© 2023 IJNRD | Volume 8, issue 11 November 2023 | ISSN: 2456-4184 | IJNRD.ORG

SpeakEZ GPT: Your Talkative AI Pal

1st Prashant Yadav *Dept of CSE Chandigarh University*Punjab, India
prashantakayadav@gmail.com

2nd Khushbu Yadav *Dept of CSE Chandigarh University*
Punjab, India
khushbuy499@gmail.com

3rd Gaurav Mehta
Dept of Electronics and Comm. EngineeringChandigarh University
Punjab, India gaurav.e14354@cumail.in

***Abstract*—SpeakEZ GPT, introduced in this paper, is a cutting-edge conversational AI system designed to enhance user interactions with the GPT-3.5 language model. This system allows users to engage in natural, voice-based conversations with the AI, making interactions more intuitive and user-friendly. By leveraging advanced technologies like speech recognition and text-to-speech synthesis, SpeakEZ GPT offers a wide range of applications in virtual assistants, customer support, education, and accessibility. This abstract provides an overview of the system's capabilities and potential, emphasizing its role as "Your Talkative AI Pal."**

***Index Terms*—Conversational AI, Voice Commands, GPT- 3.5, Speech Recognition, Natural Language Processing, Text-to-Speech, User Experience, Accessibility.**

I. Introduction

In todays world, where technology's becoming increasingly integrated into our lives conversational artificial intelligence (AI) systems have become incredibly useful, for communication and finding information. While we are now accustomed, to interacting with AI through text there is a growing desire to make these interactions feel more human like and accessible[1]. That's where SpeakEZ GPT comes in! It's a AI system that aims to be your chatty AI companion. This paper introduces SpeakEZ GPT, a project that enables voice based conversations using the renowned GPT 3.5 language model developed by OpenAI.As technology progresses it's crucial, for AI systems to keep up with the communication preferences that people have. While text based AI interfaces serve their purpose they can sometimes feel limiting for those who prefer a natural and efficient way of interacting[2]. SpeakEZ GPT aims to overcome this limitation by utilizing speech recognition technology enabling users to communicate with the AI model using voice commands. In return the AI responds through spoken words enhancing the experience and making it more engaging and user friendly. This paper provides a comprehensive exploration of the SpeakEZ GPT project, unveiling the intricacies of its architecture, the underlying technologies, and the myriad possibilities it unlocks in various domains. By enabling AI to respond to voice commands and fostering two-way natural conversations, SpeakEZ GPT has the ability to completely transform how we use technology, becoming a true companion in our digital journeys[3].

The area of computer science is concerned with creating computer programs that can understand spoken language. Note that voice recognition does not imply that the computer understands what is being said; it just suggests that it can accept dictation. Natural language processing is a subfield of computer science that deals with understanding human languages. There are several speech recognition systems on the market. Thousands of words can be recognized by the most powerful[4]. However, they often need a lengthy training session in order for the computer system to get used to a certain voice and accent. They are referred to as speaker- dependent systems. Many systems also call for the speaker to speak clearly, and deliberately, and to briefly stop between each syllable.

Discrete speech systems are what these systems are known as. Continuous speech systems, or voice recognition systems that let you speak naturally, have advanced significantly in recent years. For personal computers, a number of continuous-speech solutions are now available.Voice recognition systems have historically been employed primarily in a select few specialized scenarios due to their flaws and expensive cost. For instance, these devices are helpful when a user's hands are busy or impaired and they are unable to enter data using a keyboard.The user may just voice orders into a headset rather than typing them. However, as the price drops and the performance increases, voice recognition systems are becoming more widely utilized as a keyboard substitute[5].

The technique by which a computer converts an acoustic voice signal to text is known as automatic speech recognition. The technique of mapping an auditory speech signal to some kind of abstract meaning of the speech is known as automatic speech comprehension[6]. For the purpose of supporting a single speaker, a speaker-dependent system is created. These systems are typically simpler to create, less expensive to purchase, and more precise, but they are less adaptable than speaker-independent or adaptive systems. For every speaker of a specific kind (such as American English), a speaker-independent system is created. The precision of these systems is lower than speaker-dependent systems, and they are the most expensive and complex to construct. They are more adaptable,

