 IEEE Frontiers in Education Conference (FIE) (pp. 1-5). IEEE.

[9] Dehbozorgi, N., Maher, M. L., & Dorodchi, M. (2021, October). Does self-efficacy correlate with positive emotion and academic performance in collaborative learning?. In 2021 IEEE Frontiers in Education Conference (FIE) (pp. 1-8). IEEE.

[10] Peter J. Rousseeuw (1987). "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis" . Computational and Applied Mathematics 20: 53–65.

[11] Caliński, T., & Harabasz, J. (1974). "A Dendrite Method for Cluster Analysis." Communications in Statistics-theory and Methods 3: 1-27.

[12] Davies, David L.; Bouldin, Donald W. (1979). "A Cluster Separation Measure" IEEE Transactions on Pattern Analysis and Machine Intelligence. PAMI-1 (2): 224-227.

[13] Halkidi, Maria; Batistakis, Yannis; Vazirgiannis, Michalis (2001). "On Clustering Validation Techniques" Journal of Intelligent Information Systems, 17(2-3), 107-145.

[14] Dehbozorgi, N. (2020). Sentiment analysis on verbal data from team discussions as an indicator of individual performance (Doctoral dissertation, The University of North Carolina at Charlotte).

[15] Dehbozorgi, N., Maher, M. L., Dorodchi, M. Emotion mining from speech in collaborative learning.