Speech Fluency Analysis : Speech-Based Emotion Enhanced Learning and Assessment System

C. Mithul
Department of Computer Science
Amrita School of Computing,
Amrita Vishwa Vidyapeetham,
Bengaluru, India.
BL.EN.U4AIE21034@bl.students.amrita.edu

D. Mohammad Abdulla
Department of Computer Science
Amrita School of Computing,
Amrita Vishwa Vidyapeetham,
Bengaluru, India.
BL.EN.U4AIE21044@bl.students.amrita.edu

Vaan Amuthu Elango
Department of Computer Science
Amrita School of Computing,
Amrita Vishwa Vidyapeetham,
Bengaluru, India.
e_vaanamuthu@blr.amrita.edu

Abstract **– The Language learning is a hard and difficult thing that needs a clear knowledge of the target language. Judging the fluency level of non-natives English speakers is the first and the most important step in their language learning process. Nevertheless, it is a long and boring thing, which is also a subjective matter and needs the skills and practical experience. In this paper we offer the idea of the automation of the speaker fluency evaluation by the means of machine learning. We have created a database of annotated audio conversations in English between people of different levels of fluency. We have divided the audio conversations into 5s non-overlapped audio segments and have extracted features from them. Thus, we have cross-trained five different machine learning models to recognize the level of fluency of non-native English speakers. Our results prove that the model that performs the best is the Support Vector Machine (SVM) with a good accuracy of 94%. The 39% models have not managed to get a good accuracy, but the rest of the models have performed better with an accuracy of 89% or higher. Our project presents a new innovative technique for the automatic speaker evaluation of the speaker fluency level, which can be further enhanced by adding more features and by researching different machine learning algorithms.**

Keywords: Speech processing, Speech Fluency, adaptive learning, speech analytic, language learning, deep learning, machine learning, CNN, SVM, RNN, RF, MLP, signal processing, voice transcription, natural language processing, Language proficiency

I. INTRODUCTION

The skill to communicate effectively in a foreign language or second language is very significant in the modern world which is interconnected. Nevertheless, it can be a great task to become fluent in another language and this requires much practice and correction. Traditional methods of learning languages involve face-to-face teaching or personal tutoring which are time-consuming, costly and do not offer enough chances for individualized support as well as practice.

Language learning has been improved by current developments in artificial intelligence (AI) and speech processing technology. These systems can evaluate how fluently a learner speaks by analyzing their speaking patterns automatically then give recommendations and feedbacks customized towards them. Besides, emotional intelligence may also be incorporated into these systems so that they adapt better according to what they detect about learners' emotions from their spoken words.

First, we have built our own audio set (Fluency audio set), the details of the audio set are given in section 2. In our case, to face the general problem of fluency level monitoring of each individual during a conversation, we have built our own audio set. After that, each conversation has been split into 5 non-overlapped segments and these segments have had some features extracted (the mel coefficients + the zero crossing rate + the root-mean-square-energy + the spectral flux). Finally, the feature vectors are being fed into a classification model which is being trained for its performance assessment by means of accuracy metrics. Our defined fluency classes are three: three levels of fluency, namely the low fluency, the intermediate fluency, and the high fluency.

We have studied five ML models, i. e, multi-layer perceptron (MLP), support vector machines (SVM), random forest (RF), convolutional neural networks (CNN) and recurrent neural networks (RNN). The workflow discussed above is the one that is usually used in the context of sound events audio analysis in the conventional ML approach.

The main hypotheses that we are trying to answer here is: Considering an audio set with labels and balanced audio set meeting the prescribed fluency criteria, can we develop a model that will classify an audio segment according to the level of fluency?If that is the case, then we will be able to know whether a group in a conversation needs a suggested subject to make the conversation flow. Furthermore, we could determine if Person A has a lower fluency than Person B and needs to be shifted to a lower fluency level group (as well as the case when the opposite is true).

Our last results of the model classification have managed to get the accuracies higher than 90% (except for one model), and they are the best of the accuracies up to 94%. 39% for an SVM. For the start, we have changed the number of the Mel coefficients (MFCCs) extracted to make sure that the accuracies will be high. After that, we proofed that modifying the basic MFCCs with the addition of the zero-crossing rate (ZCR), root-mean-squared-energy (RMSE) and spectral flux onset strength envelope (SF) has improved the model performance.

