Towards Application of Speech Analysis in Predicting Learners' Performance

Dinesh Chowdary Attota
Department of Computer Science
Kennesaw State University
Marietta, GA, USA
dattota@students.kennesaw.edu

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

***Abstract*—In this work in progress, we propose a model for analysis of students' verbal conversation during teamwork to predict their academic performance based on expressed emotions. Our previous studies support the link between an individual's attitude and emotional states during the cognitive process with their performance in the given context [1], [2]. Traditionally the learners' affective states were assessed by having them fill out standard surveys. More recently the researchers have been using advanced methods to extract students' emotions from their writings by using Natural Language Processing (NLP) models. These models are applied to data collected from different sources such as discussion forums, team chats, students' reflective surveys, and journals. In this research, we take one step further by recording students audio in class as they converse about the course topic in low-stake teams and extract emotions from their conversations by NLP methods. The main contributions of the proposed model are 1) the audio transcription component 2) the multi-class emotion analysis unit and 3) the performance prediction model based on input data. SpeechBrain pre-trained models with transformer language models were applied for automated transcription of audio data and converting them to embedding vectors. NLP methods were applied for sentiment analysis. Next, we formed the feature set by combining the extracted emotions with students' formative assessment grades during the semester to implement a prediction model. We further analyzed which features in the feature set have a higher impact on the students' academic performance. The early result of this research is promising as we found high accuracy in the predicted scores of the students.**

***Index Terms*—Automatic Speech Recognition (ASR), emotion analysis, predictive model, academic performance, NLP**

I. INTRODUCTION

The application of advanced technologies for a more realistic interpretation of human speech by computers is being more popular in both academia and industrial domains. Automated Speech Recognition (ASR) is a critical component of conversational technology by which computers detect and convert spoken language into text and bring together linguistics, computer science, and other disciplines of study. Due to effective training and decoding techniques, end-to-end ASR has garnered interest as a means of directly combining acoustic and linguistic models (AMs and LMs) [3], [4]. Numerous models for ASR have been developed, including attention-based encoder-decoder architectures [5], [6], Recurrent Neural Network (RNN)-powered transducers [7], and Connectionist Temporal Classification (CTC). Transformers have taken the role of RNNs in recent years, exceeding bi-directional RNNs in terms of performance [8], [9]. Transfer learning is one of the most remarkable models and algorithms in the wide family of machine learning methods and algorithms. Transfer learning is a broad term that encompasses all strategies that use supplemental resources to enhance model learning for the target problem domain. With a high level of variation and dynamics, it is practically impossible for researchers in speech and language to train the model from a single data source [10]. We can depend on more clever algorithms that allow learning from a wide range of languages, data sets, and topics, and to constantly adapt the model.

A. Emotion Recognition:

The study of people's feelings or emotions toward a subject is called emotion recognition. It is critical to evaluate students' emotions in the educational context and notably during the collaborative learning process, in order to adapt and enhance content delivery methods [11]. Holistic evaluation of students' academic performance is critical for both students and educators, enabling them to identify the ones at risk and intervene early in order to lessen the likelihood of failure [12], [13]. When attempting to gauge a student's academic progress, formative and summative tests and assignments are often used [14]. Such evaluations provide useful information, such as trends and patterns connected to the educational process, which may be utilized to better understand the overall learning state of the students. Grade-based assessment and evaluation have some pitfalls too, especially in collaborative environments in which it's a challenge to assess an individual's contribution to the teamwork. Research suggests there is a correlation between students' emotions and their academic achievement [15]. Sentimental data such as anger, fear, joy, surprise, and sadness can be used as complementary data points to students' grades to make the evaluation process more efficient [16]. In the following sections, we discuss the related work and present the proposed model followed by a case study and data analysis.

II. RELATED WORK

Using deep neural networks to combine different types of information, such as audio and language, makes it easier for computers to recognize emotions from speech [17]. Researchers in [18] incorporate an end-to-end framework for sentiment analysis using speech recognition. In this work, they utilized a single layer multiplicative LSTM (mLSTM) [19] model with 4,096 nodes to encode the entire text input. The ASR model computes character probabilities for each frame and then extracts the final transcription (greedy decoding). They have not used any language model to rectify any spelling errors or out-of-vocabulary words. This ASR model is trained on five datasets: LibriSpeech, TED-LIUM v3, Mozilla Common Voice, and VoxForge. Each dataset contains around 1000 hours of English-read speech. The only pre-processing technique used is the conversion of recording to a WAV file using a single-channel 16-bit signed integer format with a sampling rate of 16,000. They have achieved 69.9% of accuracy.