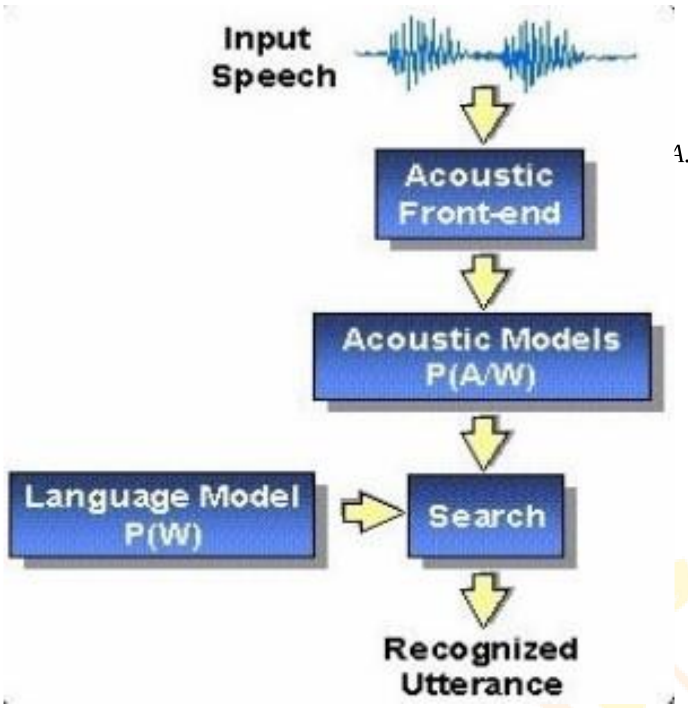

Fig. 1. Simple equations are used in Fig. 1 to depict a voice recognition system's front-end unit, model unit, language model unit, and search unit mathematically. The identification procedure is illustrated below (Fig. 1).

though.To adjust to the features of new speakers, a speaker adaptive system is designed. Among chatterer-sovereign besides chatterer-at the mercy of schemes, its complexity is in between.

II. Literature Survey of Speech Recognition

Machine recognition first appeared in the early 1920s. The Radio Rex (toy) was created in 1920[165] and was the first commercially successful voice recognition device[7].Bell Labs started exploring the ideas of voice technology as early as 1936. At the New York World's Fair in 1939, Bell Labs displayed a voice synthesis device that mimics conversing. Bell Labs eventually gave up on efforts to create speech- simulated listening and recognition, basing their decision on the false assumption that success would ultimately need arti- ficial intelligence.The 1950s saw the first attempts to develop systems for automated speech recognition by machines as a number of researchers attempted to utilize the basic principles of acoustic phonetics. In the 1950s, the majority of speech recognition systems looked at spectral resonances that were taken from output signals of an analog filter bank and logic circuits during the vowel area of each phrase. Using a single speaker, Davis, Biddulph, and Balashek created a method for isolated digit recognition. in 1952 at Bell Laboratories. The technique largely depended on spectral resonance measurements in each digit's vowel area[8]. Olson and Belar attempted to identify 10 unique syllables of a single talker as represented by 10 monosyllabic phrases in a separate experiment at RCA Laboratories in 1956. Advances in voice recognition for tiny implementations, speech recognition for continuous speech, and speech recognition for noisy situations have been documented.

A. *Robust speech recognition*

Using a variety of approaches, it was investigated how resilient speech recognition systems are to the mismatch between training and testing environments caused by background noise, voice distinctiveness, microphones, transmission channels, room reverberation, etc. For accurate speech recognition, it's vital to use techniques like the structural maximum a posteriori (SMAP) method, model decomposition, parallel model composition (PMC), and maximum likelihood linear regression (MLLR). A signal bias reduction (SBR) method based on maximum likelihood estimation is described in the work by Mazin G. Rahim for the reduction of adverse effects that occur in telephone speech recognition systems, such as background noise, channel distortions, etc. Utilize maximum likelihood (ML) stochastic matching to reduce the acoustic mismatch[9].

