SSRG International Journal of Computer Science and Engineering
ISSN: 2348–8387 / https://doi.org/10.14445/23488387/IJCSE-V10I12P101

Volume 10 Issue 12, 1-4, December 2023

Original Article

Pronunciation Error Detection and Correction

Vadlamudi Hari Priya[1], Yenuganti Rama[2], Purimetla Srimati[3], Shaik Shahanaz[4]

[1,2,3,4,5]*Department of Information Technology, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India.*

[1]*Corresponding Author : ramayenuganti2002@gmail.com*

Received: 01 October 2023 Revised: 03 November 2023 Accepted: 26 November 2023 Published: 09 December 2023

Abstract *- Accurate pronunciation plays a pivotal role in language learning and communication. This paper presents a comprehensive overview of the field of pronunciation error detection and correction. It explores various techniques, including Automatic Speech Recognition (ASR) and Natural Language Processing (NLP), to identify and correct pronunciation errors. The paper delves into the challenges associated with this task, such as accent diversity, non-native speakers, and contextual variations. Additionally, it discusses the potential applications in language education, speech therapy, and language assessment. This paper aims to contribute to developing more effective tools and systems for improving pronunciation and language proficiency.*

Keywords *- Speech recognition, Accent diversity, Speech therapy, Language assessment.*

1. Introduction

Pronunciation error detection and correction is a technology that aims to identify and rectify mispronunciations and speech errors made by individuals when speaking a language. This technology uses speech recognition and natural language processing to detect the pronunciation of words or sentences. It captures audio using the microphone as input and detects the error by comparing it with the audio present in the dictionary. It enhances word pronunciation by providing correct word suggestions. It can be used in language learning, language assessment and speech therapy to help individuals improve their pronunciation and fluency.

Using a microphone as an input source, this technology captures spoken audio and employs advanced algorithms to analyze the pronunciation of individual words or entire sentences. It compares the spoken audio against a comprehensive dictionary, such as the CMU Pronouncing. Dictionary to detect deviations from the expected pronunciation by pinpointing errors; this technology provides valuable feedback and suggestions to improve pronunciation accuracy. It offers corrective guidance by suggesting the correct pronunciation of words and phrases, thereby aiding in language learning, language assessment, and speech therapy. Additionally, it assists individuals in enhancing their overall fluency and articulation.

Furthermore, this technology can be integrated into various applications and platforms, including language learning software, educational tools, and speech therapy applications, making it accessible and beneficial across diverse learning environments. Its adaptability and ability to cater to individual learning needs make it a valuable resource in fostering better communication skills and linguistic proficiency.

1.1. Data Analysis

Pronunciation error detection and correction technology employs advanced speech recognition and natural language processing to identify and rectify spoken language errors. It detects mispronunciations by comparing audio input with comprehensive dictionaries like CMU Pronouncing Dictionary, offering corrective suggestions to enhance pronunciation accuracy. Data analysis can focus on error frequency, algorithm performance, and user proficiency, evaluating the technology's effectiveness in improving pronunciation. Comparative studies across languages or dialects and assessing the impact on language learning outcomes can provide valuable insights. Usability studies gauge user satisfaction and usability, guiding enhancements for a more effective learning experience. Overall, this analysis delves into error patterns, technology efficacy, and user experiences to optimize pronunciation improvement tools.

2. Problem Statement

Develop a real-time pronunciation feedback system that captures a user's spoken input, identifies pronunciation errors using the CMU Pronouncing Dictionary, and provides immediate suggestions for correct pronunciation. The system should operate in a continuous loop, allowing the user to receive feedback and make corrections. The primary objectives include:

1. *Speech Input:* Implement a mechanism to capture spoken input using a microphone.
2. *Pronunciation Detection:* Use the CMU Pronouncing Dictionary to identify pronunciation errors in the user's input.
3. *Correction Suggestions:* Provide suggestions for correcting pronunciation errors through text-to-speech.
4. *User Interaction:* Display feedback to the user, indicating whether the pronunciation was correct or if there were detected errors.
5. *Continuous Operation:* Allow the system to run continuously, enabling users to make multiple attempts and receive feedback for each.

