



Automatic Pronunciation Mistake Detector

¹Vinita Vikram Patil, ²Pratiksha Mohan Wadkar, ³Aishwarya Bhagwan Mohite, ⁴Ms. Priyanka Rajendra Jadhav

^{1,2,3}Students, ⁴Assistant Professor,
^{1,2,3,4} Computer Science Department,

Abstract: Machine learning is an artificial intelligence technology that educates systems to identify patterns from data and generate inferences or decisions without explicit programming. This approach is similar to teaching machines how to learn and think autonomously. Traditional English education emphasizes on writing and grammar instead of vocal language learning. As a result, while many Chinese students perform well on written exams, few can speak English successfully in everyday situations. With the increase in around the world relations, the importance placed on language as an instrument of communication has underlined the value of spoken language training. To help students practice their pronunciation outside of school, an English pronunciation detection tool is presented. Only proper pronunciation can help users convey their ideas clearly. As a result, the primary outcome of oral language learning is the development of pronunciation.

Keywords - Pronunciation Detection, Speech Recognition, Language Instruction, Communication Skills, Artificial Intelligence.

I. Introduction

Machine learning is akin to educating machines to act intelligently. By analyzing a vast array of examples, these machines can discern patterns and subsequently make informed decisions about new data based on their acquired knowledge. The development of machine learning algorithms often utilizes frameworks like TensorFlow and PyTorch, which streamline the process of building solutions.

As globalization progresses, English has gained a lot of attention as the most widely spoken language. The development of oral ability necessitates extensive oral practice, which is expected of all students studying English. Furthermore, throughout the practice phase, wrong feedback must be delivered in a timely and appropriate manner. Chinese students generally practice their pronunciation by listening to a tape, reading it out, and acting it out, similar to how they utilize a common language repeater. It was difficult to determine the relationship between the machine's speech and the children's reading because they received no feedback from the system during practice.

As computer technology has advanced and become more widely accepted, computer-aided teaching has become an important component of the application of current technology for learning in every aspect of education. Currently, much computerized assistance language learning software focuses on improving understanding of speech and language application abilities. In general, there has been very little focus on improving language verbal skills. Grammar, structure, idioms, and other elements of oral communication are usually taught independently from pronunciation in spoken language lessons. Our approach focuses on extracting key information and generating natural language descriptions that are both accurate and coherent. Automatic pronunciation of mistakes with machine learning is a based on technology approach for improving oral language fluency. This project uses machine learning algorithms to determine and provide corrections for pronunciation, supporting students in improving their ability to speak English.

II. OBJECTIVE

- To recognize and categorize pronunciation faults in spoken language.
- To provide real-time feedback from the system, allowing users to adjust their pronunciation immediately.
- To design a user-friendly interface for language learners to analyze spoken pronunciation and receive responses in a clear and engaging manner.
- To develop an automatic pronunciation detector system to provide accurate feedback on people's pronunciation in a target language.
- To use a machine to quickly understand pronunciation.
- To improve time management and productivity.

III. SYSTEM ARCHITECTURE

Following the detection of pronunciation mistakes, the text is passed through a series of corrective modules. The first of these is a grammar correction module that ensures the syntactic accuracy of the text. Next, the text undergoes a spelling mistake checker to rectify any misspelled words. Finally, a sentence correction module refines the text to enhance its overall readability and coherence. This sequential processing ensures that the final output is a well-formed, grammatically correct, and accurately transcribed version of the original spoken input.