B. *Modeling Techniques*

Eduardo proposed a collection of acoustic modeling and decoding techniques for utterance verification (UV) in HMM-based continuous speech recognition.Regarding automated voice recognition using high dimension feature vectors to condense short-term speech features, Lawerence K. discusses this topic. These were accomplished by selecting certain parameters in one of two ways: either to increase the chance of detected speech signals or, alternatively, to reduce the frequency of classification mistakes[10]. That Tat he has suggested multiple models, which are i) the FE-HMM,NC-FE-HMM, FE-GMM,NC-FE-GMM, FE-VQ and NC-FE-VQ in the FE approach, ii) the FCM-HMM, NC-FCM-HMM, FCM-GMM and NC-FCM-GMM in the FCM tackle and iii) the complex the HMM and the Gaussian mixture to serve as particular prototypes that combine the the FE and the FCM method methods for language recognition[11].

C. *Noisy speech recognition*

In the past ten years, not much research has been done on loud speech detection. Adoram Erell devised one of the crucial techniques known as minimal mean square error (MMSE) estimation of the filter log energies, which markedly improved upon prior algorithms.The resilience of the SR system to additive noise is increased by a model-based spectral estimation approach that has been developed[12]. The approach has been specifically designed for filter bank-based systems, where the estimation should aim to reduce distortion as determined by the recognizers distance. In the field of noisy robust voice recognition, nothing has been done. The resilience of the voice recognition system to additive noise has been improved via a model-based spectral estimation technique.

D. *Multimodal speech recognition*

Humans communicate with one another through a variety of media. Studies on language lucidity take demonstrated that taking equally pictorial and aural info boosts the amount of effective data transmission, particularly when the communication is compound or after message occurs in a loud setting. The custom of pictorial expression data, especially rim info, in language gratitude takes stood studied. The findings indicate that utilizing both forms of information improves recognition performance over using just audio or just visual information, especially in noisy environments[13]. The exertion completed in this study defines the behaviour of such multi-distance showing in real recognition. Jerome R. has created Large Language Talking Gratitude using Multi-span Numerical Philological Mockups.

III. Methodology

ASR, or automatic speech recognition, is a revolutionary technique that enables computers and other devices to translate spoken language into written text. This technology is a vital part of the wider subject of natural language processing(NLP) and is extremely important in a variety of applications,from voice commands and virtual assistants to accessibilitytools and transcription services[14]. "SpeakEZ GPT: YourTalkative AI Pal" creation and assessment entail a systematicresearch approach that covers a number of phases, from designthrough implementation and testing. This technique aims toguarantee the conversational AI system's reliability, potency,and usability. An overview of the project's research techniqueis provided below- Project Inception and Problem Definition: Clearly define the objectives of the project, including the goal of creating a voice-enabled interface for GPT-3.5. Identifythe key challenges and limitations of current conversational AI systems. Formulate research questions and hypotheses to guide the project. Literature Review:

Conduct a comprehensive review of existing research on conversational AI, speech recognition, natural language processing, and text-to-speech synthesis. Analyze relevant papers, articles, and case studies to gather insights and best practices. Identify gaps in the literature that the project aims to address. Data Collection and Preparation:

Gather a diverse dataset of voice recordings and text transcriptions to train and test the speech recognition component. Acquire a dataset of spoken interactions for fine-tuning the natural language processing model. Ensure data privacy and ethical considerations are addressed. System Design and Architecture:

Develop a detailed system architecture, outlining the components, their interactions, and data flows. Select appropriate technologies and frameworks for speech recognition, natural language processing, and text-to-speech synthesis. Design a user-friendly interface for voice-based interactions. Development and Implementation:

Develop and code the various components of SpeakEZ GPT, including speech recognition, NLP, and TTS. Integrate the GPT-3.5 model and fine-tune it for voice-command-based interactions. Implement a robust and scalable back-end infrastructure to support the system. Testing and Evaluation:

Conduct extensive testing to evaluate the accuracy and performance of the speech recognition component, ensuring it can handle various accents and languages. Evaluate the natural language processing capabilities by assessing the system's understanding of user intent and context. Assess the quality and naturalness of the text-to-speech synthesis. Perform user testing to gather feedback on the user experience and iterate on design and functionality. Optimization and Refinement:

Continuously optimize the system's components based on testing results and user feedback. Implement improvements in latency, response times, and accuracy. Address privacy and security concerns, ensuring user data is protected. Documentation and Reporting:

Document the development process, including system architecture, algorithms, and technologies used. Compile research findings, test results, and user feedback into a comprehensive report. Prepare documentation for users, including user guides and FAQs. Future Research and Development:

Identify avenues for future research and improvement, such as multilingual support, emotional intelligence, and expanded applications. Explore possibilities for integrating SpeakEZ GPT into various domains, including healthcare, education, and entertainment. Conclusion and Dissemination:

Summarize the project's achievements and contributions to the field of conversational AI. Share research findings and the developed system with the AI and technology communities through conferences, publications, and open-source initiatives.