Empathy-based interactive dialogue management is another approach presented by the authors in [20]. Emotions were extracted from raw speech input using the CNN deep learning model. In this approach, Deep Learning models like DNN-HMMs have been used to train raw audio for the acoustic model of the Kaldi speech recognition system. The sentiment of identified speech was examined using a CNN-based classifier and Word2Vec in experiments done with the TED-LIUM dataset. An accuracy of 65.7 was achieved by avoiding feature engineering approaches in voice emotion identification. When trained on domain data, CNN sentiment analysis yielded an 82.5 F-score.

On the other hand, research in [18] developed a strategy for integrating an ASR system with a character-level recurrent neural network for sentiment recognition. Earlier neural emotion detection models in the context of human-robot interaction [21] have been extended in their work. Experiments have been carried out to resolve the disparities in the performance of spoken sentiment identification when there is no transcript. They employed mLSTM (multiplicative Long Short Term Memory) to represent the entire textual input. This study combined five different freely available datasets to train the model. On the Stanford Sentimental Treebank, the model outperformed more sophisticated architectures in identifying emotion based on the next character prediction in the given context.

Use of Transformers based language models has been increased in recent years [22], [23], [24], [25]. Originally suggested for machine translation, the transformer model is an encoder-decoder based on self-attention. When applied to ASR, the transformer model has shown promising results and enhanced Word Error Rates(WERs) over RNN-based systems [26]. According to LibriSpeech test data, the suggested streaming transformer architecture in [18] produces a WER of 2.8% and 7.3%, which is the best documented streaming end-to-end ASR result for this job to the researchers' knowledge. Likewise, research work in [27] presented an approach for recognizing emotions in speech using transfer learning from automatic speech recognition. They attained a 71.7 percent accuracy for anger, excitement, sadness, and neutrality emotion content using speech data.

Automatic sentiment recognition in natural audio streams was the subject of a study published in [28] by a team of researchers. Text sentiment detection models were developed using POS tagging and Maximum Entropy Modelling (MEM). An algorithm called ME is used to estimate ratings based on text collected from reviews. It makes use of Stanford's Log-linear POS tagger for POS tagging. Switchboard and Fisher corpora have been used to train the speech recognition system, while Mel Frequency Cepstral Coefficients (MFCC) characteristics have been employed to train the acoustic model in this study. The baseline model without tuning had an accuracy of 92.1%, whereas the baseline mode, with certain parts of speech such as Nouns, had an accuracy of 94.8%.

III. Proposed Method

In this section, we discuss our proposed model for automatic speech transcription and extracting different types of emotions from transcribed speech. Fig 1 depicts the high-level architecture of the model, which is composed of a speech transcription unit, an emotion recognition unit, and a regression model to predict the students' performance based on the conversations captured from their audio input and score in-class activities.

A. Speech Transcription Unit

We used SpeechBrain [29], an open-source conversational AI framework that runs on PyTorch, to create voice transcription. SpeechBrain is a transformer-based end-to-end ASR that includes a large number of language translation models that create WAV to vector embeddings that can be used to predict text. SpeechBrain enables a user-friendly and flexible implementation of cutting-edge speech technologies such as speech recognition, speaker recognition, voice augmentation, speech