II. Literature Survey

Litman, D. et al. [1] we presents a comprehensive overview and assessment of the current and potential uses of Automated Speech Recognition (ASR) and Spoken Dialogue Systems (SDS) in second language speaking assessment. These technologies have evolved from primarily aiding language learning to being increasingly utilized for grading purposes. The effectiveness of tests utilizing these technologies depends on how the speaking construct is defined, the specific contexts in which the tests are used, and the challenges and future research needs identified.

The authors highlight the importance of defining the speaking construct accurately to evaluate the adequacy of tests utilizing ASR and SDS technologies. They also emphasize the need for interdisciplinary research collaboration among language testing and speech technology experts to further explore the opportunities for integrating these technologies into assessment practices. Overall, the paper suggests that there are significant opportunities for leveraging ASR and SDS technologies in language assessment, but further research and collaboration are essential to fully realize their potential.

K. Soong, et al. [2] presents a couple of revolutionary methods intended to improve the chances of mispronunciations detection. Firstly, we use a GMM-HMM (Gaussian Mixture Model-Hidden Markov Model) based acoustic models for calculating GOP (Grade of Pronunciation). To make the discrimination better, this model is additionally refined by training deep neural network (DNN). In the next step, the fundamental frequency (F0) information is implicated into a DNN-based model to detect mispronunciations due to incorrect applications of lexical stress or tone particularly in the context of second language acquisition, e.g., when learning L2 language words like "cookies". Through F0 integration feature to DNN model we intend to increase accuracy of detecting pronunciation mistakes.

In concluding, we suggest a logistic regression classifier that is based on a neural network for streamlining the training process as well as improving the generalization capabilities. The purpose of this classifier is to enable the dispensing of multiple individual classifiers without having to train them separately, which leads to the optimization of efficiency and effectiveness. Experiments with multiple learning corpora prove the usefulness of our proposed methodology in reducing mispronunciation detection and boosting detection accuracy in general Innovations such as these help us to develop new ways of recognizing and rectifying the pronunciation errors, particularly the language learning and testing processes.

R. Gretter, et al. [3] an automatic system to grade English and German proficiency of Italian students aged 9-16. Students' answers are transcribed by multilingual automatic speech recognition models adapted for non-native child speech. Features extracted from the transcripts and acoustic models are input to neural networks trained to predict human expert scores on proficiency indicators like syntax, pronunciation, fluency, etc. Reasonable correlation of 0.53-0.61 is achieved between the automatic and human scores.

The classification accuracy is about 60-67%. Feature selection, regression models, and utilizing techniques from ASR quality estimation can be used to further improve the accuracy. Limitations are the use of single human raters and lack of broad bench marking.

S. M. Siniscalchi, et al. [4] presents a new approach that illustrates the complimentary nature of speech attribute features correlated with the traditional phone features for speech error detection. The system accomplishes this by using these features in the input representation for Long Short-Term Memory (LSTM) classification units, thus achieving more accurate speech error detection. The components of the system are also robust to phone label-level noise in non-native speech data. The proposed framework performs very well, particularly when it comes to dynamic modeling of frame-level pronunciation scores that leads to substantial FAR and FRR reductions concerning systems that use only phone features (decreases of 37.90% and 38.44%, respectively).

Next steps in research are described where a Mandarin vowel mispronunciation detection system will be implemented with tongue position and lip roundness feedback mechanisms. Furthermore, developing multi-way feedback features, like merging the information with the decision trees and using detection curves, is acknowledged as one of the possible future investigations and improvement steps.

Zezario, R. E et al. [5], does a analysis machine learning techniques in assessing aphasia-related speech impairment. It covers aphasia's impact on communication, assessment methods including feature extraction and classification, machine learning algorithms such as MFCC and CNN, and their application in automatic speech assessment. It shows assessment strategies, neural network models, and the relationship between cognitive function and speech evaluation. This survey shows the importance of automatic assessment systems in evaluating speech impairment for research and clinical applications.