This research methodology ensures a structured and systematic approach to the development of "SpeakEZ GPT: Your Talkative AI Pal," with a focus on user experience, accuracy, and future scalability and improvements.

IV. System Architecture

A. *Speech Recognition*

The way people interact with computers and other technology has been completely transformed by speech recognition, a key element of contemporary technology. It method that lets a processer system or machine to comprehend verbal language and interpret it into manuscript or commands[15]. This game-changing technology has been widely used in a variety of fields, from smartphone virtual assistants to tran- scription services and accessibility solutions. Speak-to-device functionality has improved user comfort while also expanding the possibilities for accessibility and productivity. Since it was first developed, speech recognition technology has seen substantial development, thanks to developments in machine learning, artificial intelligence, and deep neural networks.

We will examine the development of voice recognition technology, its underlying ideas, and its various uses in the modern digital environment in this introduction. We will also examine the difficulties faced by academics and developers as they work to improve speech recognition's precision, adaptability, and seamless integration into daily life[16]. Speech recognition technology has the potential to further close the communication gap between humans and machines, ushering in a new era of interactivity and accessibility.

B. Natural Language Processing

NLP, or natural language processing, is a fascinating field that combines linguistics and technology to allow computers to understand, decipher, and generate meaningful and context-appropriate human language. It is an area of artificial intelligence (AI) that has shown rapid development and innovation in recent years. This advancement has major consequences for how humans communicate with machines, handle enormous volumes of text data, and improve our comprehension of language.The goal of NLP is to close the gap between the complexity of human language and the computing power of machines. Its uses are broad and varied, having an influence on many areas of our life[17]. The way we interact with digital technology has been completely transformed by NLP, from chatbots that have human-like conversations to sentiment analysis tools that measure public opinion, from language translation services that remove language barriers to recommendation systems that personalize content.

The fundamental ideas and difficulties of NLP, as well as the development of the discipline and the key developments that have made it a crucial component of contemporary computing, will all be covered in this introduction. We will get insights into how machines are learning to interpret and produce language as we dive deeper into the field of natural language processing[18]. This will open up a world of opportunities for communication, information retrieval, and knowledge sharing. NLP is a promising new development in artificial intelligence that has the potential to improve the usability, effectiveness, and human-likeness of our interactions with technology.

C. Text-to-Speech

Text-to-voice, often known as TTS, is a revolutionary technology that breathes life into printed text by translating it into real-sounding, human-like voice. This cutting-edge field has significant implications for accessibility, communication, and human-computer interaction as it lies at the confluence of computer science, linguistics, and artificial intelligence. With the help of developments in deep learning and neural networks, TTS technology has advanced significantly in recent years. Speech synthesis technology has advanced from robotic, monotonous voices to extremely lifelike, expressive systems that can transmit subtlety and emotion. TTS's potential has been broadened by this development, making it a crucial component in a variety of applications[19].

In this introduction, we'll look at the foundational ideas of TTS as well as its progress throughout time and current uses. TTS is transforming how we interact with information and technology in a variety of ways, from assisting people with visual impairments to access written content to improving the usability and enjoyment of navigation systems, enhancing the capabilities of virtual assistants, and providing narration for multimedia content.

V. AUTOMATIC SPEECH RECOGNITION

Reflex talking gratitude just earnings that the machine comprehends hominoid dialog and understands that the hominid voice issues guidelines to the mainframe directly. The computer then carries out the orders in accordance with the known and handled speech, understanding the bright interface amongst the hominid and the processor. The majority of conventional automated speech recognition models employ the probability and likelihood-based Hidden Markov Gaussian Mixture Model (HMM-GMM)[20]. The more intricate and multifaceted structure of the conventional speech recognition model is seen in Figure 2. The front-end speech preprocessing step is when the standard automated dialog gratitude model needs to make use of the relationship between the speech signal and the digital model. However, it is challenging to adapt the preprocessing models in conventional speech recognition to these many settings because of the complication of talking material abstraction caused by the diverse pronunciations of persons with varied idioms, genders, and ages[21].