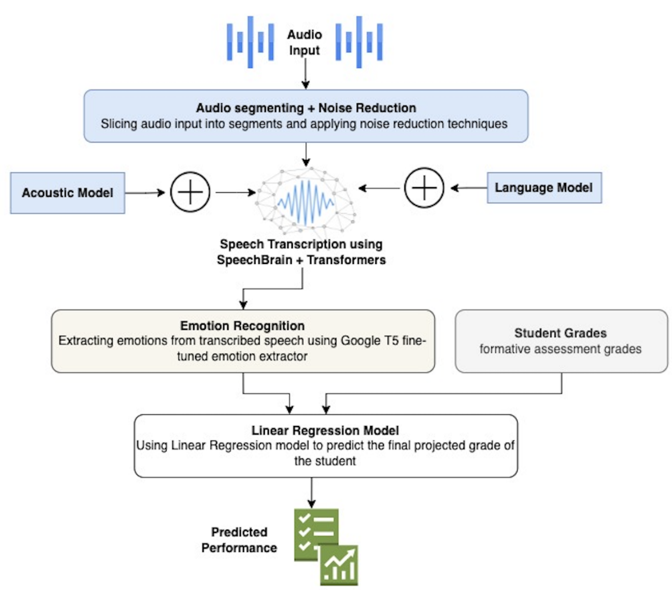

Fig. 1: High Level Architecture of the Proposed Approach

separation, language identification, and multi-microphone signal processing. This ASR has been pre-trained on the corpus of LibriSpeech [30] and is publicly available on HuggingFace. This ASR consists of three components of Unigram Tokenizer, Neural Language Model (Transformer LM), and Acoustic model with CTC.
The Unigram Tokenizer converts the words into subwords and is trained using LibriSpeech train transcriptions [31]. The tokenization starts with a large vocabulary and gradually reduces the size of the vocabulary until it reaches the target vocabulary size. A Unigram model examines each token independent of the previous one such that the probability of token X given the preceding context is just X. So a Unigram language model would always anticipate the most popular token. A token's likelihood is its frequency in the original corpus divided by the total of all tokens' frequencies in the lexicon (to make sure the probabilities sum up to 1).
The Neural Language Model consists of a deep learning model, which is trained on a dataset that consists of 10M words. The neural language model predicts the most likely sequence of words among numerous text strings. The output of the previous tokenization component will be passed to this language model to predict the probabilities of words at different timestamps [32].

Finally, the acoustic model analyzes the waveform of speech and predicts the most likely phonemes in the speech. The language model generates a matrix containing the character probabilities for each timestamp. The matrix that is generated by the neural language is decoded with the Connectionist Temporal Classification (CTC) algorithm [33]

B. Emotion Recognition

We applied Google's Text-To-Text-Transfer Transformer (T5) basic fine-tuning model to extract different emotion classes of joy, anger, fear, sadness, surprise, and love. T5 is based on the research reported in [34], which conducts a large-scale empirical survey to ascertain the most effective transfer learning approaches. Colossal Clean Crawled Corpus (C4) dataset [35], which is two orders of magnitude bigger than Wikipedia, is used to train the T5 model. The trained model is flexible enough to be used for a wide range of important downstream tasks. T5 generates text-to-text formats, where the inputs and outputs are both text strings, unlike BERT, which only outputs a class label or a span of the input. Using graph representations of the text, the pre-trained model of T5 maintains semantic relations between words in a phrase and extracts patterns. CNN uses the extracted patterns to infer the underlying emotion of the phrase they are fed into to get their predictions. The self-attention feature analyzes a sentence and places it into one of many emotion categories based on the keywords that are found in it [36]. These keywords correspond to the level of interest that a student has in a certain topic or assignment. Thus, the pre-trained T5 model enabled us to extract six distinct types of emotions from the student's conversation. In this study, the emotion of "love" is intertwined with passion and interest.

C. Linear Regression Prediction Model

We used multiple linear regression (MLR) to predict the performance of the students based on the extracted emotions combined with tests and assignment grades. Multiple linear regression figures out the relationship between two or more independent variables and a dependent variable by fitting a linear equation to the data. Each independent variable x has a corresponding value in the dependent variable y. The equation 1 represents the regression line for n observations.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_n x_{in} \quad (1)$$

The least-squares model finds the line that best fits the observed data by minimizing the sum of the squares of the vertical deviations between each data point and the line (vertical deviation is 0 when the point lies on the fitted line). The variance can be calculated using Mean Squared Error (MSE) shown in equation 2.

$$v^2 = \frac{\sum e_i^2}{n - p - 1} \quad (2)$$

D. Case study

To test the developed model we conducted a case study by collecting speech data from a CS1 active learning class [37], [38] where students worked in the same low-stakes teams throughout the semester [39]. The class met twice a week for 75 minutes. An average of 40 minutes was dedicated to teamwork in each class that we used to record students' speeches. For this purpose, a recorder with dual microphones was connected to the members of the team and the total number of 28 students were recorded during 5 weeks of the semester. Every class began with a mini-quiz on the prep material [40]. After a brief poll quiz on prep-work students were given a mini-lecture if they didn't comprehend the subject. Then, after completing a graded class activity, students were requested to complete an exit form (minute paper) separately [41]. Students were tested in 4 tests, 4 major assignments, and 4 lab tests. The students' final scores were based on their performance in the final exam as well as class and lab tests, assignments, class activities, polls, and prep work quizzes. This study uses students' final grades as the performance metric to determine if emotions and other low-stake grade data points can predict their performance.