Ensure that the system is user-friendly, responsive, and capable of handling a variety of spoken words. Consider edge cases, such as words not present in the CMU Pronouncing Dictionary, and handle them gracefully.

3. Materials and Methods

3.1. Materials

3.1.1. Pyttsx3

This is a Python library for text-to-speech conversion, allowing the system to provide audio suggestions for pronunciation correction.

3.1.2. Speech_recognition

It provides the ability to capture the user's spoken input through a microphone.

3.1.3. NLTK (Natural Language Toolkit)

Used for working with natural language data.

3.1.4. Cmudict

Part of NLTK it contains the CMU Pronouncing Dictionary, which is used to map words to their phonetic representations.

3.2. Methods

3.2.1. Initialization

- The script initializes the speech recognizer (recognizer).
- It downloads and initializes the CMU Pronouncing Dictionary.
- It defines a function (speak) for text-to-speech conversion.

3.2.2. Capture_input() Function

- Captures the user's spoken input through a microphone.
- Adjusts for ambient noise to improve speech recognition.
- Uses the Google Web Speech API to recognize and transcribe the user's input.

3.2.3. Detect_and_correct_pronunciation() Function

- Takes user input as text and splits it into words.
- Check if each word is in the CMU Pronouncing Dictionary.
- If a word is in the dictionary, it replaces the word with its phonetic representation; if not, it keeps the original word.

4. Results and Discussions

4.1. Results

- Pronunciation Error Detection and Correction is capable of capturing spoken input from users and analyzing it for pronunciation errors in real time.
- When a user speaks a word, the script transcribes the spoken word and compares it to the CMU Pronouncing Dictionary to detect and correct pronunciation errors.
- If the user's pronunciation is correct, the script informs them of this.
- If the pronunciation is incorrect, the script suggests a corrected pronunciation using text-to-speech feedback.

4.2 Input

Data input for pronunciation error detection involves audio recordings of varied proficiency levels annotated for specific errors. Comparison with pronunciation dictionaries aids in error identification. User feedback, performance metrics, and usability studies supplement data for technology evaluation, enhancing correction tools.

Audio samples, annotated for errors, are compared against pronunciation dictionaries to identify mistakes. User feedback and performance metrics bolster evaluation, gauging effectiveness and accuracy. Additionally, user studies on usability aid in refining the technology, ensuring enhanced pronunciation correction tools. Aggregated data from diverse sources facilitates comprehensive analysis, refining the accuracy and efficacy of pronunciation error detection and correction technology.

Fig. 1 Shows about taking input from the user through a microphone

4.2. Ask for Speech

System prompts please speak a word for taking input from the user through voice. Meanwhile, the microphone will be activated to capture input from the user. Microphone will be activated, and the user will provide the input.

4.3. Detecting Error

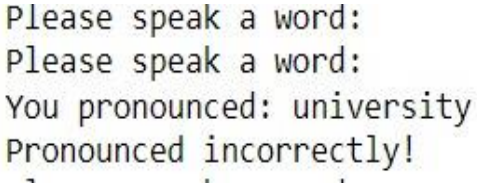

Fig. 2 Shows about the output when the user pronounces the wrong pronunciation

After capturing the input, it processes the input and detects if there is any error in the pronunciation.

If there is any error, it displays pronounced incorrectly and provides the correct pronunciation in the form of voice through the speaker.

```
Please speak a word:
You pronounced: hello
Correct pronunciation: hello
Pronounced correctly!
```

Fig. 3 Shows about the output when the user pronounces the correct pronunciation

After capturing the input, it processes the input and detects if there is any error in the pronunciation.

If there is no error it displays pronounced correctly.

4.4. Discussion

- It provides a basic yet functional tool for real-time pronunciation error detection and correction.
- It offers practical utility for language learners and individuals seeking to improve their pronunciation.
- The use of the CMU Pronouncing Dictionary is valuable for correcting common words, but it may not cover all words in the English language.
- The incorporation of text-to-speech feedback is a helpful feature, especially for language learners who want to hear the correct pronunciation.
- However, it is important to note that the script primarily focuses on individual word pronunciation and may not handle complex sentences or phrases as effectively.
- Future improvements could include expanding the dictionary or incorporating other pronunciation resources to enhance accuracy.
- Additionally, integrating machine learning techniques could provide more advanced error detection and correction capabilities, making the tool more robust and adaptable to various accents and dialects.