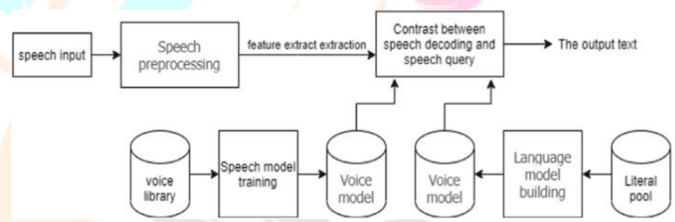

Fig. 2. Traditional voice recognition software

ASR, or automatic speech recognition, is a revolutionary technique that enables computers and other devices to translate spoken language into written text. This technology is a vital part of the wider subject of natural language processing (NLP) and is extremely important in a variety of applications, from voice commands and virtual assistants to accessibility tools and transcription services. ASR systems analyze and decode the acoustic patterns of human speech using cutting-edge algorithms, machine learning strategies, and neural networks. These systems recognize phonemes, words, and sentences to deconstruct spoken words and phrases into linguistic elements. The objective is to correctly convert spoken language into text so that computers can comprehend and interpret spoken input.ASR has several uses and is constantly being developed[22]. It is essential to voice-activated technology since it enables users to communicate verbally with gadgets and programs like Siri and Google Assistant. ASR is especially essential in contact centers where it enables analytics and voice-based customer support, as well as in the healthcare sector where it helps with medical transcription.

In this overview of automatic speech recognition, we'll look at the fundamental ideas, problems, and developments. We'll explore the subtleties of how ASR systems operate, the challenges of dealing with different languages and accents, and the continuous initiatives to make speech recognition more precise and adaptable. ASR is a vibrant[23], developing discipline that has the potential to revolutionize how we use technology and obtain information through voice-activated interfaces.

VI. Future Scope

The development of SpeakEZ GPT represents a significant step forward in the realm of conversational AI, opening up a world of possibilities for the future. As technology continues to advance and user expectations evolve, the future scope of SpeakEZ GPT holds immense promise. Here are several areas of future development and application for this talkative AI companion- Multilingual Support: Expanding SpeakEZ GPT's language capabilities to encompass a broader range of languages will enhance its accessibility and usefulness on a global scale[24]. This will involve training the system on diverse linguistic nuances and accents.

Enhanced Emotional Intelligence: Future iterations of SpeakEZ GPT could be designed to recognize and respond to user emotions, making interactions more empathetic and responsive. This could be especially valuable in applications like mental health support or companionship.

Voice Customization: Allowing users to customize the voice of SpeakEZ GPT could add a personal touch to interactions. Users could choose from a variety of voices or even create their own voice profiles.

Improved Context Awareness: Developing the ability for SpeakEZ GPT to maintain context over extended conversations and reference past interactions will result in more natural and coherent dialogues.

Real-time Language Translation: Expanding SpeakEZ GPT's language translation capabilities can facilitate seam- less communication between individuals who speak different languages.

Voice-Controlled Smart Homes: SpeakEZ GPT could serve as a central hub for controlling smart home devices, enabling users to manage their homes through voice commands effortlessly.

Education and Training: The system can be further harnessed to create immersive educational experiences, provid- ing personalized tutoring and adapting to individual learning styles.

Medical and Healthcare: In healthcare, SpeakEZ GPT could assist with tasks such as monitoring patient conditions, offering medical advice, or providing emotional support for patients.

Entertainment and Content Creation: SpeakEZ GPT can be employed to create audio content, generate music, or assist in scriptwriting, enhancing creativity in various entertainment industries.

Business Applications: SpeakEZ GPT can find applications in business environments for tasks such as automated call center support, voice-controlled data analysis, and virtual meetings.

Privacy and Ethical Considerations: Future developments should prioritize robust privacy measures and ethical considerations, ensuring that user data is protected and that AI interactions adhere to responsible guidelines.

User Experience Enhancements: Continual improvements in user experience, including reducing response times, improving naturalness in conversation, and refining user interfaces, will be crucial.

Cross-Platform Integration: Integrating SpeakEZ GPT into various platforms and devices, such as cars, appliances, and wearables, will extend its utility and reach.

