



**Chandigarh Engineering College Jhanjeri**  
**Mohali-140307**  
**Department of Computer Science & Engineering**

# **Automated Speech Recognition**

Project-I

**Bachelor of Technology**  
Artificial Intelligence And Machine learning



**SUBMITTED BY:**

Name : Shreyansh Tiwari

Roll No: 2129990

FEB 2024

**Under the Guidance of**

Mr.Saroj Kumar Gupta

**Department of Computer Science & Engineering**  
**Chandigarh Engineering College Jhanjeri**  
**Mohali - 140307**



### **Table of Contents**

<b>SNo</b>	<b>Contents</b>	<b>Page No</b>
1.	Introduction	
2.	Brief Literature survey	
3.	Problem formulation	
4.	Objectives	
5.	Methodology/ Planning of work	
6.	Facilities required for proposed work	
7.	References	

## **Introduction**

The introduction provides an overview of the project, its significance, and the technologies used.



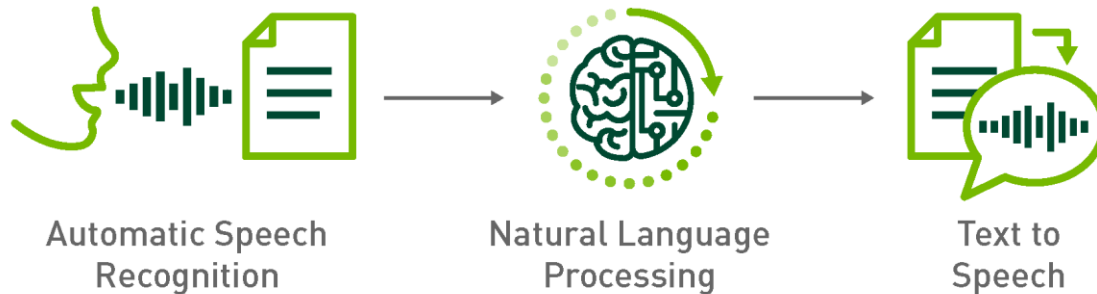
Voice-controlled assistants have gained widespread popularity in recent years, offering users a convenient and hands-free way to interact with technology. This project introduces a voice-controlled assistant implemented in Python, leveraging various libraries such as pyttsx3, speech\_recognition, and pyautogui. The assistant aims to perform a range of tasks, including opening applications, retrieving information from the internet, and executing system commands, all through voice commands. This project highlights the potential of natural language processing (NLP) and automation techniques to enhance user experience and productivity in everyday computing tasks.