Budde, K.S, et al. [6] based on the fundamental contribution of Brown et al. (2005), where they described subcortical regions as "neural signatures of stuttering" following their meta-analysis about the functional neuroimaging studies about PWS population. Brown and his colleagues located not only the excessive activity in the right inferior premotor cortex and cerebellum but also the underactivation in auditory cortex as possible neural surrogates. This study focuses on this proposition with a bigger data group (17 papers taken into account compared to only 8 papers were considered in the Brown et al.) and more reliable meta-analytic techniques. It complements what happened before, drawing the researchers' attention to the fact that those changes also are seen in the SMA.

Along with these, the researchers also explain the unique spatiotemporal patterns related to trait stuttering (between-group differences between PWS and fluent controls) and state-stuttering (within-group differences between dysfluent and induced fluent speech in PWS). This facet of the subject was not noticed by the prevailing literatures. This paper holds that the discussed neuron signatures may benefit the elaboration of the input-output of data-driven network models and serve as the markers for stuttering genes discovery in research, but no particular schemes or metrics offered. The essay stresses the necessity for large-scale studies and meta-analyses used as compared to narrow-adherence to the sample case in some recent arguments.

Amir, O., et al. [7] made using 277 listeners of 4 groups (naïve speakers, speech-language pathology students, practicing speech-language pathologists, and speech-language pathologists specializing in stuttering). Listener rating scores were assessed using 15-second speech samples from 56 speakers (20 people who stutter, 16 people who stutter using fluency-shaping techniques, and 20 non-stuttering speakers) using the 11-point stuttering severity scale. Furthermore, SES and %SS indices were also calculated for each sample of speech. The four listener groups provided equal measures of stuttering severity; meanwhile, there were significant differences between the three speaking groups. SES resulted in distinctive speaker groups, while %SS grouped the stuttering group together with the other two.

Authors discovered a statistically significant positive correlation ($r = 0.92$) between socio-economic status and the subjective rating of stuttering severity scales. The association between %SS and ratings was also seen to be lower ($r = 0.81$) as was the association between %SS and SES ($r = 0.80$. The authors posit that the SES is a dependable, objective, and likely automatic technique for assessing speech intelligibility that is also not influenced by listener prejudice. Although SES was calculated manually in this research, the authors admit that the current trends on development of algorithms for automated segmentation and SES calculation is going on.

Wang, X. [8] targeted at the issue of identifying fluency measures that can be used as a basis for the development of an automated machine-learning system for the quality assessment of English-Chinese consecutive interpreting. It examines the past research of interpreting both regarding measuring the fluency of interpreted material and assessment of spoken language. The current fluency measures involve phonation time ratio, speech rate, articulation rate, average length of runs and mean pauses (Towell et al., 1990). Fluency is measured by breakdown fluency (filled/unfilled pauses); speed fluency (mode of speech); and repair fluency (reformulations; false starts) (Tavakoli & Skehan, 2005). In interpreting studies, filled/unfilled pauses can be manifestations of disfluency popularized by (Pöchhacker, 2016). Acoustic features, such as speech rate, articulation rate, phonation ratio and durations of pauses were found to correlate with fluency ratings by people (Han et al., 2020; Yu and van Heuven, 2017).

Our work postulates new fluency parameters which can be extracted by statistical analysis of interpreting data. This statistic confirms the mean value of 0.25 of seconds as the cut-off for unfilled pauses since 80% of the values fall below this level. Defining outliers is the case of the exceptionally long/slow versus the non-exceptionally similar long/slow instances. The regression models based on the parameters we obtained have explained from 48 to 83.6% of the variance of the variables we have used the gold standard for the assessment of fluency, namely articulation rate and phonation ratio. A penny-wise model about two additional parameters (mean length of filled pauses, number of especially long unfilled pauses) explains 16.6 percent of human fluency ratings variance, outperforming models about only established measures (6.6 percent of human fluency ratings variance explained). The study figurers out the number of filled and unfilled pauses, the articulation speed and the pause lengths as key features that can be used in machine learning models for building automated fluency assessment systems.

III. Methodology

A dataset called the " Avalinguo audio set" was used to perform speech fluency analysis. It consists of fast, unscripted conversation recordings between non-native English speakers with different fluency levels, collected from sources like friends/family, language learning centers, and online videos. These audio recordings were segmented into 5-second non-overlapping clips which resulted in a total of 1,424 audio clips. Each clip was labeled into one of three fluency levels - low, intermediate, or high fluency - based on defined fluency metrics which examines factors like pausing, hesitation, and speech naturalness.