The future of SpeakEZ GPT as a talkative AI pal is bright, with potential applications across diverse domains and the capacity to reshape the way we interact with technology. As technology and AI research progress, this conversational AI system will continue to evolve, offering innovative solutions to meet the ever-changing needs of users in an increasingly interconnected world.

VII. Challenges

The development of SpeakEZ GPT, a conversational AI system designed to facilitate voice-based interactions with the GPT-3.5 model, has been a journey marked by numerous challenges. While this innovative technology has the potential to revolutionize user experiences, several key obstacles had to be overcome during its creation[25]. In this section, we will outline some of the primary challenges encountered in the development of SpeakEZ GPT-

Speech Recognition Accuracy-One of the central challenges in developing SpeakEZ GPT was achieving high speech recognition accuracy. Understanding and transcribing voice commands accurately, especially in noisy environments or with diverse accents, required advanced speech recognition models and extensive training data.

Natural Language Processing Integration: Integrating natural language processing (NLP) capabilities seamlessly into the system was a complex task. Ensuring that the AI could understand user intent from transcribed voice input and generate contextually relevant responses posed significant challenges.

Text-to-Speech Realism: The quality of the text-to-speech (TTS) synthesis was crucial for a natural and engaging conversation. Striking the right balance between synthesizing speech that is both human-like and clear was an ongoing challenge.

Latency and Responsiveness: Users expect near-instantaneous responses in voice interactions. Minimizing latency between user input and AI response was a constant consideration during system optimization.

Privacy and Security: Ensuring the privacy and security of user data and voice recordings was paramount. Developing robust privacy protocols and encryption measures added complexity to the project.

Scalability: As the system gained popularity, scaling to accommodate a growing user base without compromising performance became a significant challenge. Cloud infrastructure and load balancing solutions were essential.

Diverse Use Cases: SpeakEZ GPT is intended for use in various domains, each with its unique requirements. Tailoring the system to excel in applications such as virtual assistants, customer support, and education posed challenges in adaptability and versatility.

Continuous Learning: AI models like GPT-3.5 benefit from continuous learning and updates. Maintaining the system's knowledge base and adapting to evolving user needs presented ongoing challenges.

User Training and Familiarization: Ensuring users could effectively and intuitively interact with the voice-based system required user training and onboarding solutions.

In overcoming these challenges, the development team behind SpeakEZ GPT has strived to create a powerful and user-friendly conversational AI system. While these challenges were formidable, they also represent opportunities for innova- tion and improvement as SpeakEZ GPT continues to evolve and provide users with a truly talkative AI companion.

VIII. Conclusion

The most common and practical form of interpersonal communication is speech.Over the earlier 50 centuries, there has been a lot of interest in research into voice and recital recognition as a first step toward genuine human-machine communication, whether it be out of scientific curiosity to generate automatons that look like individuals or a longing to mechanize labour through apparatuses. We have also run into a few real-world obstacles that prevent the general adoption of applications and services. Human individuals perform one to two orders of magnitude less inaccurately than robots in the majority of voice recognition tests. Nowadays, there is a growing interest in figuring out how to close this performance difference. There is moderately diminutive that is identified nearby in what way hominid dialog is administered.

Despite the significance of these study areas, acoustic-phonetics, psychoacoustics, speech perception, and linguistics research will result in the greatest gains. Future technologies must effectively depict, store, and retrieve the knowledge required for natural discourse. An extensive review of speech recognition research and some year-by-year advancement are both included in this publication. Despite the enormous progress made over the past 20 years, there is still more to be done. A strong speech recognition system, in our opinion, should be able to perform in a wide range of conditions, including those caused by speaker variability and environmental influences. Speech recognition is a complicated and fascinating topic in and of itself.

We take completed an sweat to contemporary a detailed indication of the expansion of singing acknowledgement knowledge done the earlier 60 centuries in this daily. Meanwhile individuals usage speech credit on a even root, it is one of the zones of reproduction acumen that mixes the greatest. Talking credit has involved genii as a noteworthy arena of training, has jammed culture scientifically, and is predicted to raise level additional in this ground of humanoid-mechanism contact. We hope that the ASR research community will read this publication and gain insight and motivation.