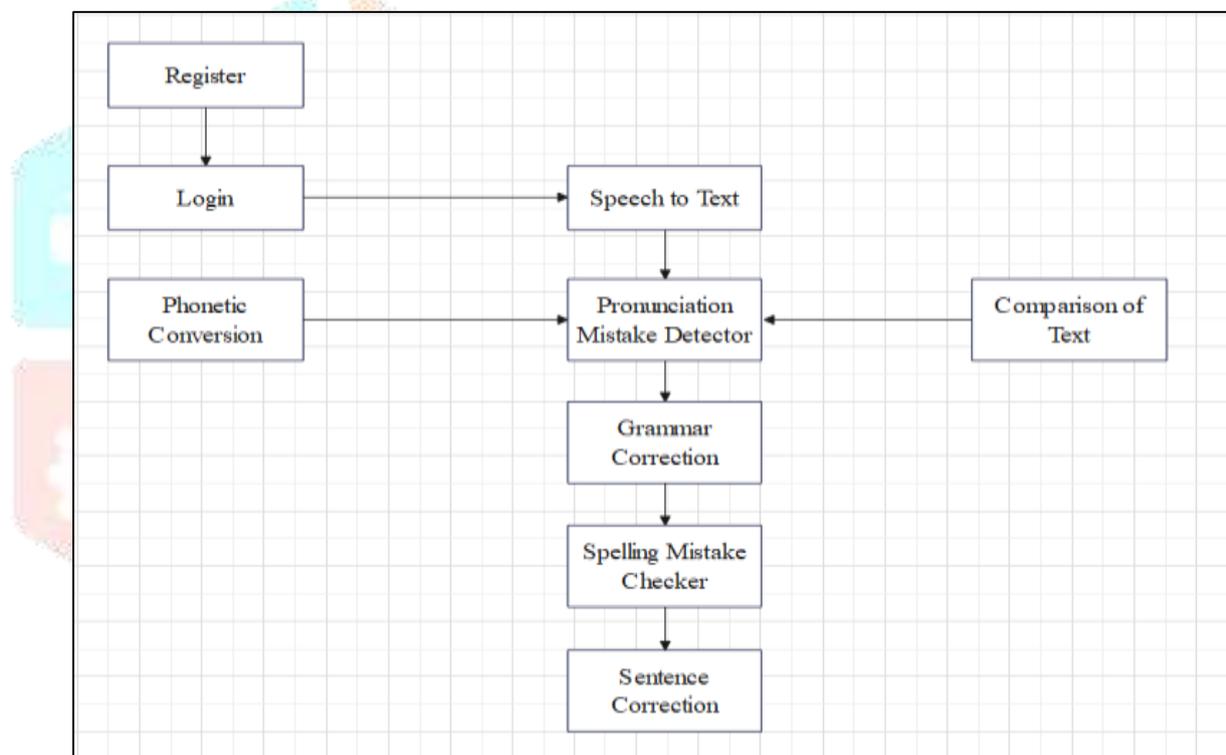


Fig 1: System Architecture

The flowchart depicts a comprehensive system designed to process and correct spoken language inputs. The process begins with the user registration, followed by logging in to the system. Once logged in, the user's speech is converted to text through a speech-to-text module. Concurrently, a phonetic conversion is carried out to aid in identifying pronunciation errors. The resulting text is then fed into a pronunciation mistake detector, which also incorporates a comparison of the converted text to a reference text to identify deviations.

I. RESEARCH METHODOLOGY

3.1 Data Flow Diagram

Data flow diagrams (DFDs) are popular because they visually show how data moves through a system. They use symbols to represent things like data flows, processes, and data stores. This makes them easy to understand for both technical and non-technical people. DFDs help you see the big picture of how a system works, from start to finish. This can be helpful for understanding existing systems or designing new ones.

The diagram represents a data flow diagram for an "Automatic Pronunciation Mistake Detector." It includes a central module that interacts with several components: "Login" (allowing users to log in with credentials), "Register" (allowing users to register with details), "Pronunciation Mistake Detector" (using text-to-speech and Langchain for detecting pronunciation errors), "Grammar Correction" (presumably for correcting grammatical errors), and "Sentence Correction" (likely for correcting sentence structure and other linguistic issues). The system integrates tools like Langchain for both grammar and sentence corrections, providing a comprehensive platform for improving language usage and pronunciation.