Acoustic features were extracted using the librosa audio processing library from 5-second audio clips,. The Mel-frequency cepstral coefficients (MFCCs) were computed, which captured short-term spectral envelope shape of the audio signal.Other features like zero-crossing rate, root mean square energy, and spectral flux onset strength were also extracted. These features help model temporal and energy characteristics of the speech signal relevant for fluency assessment. To determine the acoustic features, experiments were run to assess model performances when varying: 1) The number of MFCCs extracted, 2) Adding features to the MFCC baseline.

Five different machine learning models are used for fluency classification task using the extracted acoustic features as input and the manually labeled fluency levels as targets: Support Vector Machines (SVMs), Random Forests (RFs), Multilayer Perceptrons (MLPs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs - LSTMs). The dataset was split into 70% training and 30% test sets. The models were trained on the training set and their classification performance was evaluated on test set in terms of accuracy metrics. For neural network models, architectures were designed with 2-4 hidden layers and appropriate activation functions. Models like SVMs and RFs were constructed using standard implementations with tuned hyperparameters.

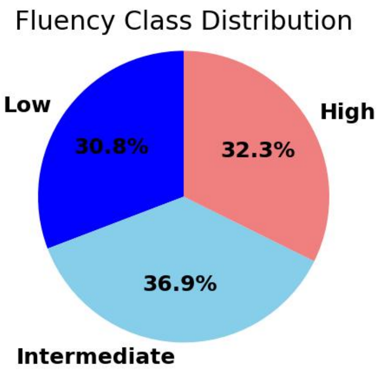

Figure 1: The Fluency audio set is distributed according to class. There are 118.65 minutes of recorded audio in this collection.

Experimental Framework

The procedure we designed for sound event analysis is based on the same machine learning (ML) approach as the standard one. The process involves three primary steps: feature extraction, classification, and output of the predicted label with higher probability for individual segments (frames) are the three main steps in the automatic speech recognition process. At first, we extract the features of each audio frame which, later on, is used as input to a classifier that gives the most probable class based on the previous training. This method of audio analysis with ML is the same as the general steps of audio analysis with ML, which includes getting and preparing the audio data, extracting the audio features, selecting and training the ML model, and finally using the trained model for the prediction.

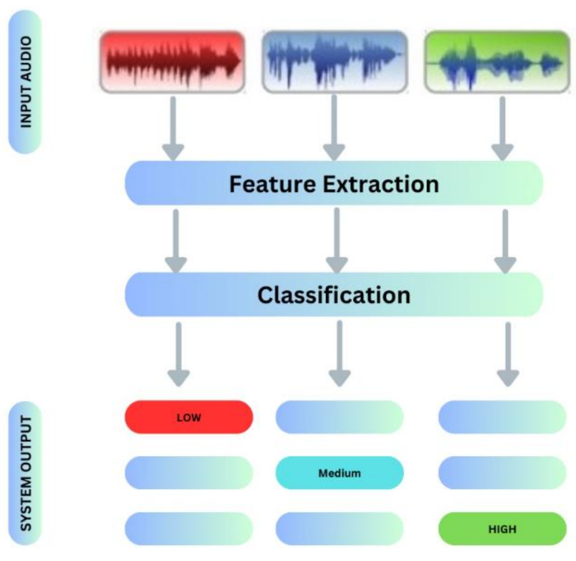

Figure 2: Pipeline of an audio classification system.

Audio segmentation

The audio segmentation is a significant preprocessing stage in many audio applications, for example, speech recognition. The process of splitting an audio stream into homogeneous regions, where each region has a unique acoustic feature, is what characterizes the segmentation. The aim is to the regions that are different in nature to be dealt with in a different way, for instance, the separation of music from noise or speech from non-speech.

In our method, we break down audio files into 5-second non-overlapping segments and then, by ourselves, assign each segment to one person. If a section has more than one speaker, we allocate it to the person who speaks the most in that section. We then classify each segment into the groups on the basis of its fluency level. We made use of the Pydub Python module to divide the audio files into sounds.

In addition to this, we examined the use of voice activity detection (VAD), which generate audio segments when it recognizes a different voice or a pause in the conversation. On the other hand, VAD eliminates the silence from the conversation, which is not the case for us when we are trying to detect the silences and pauses in each audio segment. Consequently, we abandoned this strategy.