References

[1] Yawen, G. (2022) Research on Speech Recognition Methods in the Con- text of Deep Learningfor Artificial Intelligence. Software, 43(05):122- 124

[2] Zeru, L. (2022) Speech Recognition Method and Development Trend Based on Artificial Intelligence Deep Learning. NEW GENERATION OF INFORMATION TECHNOLOGY,5(1): 104-106

[3] Zhenghong, WW. Su, P. Kun, Z. (2022) Research on Speech Recognition Based on Big Data and Deep Learning. Software, 43(01):133-135.

[4] Rodr´ıguez, ME. Ru´ız-Mezcua, B. Garc´ıa-Crespo, A. et al. (1997) Speech/speaker recognition using a HMM/GMM hybrid model. Springer-Verlag, 1997:227-234.

[5] Jinyin, C. Linhui, Y. Haibin, Z. et al. (2020) A black-box adversarial attack method for speech recognition systems. Small Microcomputer Systems,41(5):1019-1029.

[6] Yimin, H. Huiqiong, Z. Zhengyi, W. (2017) A Review of Research Ad- vances in Deep Learningfor Speech Recognition. Application Research of Computers, 34(8): 2241-2246

[7] Jia, W. Dongmei, L. (2020) A review of deep learning applications in speech recognition Technology, 13(16): 191-197.

[8] David, E., and Selfridge, O., Eyes and ears for computers, Proc.IRE 50:1093-1101,1962.

[9] Stefan Windmann and Reinhold Haeb-Umbach, Approaches to Iterative Speech Speech, And Language Processing, Vol. 17, No. 5, July 2009.

[10] Mathias De Wachter et.al., Template-Based Continuous Speech recog- nitionIEEE Transactions On Audio, Speech Vol. 15, No. 4, May,2007 1377.

[11] Hui Jiang, Xinwei Li, and Chaojun Liu,Large Margin Hidden Markov Models for Audio, Speech, And Language Processing, Vol. 14, No. 5, September 2006.

[12] .P.Hoffbeck and D.A.Landgrebe,Covariance Matrix Estimation and Clas- sification with Limited Training Data, IEEE Trans. Pattern Analysis and Machine Intelligence, vol.18, no.7,pp.763-767, July 1996.

[13] K.L.Oehler and R.M.Gray, Combining Image compression and Classi- fication Using vector quantization, IEEE Trans. Pattern Analyssis and Machine Intelligence, Vol.17, no.5,pp.461-473,1995.

[14] B.Juang, W.Chou, and C.H.Lee., Minimum classification error rate methods for speech recognition, IEEE Trans.Speech and Audio Proc. T-SAP,5(3):257- 265, May 1997.

[15] L.R.Rabiner, A tutorial on hidden Markov models and selected applica- tions in speech recognition, Proc.IEEE 77(2):257-286.February 1989.

[16] .L.Rabiner, B.Juang, S.Levinson, and M.Sondhi, Recent developments isolated word recognition, In Proc.of IEEE Trans.ASSP, 34(1):52.

[17] .Anil K.Jain, et.al., Statistical Pattern Recognition:A Review, IEEE Transactions on Pattern Analysis,Vol.22, No.1, January 2000.

[18] .Waleed H. Abdulla,HMM-based techniques for speech segments extrac- tion, Scientific Programming 10 (2002) 221239 ,IOS Press.

[19] Anant G. Veeravalli et.al, Phoneme Recognition Using Hidden Markov Models, Abstract Submitted to Huntsville Simulation Conference 2004.

[20] Xiaodong Cui et.al., A Study of Variable-Parameter Gaussian Mixture Hidden On Audio, Speech Processing, Vol. 15, No. 4, May 2007.

[21] .Hui Jiang, et.al., Large Margin Hidden for Speech Recognition Au- dio,Speech, And Language Processing, Vol. 14, No5,September 2006.

[22] Asghar.Taheri , Mohammad Reza Tarihi et.al., Fuzzy Hidden Markov Models for Speech Recognition Computing February 2005 ,ISSN, 1305-5313.

[23] N.E.Huang et.al.,, The empricial mode decomposition and the Hilbert spectrum for nonlinear, London, UK,Vol.454,pp.903-995,1998.

[24] Albs Sloin et.al., Support Vector Machine Training for Signal Process- ing, Vol.56,NO.1,Jan.2008.

[25] .Andras Kocsor et.al.,Kernel Based Feature Extraction with a Speech Signal Processing,Vol.52, NO.8, Aug.2004.