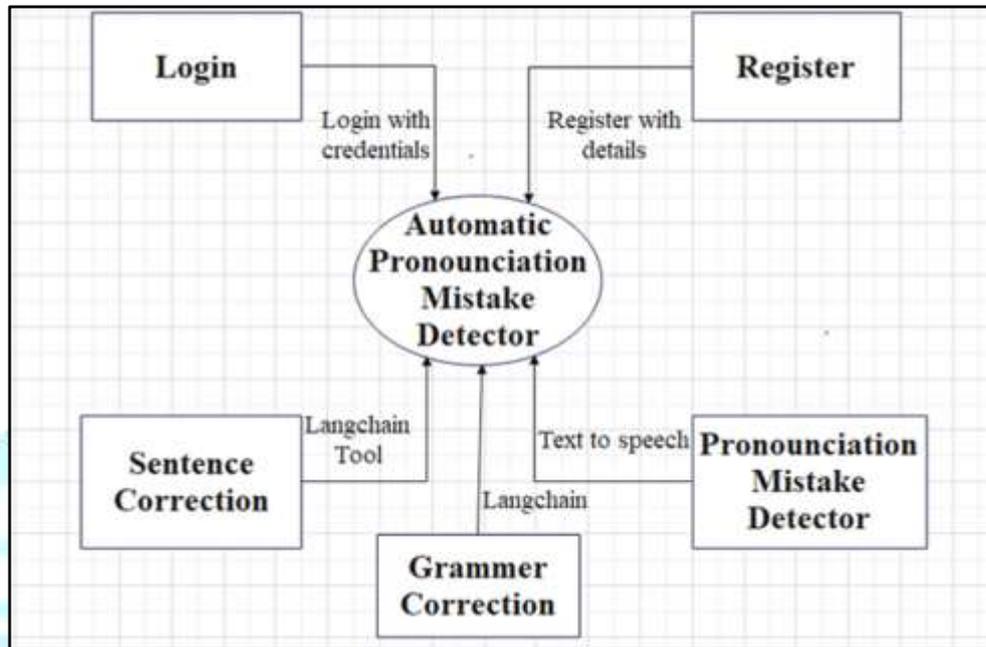


Fig 2: Data Flow Diagram

3.2 Use Case Diagram

The diagram illustrates the interaction of a user with different components of a language tool system. The user can access five main functionalities: "Register" to set up a new account, "Login" to access the system, "Sentence Correction" for refining sentence construction, "Grammar Correction" to address grammatical mistakes, and "Pronunciation Mistake Detection" for identifying and rectifying pronunciation errors. Each of these features is readily available to the user, highlighting a user-friendly design focused on improving language proficiency.

The proposed system is designed to improve users' communication skills by offering several essential features. These features include user registration and login, sentence structure refinement, grammar improvement, speech-to-text conversion, identification of pronunciation errors, and pronunciation comparison. This tool is particularly useful for those who wish to enhance the clarity, grammatical correctness, and pronunciation of their written and spoken language.