Audio segmentation is the main step in many audio applications such as speech recognition, music analysis and content indexing. It presupposes the determination of the limits between the different sounds or events in an audio stream, that is, the separation of them into the separate segments. This process is vital for a lot of things, including speaker identification, automatic speech recognition, and automatic music analysis.

Feature Extraction

In this step, we take the features from the audio segments that have been created with the Python package Librosa. Librosa is a very popular library for audio feature extraction, hence it is feasible to get over thirty audio features. We have developed a Python script for the feature extraction, which means that we have experimented with different numbers of Mel-Frequency Cepstral Coefficients (MFCCs) extracted to get the best number of MFCCs. Besides, we have introduced Zero-Crossing Rate (ZCR), Root Mean Square Energy (RMSE), and Spectral Features (SF) to make sure that this will improve the model performance.

Librosa for Feature Extraction

Librosa is a good instrument for extracting the features, thus we can compute a lot of audio features. We have used Librosa to extract the features from our audio segments and these features are MFCCs, ZCR, RMSE, and SF. Librosa has the capability of displaying the audio spectrograms using its display function. (shown in Figure 3) The specshow function, which we have been using to output our audio data.

Audio feature extraction is the first and the major phase of audio analysis, and Python has many libraries that can be used for this purpose. Librosa is one of the most widely used libraries for the extraction of audio features and we have used it to extract the features from our audio segments. The extracted features can be used in many areas, for instance, the classification, prediction and recommendation algorithms.

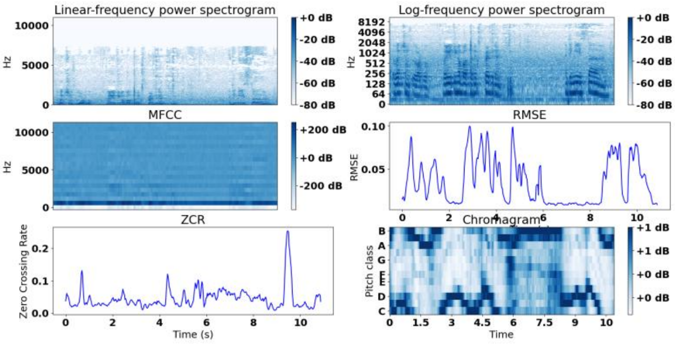

Figure 3: Spectral feature plots (single frame).

Audio Classification

The task that appears to be so simple and common is the foundation of many applications such as music information retrieval, audio tagging, and real-time audio processing. The principal aim of this project is to compare the efficiency of several classification frameworks like MLPs, CNNs, RNNs, SVMs, and RFs in the context of the audio classification tasks.

The MLP architecture which is being proposed has two hidden layers with 512 neurons in each of them and, in the end, the output layer has three neurons which are the three fluency classes. The CNN configuration has four hidden layers, the first two of which have 64 convolution filters and the other two have 32 filters. The LSTM is the network that has two hidden layers and three output neurons which is the RNN.

Neural Network	Hidden layers	Neurons	Activation
MLP	2	512x512x3	relu, softmax
CNN	4	64x64x32x32x3	relu, softmax
RNN	2	256x32x3	relu, softmax

Table 1: Neural networks architectures.

The latest findings in the field of deep learning have revealed a good future for audio applications, thus the tasks like music transcription, beat detection, and real-time audio processing have become more accurate through the use of these algorithms. The essay of the four inference engines that can be used for the on-the-fly audio classification on the CPU of an embedded single-board computer were presented, which are TensorFlow. The research has shown that all the inference engines can run the neural network models in real-time with the custom code but the execution time varies between the engines and the models.

Experimental Results

This section of our report will be dedicated to the fusion of all the outcomes from feature extraction and classifications. We tried different data exploring instruments, for instance, the Anaconda one which worked through Python 2. In addition to (1) Keras (backed up by TensorFlow), (2) Scikit-learn, (3) Librosa, (4) Pydub, and (5) Pandas, we will use such libraries to carry out our experiments. We concluded which models were the most effective. As this paper shows in Table 1, these were the architectures and hyperparameters that came out on top, based on the classification accuracies.