For extracting emotions from students' speech we used audio files, each containing the recordings of conversations between two students in a team. To accommodate the embedding vectors of the transcription model in GPU memory, we segmented the original audio file into multiple parts with a duration of 30 seconds for each part. The audio segments were given as input to the SpeechBrain thereby extracting embedding vectors and passed to decoders to extract the relevant text. The extracted transcribed text of an audio chunk is appended with the transcription of previous audio chunks in each dataset. The transcriptions were passed to the Google T5 fine-tuned pre-trained model for emotion extraction. A sample

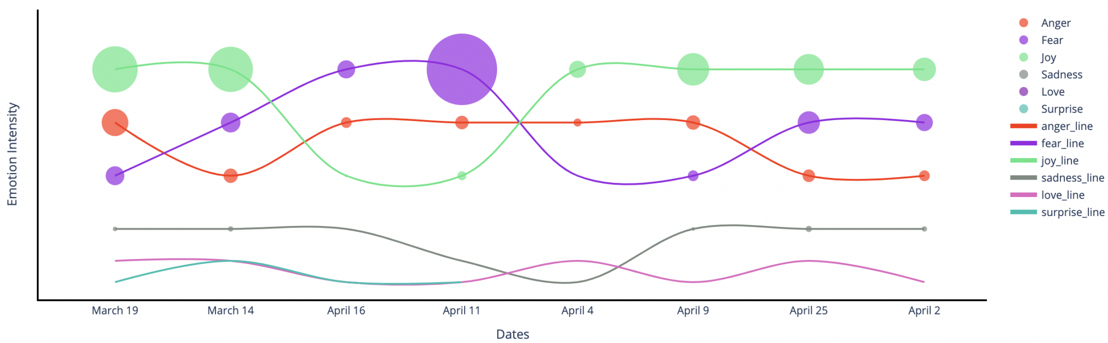

Fig. 2: Emotion Trends of a Sample Team

visualization of the trends of the emotion of a single team over the semester is presented in figure 2. In this plot, each emotion is color coded and the size of the circles reflects the intensity of the emotion in the team conversation at the specified time.

IV. Data Analysis

We evaluated if the proposed model can predict students' final grades in the course based on different data points collected throughout the semester. These data points included varying emotion classes, personality traits of the students and their peers in a team (measured by Big Five Personality tool [1], lab tests, lecture tests, assignment tests, prep-work quiz grades, and class activity scores. For the prediction model, we used the scikit-learn [42] library to implement multiple linear regression on different feature sets (combinations of data points). Our analysis showed that students' personality has the least weight in predicting the performance and the feature set of *[anger, fear, joy, surprise, sadness, love, total_number_of_words, class_activity_grade, preparation_grade, assignment_grade, lecture_test_grade, lab_test_grade]* predicted the final score more accurately. Furthermore, we implemented the variable clustering technique [43] to find out which variables in the feature set have more predictive power. Variables were divided into multiple sets of disjoint clusters. Each cluster has a linear combination of its corresponding variables. The clustering method uses $(1 - R^2)$ measure to find the most significant variable. The $(1 - R^2)_{ratio}$ is defined below in equation 3.

$$1 - R^2 = \frac{(1 - R^2)(own)}{(1 - R^2)(nearest)} \tag{3}$$

For a variable where $1 - R^2$ goes nearest to zero is the best representation of a cluster to use. When we applied the variable clustering technique to the feature set we observed that the variables of [*joy, lecture_test_grade, preparation_grade, love*] have more predictive power in predicting the students' academic performance.

The preliminary result shows that the students' positive emotions (joy and love) and their preparedness before attending the class as well as their score in the main lecture tests can determine if they will earn high final grades or not. The early result of applying our model to students' data using the selected feature set, shows impressively good performance and high accuracy. The difference between the predicted scores and the actual scores of the students ranges anywhere between -1 to 1. One reason for this high accuracy could be the limited data we had on a small sample size. More analysis should be done to come up with generic conclusions.

V. Conclusion

In this paper, we proposed a model to predict students' performance based on the emotions they express in their conversations as they worked in teams as well as their formative assessment scores during the semester. We used SpeechBrain to transcribe the recorded speech and by using a transformer-based emotion recognizer (T5) extracted the emotion classes. Data shows the performance of the model is promising as the predicted values are very close to students' actual grades. However, this work is considered as proof of concept and it's early to provide solid conclusions. The focus of this paper was to present the core functionality of the model with a limited sample size. In future work, we will work on a larger sample size of students in different classes and will be extending the functionality of the model by creating a dynamic dashboard that allows educators to visualize the various emotional patterns of the students as they work in class. This system allows instructors to adjust the content delivery pace and methods according to both educational and emotional feedback presented to them. This research has the potential to help instructors get better insights into the students' progress earlier in the semester and apply required interventions accordingly.