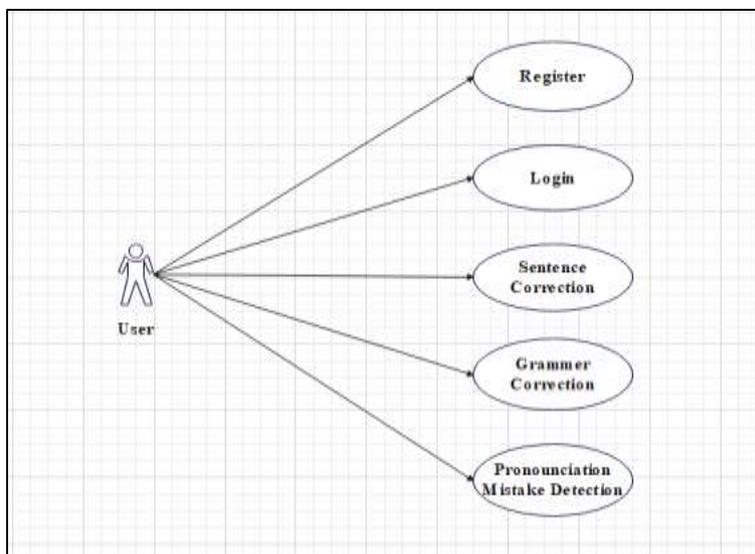


Fig 3: Use Case Diagram

IV. RESULTS AND DISCUSSION

4.1 Home Page

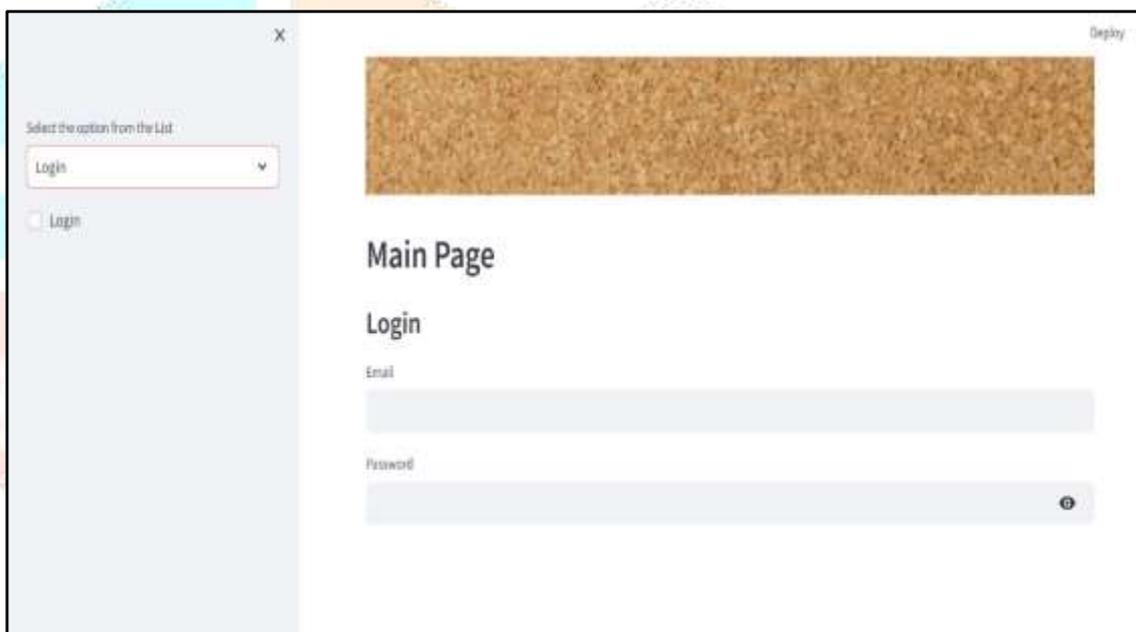


Fig 4: Home page

The figure illustrates a login page featuring a minimalistic and clean user interface. A sidebar on the left includes a dropdown menu labeled "Select the option from the List," currently displaying "Login," and a checkbox labeled "Login" underneath. The main section is titled "Main Page" with a "Login" subtitle. The login form comprises two fields for "Email" and "Password," with an option to toggle the visibility of the password. The background has a corkboard texture at the top, and a "Deploy" button is positioned at the top right corner.