Initially in the analysis of the cycle accuracy determination, we looked at the influence of changing the number of Mel coefficients (N_mel) on the result. Rising N_mel produces a model that is quite complicated, which makes it imperative that people should discern the contradiction point of that and the maximum achievable accuracy. In our analysis area we verified that there exists the phenomenon of diminishing returns when having more Mel filters. Increasing the number of Mels results in minor improvements and may also get worse already. The experiment helped us to figure out the Purple Melin for the dataset.

In addition, we illustrated how appending the features including the ordinary zero-crossing rate (ZCR), the root-mean-square energy (RMSE), and the mel-spectral features (SF) to the base Nmel increases the accuracy in almost all instances. It can be explained through the notion of feature extraction, a process of turning raw data in numerical features that are analyzable without loss of information in original data. An efficient feature extraction is of primary importance for any machine learning and deep learning applications to achieve good performance.

Model	5	10	12	20
SVM	86.00%	89.49%	92.06%	94.39%
RF	84.80%	89.00%	90.42%	92.29%
MLP	78.00%	88.78%	89.01%	92.05%
CNN	80.00%	85.04%	87.61%	93.69%
RNN	78.90%	85.04%	86.44%	87.00%

Table 2: Accuracy performance of the classification models for different *N_mel* values.

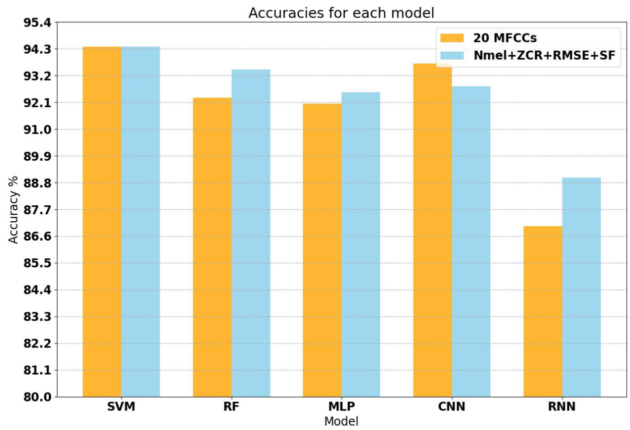

Figure 4: Accuracy comparison when using *Nmel* = 20 (orange bars) and with extra features (skyblue bars). For sake of visualization, the accuracy of the plots starts at 8%.

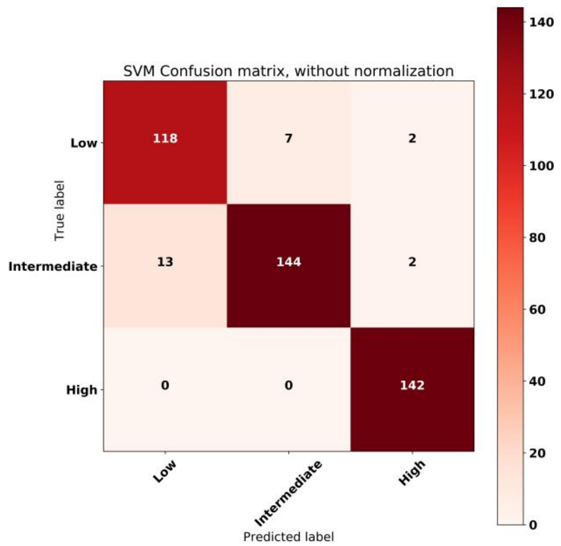

Figure 5: Confusion matrix for the SVM trained and tested with the Fluency audio set.

IV. Conclusion

The study offered speech system which is based on conversation fluency classification and emotional state identification for an adaptive virtual language learning environment. The program can adapt teaching content, feedback methods, difficulty levels as well as learning activities by assessing automatically the level of fluency in speaking from prosodic features exhibited during a dialogue with other individual and recognizing emotions such as boredom or excitement among others indicated by this person at different times. This is achieved through appraising how well one talks that is non-native English speakers in relation to what they say about their feelings when talking too much or too little during these conversations at various points in time while using different words. A feasibility test carried out on a unique dataset of speeches made by people who are not native English users demonstrated that it was possible to do so using multiple modes of communication, where support vector machines turned out to be quite accurate in terms of determining whether someone was fluent or not. What makes intelligent tutoring systems adaptive here is their ability to combine speech processing technology with machine learning algorithms plus affective computing techniques thus fostering motivation among learners, promoting sustained involvement throughout lessons and speeding up the process towards achieving proficiency in foreign languages. Such works indicate exciting prospects for applying AI within language education so that each student could get personalized support through online platforms where they interact via avatars representing themselves.