References

[1] N. Dehbozorgi, M. Lou Maher, and M. Dorodchi, "Sentiment analysis on conversations in collaborative active learning as an early predictor of performance," in *2020 IEEE Frontiers in Education Conference (FIE)*, 2020, pp. 1–9.

[2] N. Dehbozorgi, "Sentiment analysis on verbal data from team discussions as an indicator of individual performance," Ph.D. dissertation, The University of North Carolina at Charlotte, 2020.

[3] S. Karita, N. Yalta, S. Watanabe, M. Delcroix, A. Ogawa, and T. Nakatani, "Improving transformer-based end-to-end speech recognition with connectionist temporal classification and language model integration," 09 2019, pp. 1408–1412.

[4] Y. Wang, A. Mohamed, D. Le, C. Liu, A. Xiao, J. Mahadeokar, H. Huang, A. Tjandra, X. Zhang, F. Zhang, C. Fuegen, G. Zweig, and M. L. Seltzer, "Transformer-based acoustic modeling for hybrid speech recognition," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 6874–6878.

[5] S. Zhang, E. Loweimi, P. Bell, and S. Renals, "On the usefulness of self-attention for automatic speech recognition with transformers," in *2021 IEEE Spoken Language Technology Workshop (SLT)*. IEEE, 2021, pp. 89–96.

[6] H. Le, J. Pino, C. Wang, J. Gu, D. Schwab, and L. Besacier, "Dual-decoder transformer for joint automatic speech recognition and multilingual speech translation," *arXiv preprint arXiv:2011.00747*, 2020.

[7] J. Li, R. Zhao, H. Hu, and Y. Gong, "Improving rnn transducer modeling for end-to-end speech recognition," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2019, pp. 114–121.

[8] E. Tsunoo, Y. Kashiwagi, T. Kumakura, and S. Watanabe, "Towards online end-to-end transformer automatic speech recognition," *arXiv preprint arXiv:1910.11871*, 2019.

[9] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[10] J. Cho, M. K. Baskar, R. Li, M. Wiesner, S. H. Mallidi, N. Yalta, M. Karafiát, S. Watanabe, and T. Hori, "Multilingual sequence-to-sequence speech recognition: Architecture, transfer learning, and language modeling," in *2018 IEEE Spoken Language Technology Workshop (SLT)*, 2018, pp. 521–527.

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

[12] S. M. Jayaprakash, E. W. Moody, E. J. Lauría, J. R. Regan, and J. D. Baron, "Early alert of academically at-risk students: An open source analytics initiative," *Journal of Learning Analytics*, vol. 1, no. 1, pp. 6–47, 2014.

[13] D. Wiliam*, C. Lee, C. Harrison, and P. Black, "Teachers developing assessment for learning: Impact on student achievement," *Assessment in education: principles, policy & practice*, vol. 11, no. 1, pp. 49–65, 2004.

[14] S. M. Brookhart, "Successful students' formative and summative uses of assessment information," *Assessment in Education: Principles, Policy & Practice*, vol. 8, no. 2, pp. 153–169, 2001. [Online]. Available: https://doi.org/10.1080/09695940123775

[15] R. Pekrun, T. Goetz, W. Titz, and R. P. Perry, "Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research," *Educational psychologist*, vol. 37, no. 2, pp. 91–105, 2002.

[16] N. Dehbozorgi and D. P. Mohandoss, "Aspect-based emotion analysis on speech for predicting performance in collaborative learning," in *2021 IEEE Frontiers in Education Conference (FIE)*, 2021, pp. 1–7.

[17] V. Rozgić, S. Ananthakrishnan, S. Saleem, R. Kumar, and R. Prasad, "Ensemble of svm trees for multimodal emotion recognition," in *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, 2012, pp. 1–4.

[18] E. Lakomkin, M. A. Zamani, C. Weber, S. Magg, and S. Wermter, "Incorporating end-to-end speech recognition models for sentiment analysis," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 7976–7982.

[19] B. Krause, L. Lu, I. Murray, and S. Renals, "Multiplicative lstm for sequence modelling," *arXiv preprint arXiv:1609.07959*, 2016.

[20] D. Bertero, F. B. Siddique, C.-S. Wu, Y. Wan, R. H. Y. Chan, and P. Fung, "Real-time speech emotion and sentiment recognition for interactive dialogue systems," in *Proceedings of the 2016 conference on empirical methods in natural language processing*, 2016, pp. 1042–1047.