4.2 Register Page

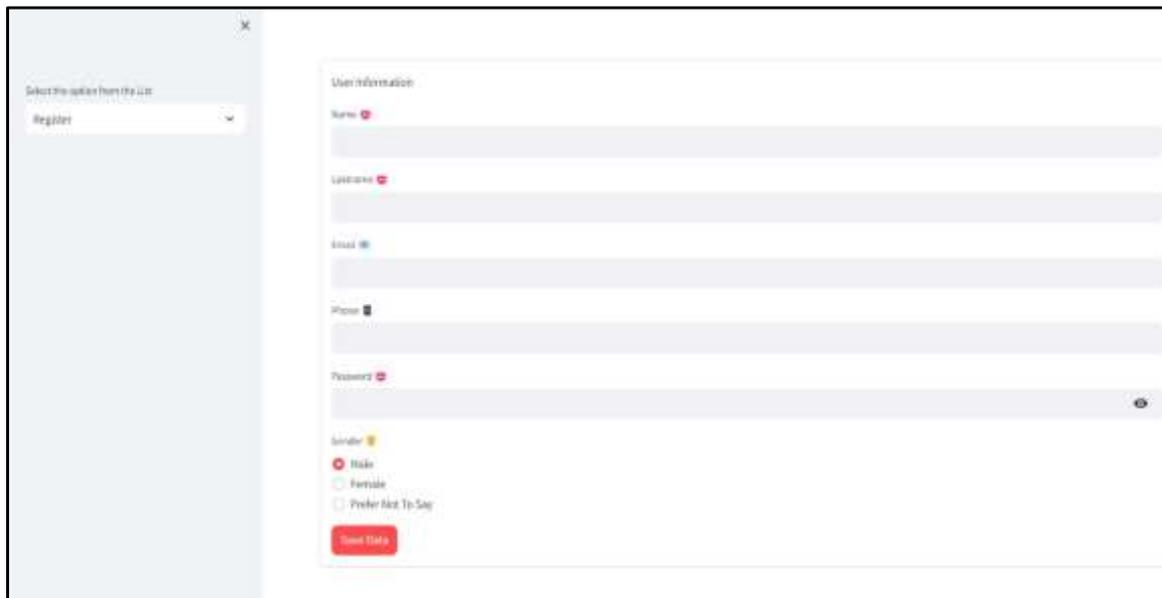
The screenshot shows a user registration form. On the left is a navigation pane with a dropdown menu labeled "Select the option from the List" and "Register" selected. The main form area is titled "User Information" and contains several input fields: "Name", "Lastname", "Email", "Phone", and "Password". Each field has a small red icon to its left. Below the "Password" field is a "Gender" section with three radio button options: "Male" (selected), "Female", and "Prefer Not To Say". A red "Save Data" button is located at the bottom of the form.

Fig 5: Register Page

The figure shows a user registration form with fields for entering personal information. The form includes input fields for name, last name, email, phone number, and password, along with an option to select the user's gender from the choices: Male, Female, or Prefer Not to Say. The form also has a "Save Data" button to submit the entered information. The interface has a clean design with icons next to each field and a navigation pane on the left side, allowing users to select options from a list.

4.3 Login Page

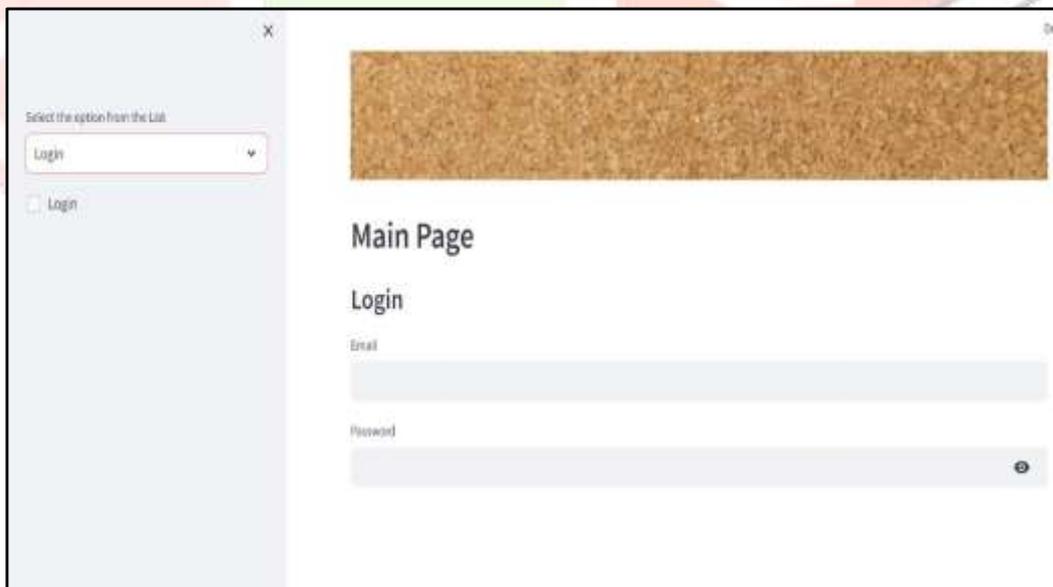
The screenshot shows a login interface. On the left is a navigation pane with a dropdown menu labeled "Select the option from the List" and "Login" selected. Below the dropdown is a checkbox labeled "Login". The main content area has a header "Main Page" and a section titled "Login". It contains two input fields: "Email" and "Password". The "Password" field has a small eye icon to its right. The background of the main content area is a textured brown color.