V. References

[1] Litman, D., Strik, H., & Lim, G. S. (2018). Speech technologies and the assessment of second language speaking: Approaches, challenges, and opportunities. Language Assessment Quarterly, 15(3), 294-309.

[2] W. Hu, Y. Qian, F. K. Soong, and Y. Wang, “Improved mispronunciation detection with deep neural network trained acoustic models and transfer learning based logistic regression classifiers,” Speech Communication, vol. 67, pp. 154–166, 2015.

[3] R. Gretter, M. Matassoni, K. Allgaier, S. Tchistiakova and D. Falavigna, "Automatic Assessment of Spoken Language Proficiency of Non-native Children," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019, pp. 7435-7439,

[4] W. Li, N. F. Chen, S. M. Siniscalchi, and C.-H. Lee, “Improving mispronunciation detection for non-native learners with multisource information and lstm-based deep models,” in Proc. Interspeech 2017, 2017, pp. 2759–2763.

[5] Zezario, R. E., Fu, S. W., Chen, F., Fuh, C. S., Wang, H. M., & Tsao, Y. (2022). Deep learning-based non-intrusive multi-objective speech assessment model with cross-domain features. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31, 54-70.

[6] Budde, K.S., Barron, D.S. and Fox, P.T., 2014. Stuttering, induced fluency, and natural fluency: a hierarchical series of activation likelihood estimation meta-analyses. Brain and language, 139, pp.99-107.

[7] Amir, O., Shapira, Y., Mick, L. and Yaruss, J.S., 2018. The Speech Efficiency Score (SES): A time-domain measure of speech fluency. Journal of fluency disorders, 58, pp.61-69.

[8] Wang, X. and Wang, B., 2024. Identifying fluency parameters for a machine-learning-based automated

interpreting assessment system. Perspectives, 32(2), pp.278-294.

[9] H. Huang, H. Xu, Y. Hu, and G. Zhou, “A transfer learning approach to goodness of pronunciation based automatic mispronunciation detection,” The Journal of the Acoustical Society of America, vol. 142, no. 5, pp. 3165–3177, 2017.

[10] Busse, V., Jungclaus, J., Roden, I., Russo, F. A., & Kreutz, G. (2018). Combining song—and speech-based language teaching: An intervention with recently migrated children. Frontiers in Psychology, 9, 2386.

[11] K. R. Jothi and V. L. Mamatha, "A Systematic Review of Machine Learning based Automatic Speech Assessment System to Evaluate Speech Impairment," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp. 175-185, doi: 10.1109/ICISS49785.2020.9315920.

[12] M. A. Zissman, “Comparison of four `approaches to automatic language identification of telephone speech,” IEEE Transactions on speech and audio processing, vol. 4, no. 1, p. 31, 1996

[13] Cao, Dongbo, and Ying Guo. "Algorithm research of spoken English assessment based on fuzzy measure and speech recognition technology." International Journal of Biometrics 12, no. 1 (2020): 120-129..

[14] W. Hu, Y. Qian, and F. K. Soong, “A new dnn-based high quality pronunciation evaluation for computer-aided language learning (call).” in Interspeech, 2013, pp. 1886–1890

[15] K. Zechner, J. Sabatini and L. Chen, "Automatic scoring of childrens read-aloud text passages and word lists", *Proc. of the Fourth Workshop on Innovative Use of NLP for Building Educational Applications*, 2009.

[16] Pati, Peeta Basa, and V. Shreyas. "Speech to equation conversion using a poe tagger." In *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, pp. 1-4. IEEE, 2022.

[17] Pattanayak, Satyabrata, Peeta Basa Pati, and Tripty Singh. "Performance Analysis of Machine Learning Algorithms on Multi-Touch Attribution Model." In *2022 3rd International Conference for Emerging Technology (INCET)*, pp. 1-7. IEEE, 2022.

[18] Peeta Basa Pati, S. Raju, S., Pati, N., and Ramakrishnan, A. G., “Gabor filters for document analysis in Indian bilingual documents”, International Conference on Intelligent Sensing and Information Processing, 2004. Proceedings of. pp. 123-126, 2004.