[21] E. Lakomkin, M. A. Zamani, C. Weber, S. Magg, and S. Wermter, "Emorl: continuous acoustic emotion classification using deep reinforcement learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 4445–4450.

[22] B. Xue, J. Yu, J. Xu, S. Liu, S. Hu, Z. Ye, M. Geng, X. Liu, and H. Meng, "Bayesian transformer language models for speech recognition," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 7378–7382.

[23] Z. Dai, Z. Yang, Y. Yang, J. Carbonell, Q. V. Le, and R. Salakhutdinov, "Transformer-xl: Attentive language models beyond a fixed-length context," *arXiv preprint arXiv:1901.02860*, 2019.

[24] K. Irie, A. Zeyer, R. Schlüter, and H. Ney, "Language modeling with deep transformers," *arXiv preprint arXiv:1905.04226*, 2019.

[25] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[26] S. Karita, N. Chen, T. Hayashi, T. Hori, H. Inaguma, Z. Jiang, M. Someki, N. E. Y. Soplin, R. Yamamoto, X. Wang *et al.*, "A comparative study on transformer vs rnn in speech applications," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. IEEE, 2019, pp. 449–456.

[27] S. Zhou and H. Beigi, "A transfer learning method for speech emotion recognition from automatic speech recognition," *arXiv preprint arXiv:2008.02863*, 2020.

[28] L. Kaushik, A. Sangwan, and J. H. L. Hansen, "Sentiment extraction from natural audio streams," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 8485–8489.

[29] M. Ravanelli, T. Parcollet, P. Plantinga, A. Rouhe, S. Cornell, L. Lugosch, C. Subakan, N. Dawalatabad, A. Heba, J. Zhong, J.-C. Chou, S.-L. Yeh, S.-W. Fu, C.-F. Liao, E. Rastorgueva, F. Grondin, W. Aris, H. Na, Y. Gao, R. D. Mori, and Y. Bengio, "SpeechBrain: A general-purpose speech toolkit," 2021, arXiv:2106.04624.

[30] V. Panayotov, G. Chen, D. Povey, and S. Khudanpur, "Librispeech: An asr corpus based on public domain audio books," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 5206–5210.

[31] T. Kudo, "Subword regularization: Improving neural network translation models with multiple subword candidates," *arXiv preprint arXiv:1804.10959*, 2018.

[32] K. Li, Z. Liu, T. He, H. Huang, F. Peng, D. Povey, and S. Khudanpur, "An empirical study of transformer-based neural language model adaptation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 7934–7938.

[33] H. Scheidl, S. Fiel, and R. Sablatnig, "Word beam search: A connectionist temporal classification decoding algorithm," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 2018, pp. 253–258.

[34] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *arXiv preprint arXiv:1910.10683*, 2019.

[35] J. Dodge, M. Sap, A. Marasovic, W. Agnew, G. Ilharco, D. Groeneveld, and M. Gardner, "Documenting the english colossal clean crawled corpus," *arXiv e-prints*, pp. arXiv–2104, 2021.

[36] E. Saravia, H.-C. T. Liu, Y.-H. Huang, J. Wu, and Y.-S. Chen, "Carer: Contextualized affect representations for emotion recognition," in *Proceedings of the 2018 conference on empirical methods in natural language processing*, 2018, pp. 3687–3697.

[37] N. Dehbozorgi, S. MacNeil, M. L. Maher, and M. Dorodchi, "A comparison of lecture-based and active learning design patterns in cs education," in *2018 IEEE Frontiers in Education Conference (FIE)*, 2018, pp. 1–8.

[38] N. Dehbozorgi, "Active learning design patterns for cs education," in *Proceedings of the 2017 ACM Conference on International Computing Education Research*, 2017, pp. 291–292.

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

[40] M. Maher, N. Dehbozorgi, M. Dorodchi, and S. Macneil, "Design patterns for active learning," *Faculty Experiences in Active Learning: A Collection of Strategies for Implementing Active Learning Across Disciplines*, pp. 130–158, 2020.

[41] N. Dehbozorgi and S. MacNeil, "Semi-automated analysis of reflections as a continuous course," in *2019 IEEE Frontiers in Education Conference (FIE)*, 2019, pp. 1–5.

[42] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[43] R. Sanche and K. Lonergan, "Variable reduction for predictive modeling with clustering," in *Casualty Actuarial Society Forum*. Citeseer, 2006, pp. 89–100.