5. Conclusion

Pronunciation Error Detection and Correction demonstrates a practical approach to real-time pronunciation error detection and correction. Combining speech recognition, the CMU Pronouncing Dictionary, and text-to-speech capabilities, the script offers a valuable tool for language learners and individuals aiming to enhance their spoken communication skills. While the focus of this implementation centers on individual word pronunciation, it sets the foundation for more advanced and comprehensive error detection and correction systems. The importance of accurate pronunciation in language learning and effective communication cannot be overstated. This script addresses a critical aspect of language proficiency by providing immediate feedback and suggestions for improving pronunciation. Future work in this field could include expanding the dictionary or integrating machine learning techniques to handle a broader range of words, phrases, and accents. The tool's relevance extends to language education, speech therapy, and various language assessment applications.

This technology operates by converting spoken language into text, leveraging speech recognition algorithms to transcribe audio input into words or phrases. It then utilizes the CMU Pronouncing Dictionary or similar linguistic resources to compare the pronunciation of these transcribed words with their expected phonetic representations. Additionally, text-to-speech capabilities are employed to provide immediate feedback to users. When a mispronunciation is detected, the system can suggest the correct pronunciation either through auditory feedback or by displaying the correct phonetic transcription.

Moreover, the tool's functionality could extend beyond single-word corrections to handling contextual errors within sentences or phrases, enhancing its utility in real-world language communication scenarios. To improve accuracy, future iterations might explore incorporating machine learning techniques. These could enable the system to adapt and learn from a broader dataset of pronunciation variations, considering diverse accents, regional dialects, and nuances in speech patterns. Furthermore, integrating this technology into mobile applications or online platforms could increase accessibility, allowing users to practice pronunciation and receive immediate guidance anywhere and anytime.

By addressing not only language learners but also individuals seeking to refine their speaking abilities, this technology stands as a valuable asset in fostering clear and effective communication across linguistic barriers.

Acknowledgement

The authors would like to express their gratitude to the developers and contributors of the SpeechRecognition library, the NLTK library, and the CMU Pronouncing Dictionary. These open-source tools and resources have played a vital role in successfully executing the Pronunciation Error Detection and Correction project. Additionally, the authors appreciate the support and guidance received from academic institutions and communities that foster an environment.

References

[1] H. Lu, and L. K. Hansen, "Phoneme Error Detection and Discrimination Using Hidden Markov Models," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 4, pp.377-388, 2003.

[2] S. Zhang, and R. Artstein, "Automatic correction of pronunciation errors: How and when?," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4656-4659, 2011.

[3] L. Wang, and A. Narayanan, "Phonetic-based Detection of Pronunciation Errors in Non-native English Speech," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 7734-7738, 2013.

[4] C.Li, and Y, Zhao, "Phoneme Error Detsection with Bidirectional Long Short-Term Memory," *IEEE Spoken Language Technology Workshop (SLT)*, pp. 266-271, 2017.

[5] Helmer Strik et al., "Comparing Different Approaches for Automatic Pronunciation Error Detection," *Speech communication*, vol. 51, no. 10, pp. 845-852, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[6] Khiet Truong et al., "Automatic Pronunciation Error Detection: An Acoustic Phonetic Approach," *InSTIL/ICALL 2004 Symposium on Computer Assisted Learning*, pp. 1-4, 2014. [Google Scholar] [Publisher Link]

[7] Renlong Ai, "Automatic Pronunciation Error Detection and Feedback Generation for CALL Applications," *Learning and Collaboration Technologies*, vol. 9192, pp. 175-186, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[8] Yuhua Dai, "An Automatic Pronunciation Error Detection and Correction Mechanism in English Teaching Based on an Improved Random Forest Model," *Journal of Electrical and Computer Engineering*, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]