Fig 6: Login Page

The figure illustrates a login interface for a website or application. At the top, a header reads "Main Page," followed by a section titled "Login." Users are asked to input their email and password into two designated fields. An eye icon is situated next to the password field, presumably to toggle the visibility of the password. On the left, there is a navigation pane with a dropdown menu labeled "Select the option from the List," with "Login" as the current selection. Below the dropdown, a checkbox labeled "Login" is displayed. The layout is simple and easy to navigate.

4.4 Detect Pronunciation Error

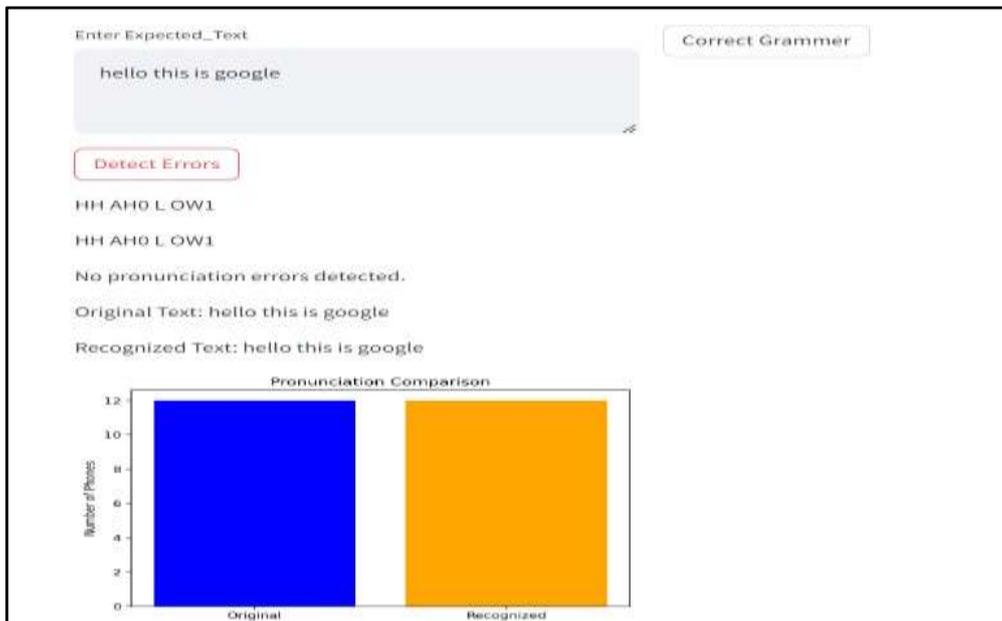


Fig 7: Generate Summary Page

The figure shows an interface for checking pronunciation and grammar. It features a text box labeled "Enter Expected_Text" where the phrase "hello this is google" is entered. There is a "Correct Grammer" button at the top right, which appears to be a typo for "Correct Grammar." A "Detect Errors" button is provided below the text box. Upon clicking this button, the phonetic transcription for the entered text ("HH AH0 L OW1") is displayed twice, indicating that no pronunciation errors were detected. The original text and recognized text are both displayed as "hello this is google." Below these results, a bar chart titled "Pronunciation Comparison" compares the number of phonemes between the original and recognized text, showing equal counts. The interface aims to help users verify and compare pronunciation accuracy.

4.5 Grammar Correction



Fig 8: Grammar Correction

The image depicts a web interface for a tool designed to detect pronunciation mistakes and correct grammar and spelling. The user, logged in as vini@gmail.com, enters the text "hallo this is gogle" into a text box labeled "Enter Expected_Text." After clicking on the "Correct Grammer" button, the tool corrects the input to "Hello this is Google." There is also a "Detect Errors" button available for use. The interface highlights misspelled words with a red underline and corrects both spelling and grammar mistakes.

4.6 Sentence Correction

Sentence Correction

Enter a sentence to correct errors:

Input Text

i is a teacher

Correct Sentence

Corrected Sentence:

I am a teacher

Fig 9: Sentence Correction

The figure shows a web interface for a "Sentence Correction" tool. The user is prompted to enter a sentence to correct errors in the "Input Text" field. The example input provided is "i is a teacher." After clicking the "Correct Sentence" button, the tool outputs the corrected sentence, "I am a teacher." This tool effectively identifies and corrects grammatical mistakes in user-provided text.

IV. CONCLUSION

The system implements the concept of an Automatic Pronunciation Mistake Detector. It takes voice data from the user, checks those words against a dictionary, and determines if the words are pronounced correctly. This system has partially completed the recognition of words and can declare whether the pronunciation is correct or if a mistake is found. Additionally, it offers feedback on common pronunciation errors and suggests corrections to help users improve their spoken language. Future enhancements may include support for multiple languages and dialects to broaden its usability and effectiveness.

REFERENCES

- [1] Bandar Ali Al-Rami and Yousef Houssni Zrekat, "A framework for pronunciation error detection and correction for non-native Arab speakers of English language", ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print) © 2023 by the authors; licensee Growing Science, Canada. doi: 10.5267/j.ijdns.2023.5.004
- [2] 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 Volume 2022, Article ID 6011993, 2022.
- [3] Lei Chen, Chenglin Jiang, Yiwei Gu, Yang Liu, and Jiahong Yuan, "Automatically Detecting Reduced-formed English Pronunciations by Using Deep Learning", Proceedings of the 17th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2022), pages 22 – 26 July 15, 2022 © 2022 Association for Computational Linguistics.
- [4] Liang Meng, Jinal Upadhyay, Sumit Kumar et al., "Nonlinear Network Speech Recognition Structure in a Deep Learning Algorithm", Computational Intelligence and Neuroscience Volume 2022, Article ID 6785642.
- [5] Zhan Zhang, Yuehai Wang, Jianyi Yang, "Text-conditioned Transformer for automatic pronunciation error detection", <https://doi.org/10.1016/j.specom.2021.04.004> Received 26 August 2020; Received in revised form 11 February 2021; Accepted 18 April 2021 Available online 24 April 2021 0167-6393/© 2021 Elsevier B.V. All rights reserved.
- [6] Fengming Jiao, Xin Zhao, Jiao Song, "A Spoken English Teaching System Based on Speech Recognition and Machine Learning", iJET – Vol. 16, No. 14, 2021 Jitang College of North China University of Science and Technology, Tangshan, China, 2021.

- [7]Zhang Gang, “Quality evaluation of English pronunciation based on artificial emotion recognition and gaussian mixture model”, Journal of Intelligent & Fuzzy Systems 40 (2021) 7085– 7095 DOI: 10.3233/JIFS-189538 IOS Press, Henan Medical College, Henan Zhengzhou, China 2021.
- [8]Zhang Shufang, “Design of an Automatic English Pronunciation Error Correction System Based on Radio Magnetic Pronunciation Recording Devices”, Journal of Sensors Volume 2021, Article ID 5946228, 12 pages <https://doi.org/10.1155/2021/5946228>
- [9] Cai Xiao-ming, “Automatic Detection of English Pronunciation Errors based on Statistical Pattern Recognition”, December 25-27, 2020 Harbin, China IEEE International Conference on Industrial Application of Artificial Intelligence (IAAI), 2020.
- [10] Rahib Abiyev, John Bush Idoko, Murat Arslan, “Reconstruction of Convolutional Neural Network for Sign Language Recognition”, Proc. of the 2nd International Conference on Electrical, Communication and Computer Engineering (ICECCE) 12-13 June 2020, Istanbul, Turkey, 2020.