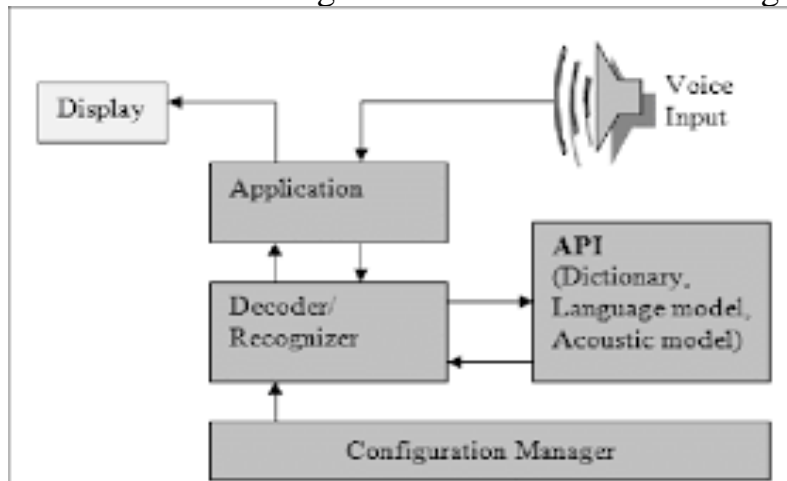
## **Brief Literature Survey**

From the year 2000 Speech recognition Technology had achieved close to the 80 percent accuracy. For almost of all the Decade There aren't a lot of Advancements till google has come with a start of google search voice. As it was an application which put speech recognition into hands of lakhs of people . This was also Significant because that the processing power would be offloaded to its data Centres. Not only for that, Google Application was collecting data from many billions of the searches which could help this to predict what a human is actually Saying. That time Google's English voice search system, included 240 billion words from user searches. 2010s In 2012 Apple Launched SIRI which was as same as the Google's VOICE SEARCH. The early part of the decade saw an explosion of the other voice Recognition Applications. And with Amazon's ALEXA, Google Home we've seen consumers Becoming More and More comfortable talking to Machines. Today, Some of the Largest Technical Companies are competing to herald the speech accuracy title. In 2015, IBM achieved a word ERROR RATE pf 6.8%. IN 2016 Microsoft overpassed IBM with a 5.8 % claim. Shortly After that IBM improved their Rate to 5.4 %. However it's Google that claims the lowest Ratio rate at 4.8percent.

Research in the domains of voice-controlled assistants and natural language processing (NLP) has experienced notable advancements in recent years, revolutionizing human-computer interaction. Studies, such as those referenced in, underscore the remarkable progress made in speech recognition algorithms. These advancements have facilitated the accurate transcription of human speech into text, serving as the cornerstone for the development of sophisticated voice-controlled applications. By effectively converting spoken language into machine-readable text, these algorithms have propelled the evolution of voice-controlled assistants, enabling users to interact with technology in a more intuitive and natural manner.



Moreover, research in human-computer interaction (HCI), as highlighted in , has emphasized the significance of designing user interfaces that are intuitive and user-friendly for voice-controlled systems. Recognizing the importance of accessibility and usability, HCI studies have contributed invaluable insights into interface design principles tailored specifically for voice-based interactions. By prioritizing intuitive design, researchers aim to enhance the overall user experience and promote wider adoption of voice-controlled technologies across diverse user demographics.



Furthermore, the literature on automation and scripting, as referenced in , has played a pivotal role in enhancing the efficiency and functionality of voice-controlled assistants. Through the exploration of techniques for executing system commands and automating repetitive tasks, researchers have empowered these assistants to perform a myriad of functions seamlessly. From opening applications to retrieving information from the internet, these automation techniques have expanded the capabilities of voice-controlled assistants, making them indispensable tools for productivity and convenience in various contexts.



Collectively, these research endeavors have propelled the development of voice-controlled assistants and NLP technologies, reshaping the way individuals interact with and utilize technology in their daily lives. As researchers continue to innovate and refine these technologies, the potential for further advancements in voice-controlled assistants and NLP remains vast, promising even greater levels of functionality, accessibility, and user satisfaction in the future.

## **Problem Formulation**

Despite the convenience offered by existing voice-controlled assistants in executing tasks via voice commands, their functionality is often constrained by limitations in adaptability and resilience to diverse user interactions. Key challenges encompass compromised accuracy in recognizing commands, insufficient support for intricate tasks, and a reliance on stable internet connectivity. To mitigate these issues, this project endeavors to pioneer a novel voice-controlled assistant that transcends these limitations.

The primary objective of this project is to develop a sophisticated voice-controlled assistant capable of accurately deciphering user commands across a spectrum of complexities. By harnessing advanced speech recognition algorithms and leveraging natural language processing (NLP) techniques, the assistant aims to achieve unparalleled precision in understanding user inputs. Furthermore, the assistant will be engineered to tackle a wide array of tasks, ranging from basic commands like opening applications to more intricate operations such as retrieving specific information from databases or executing system-level commands.

A paramount focus of this project lies in enhancing the assistant's robustness and adaptability, ensuring seamless performance even in diverse and dynamic user environments. By minimizing dependencies on external factors like internet connectivity, the assistant seeks to guarantee uninterrupted functionality, thereby bolstering user confidence and satisfaction. Moreover, the assistant will be designed to seamlessly integrate into users' existing workflows, facilitating a frictionless user experience and enhancing productivity across various domains.

Through meticulous development and rigorous testing, this project aims to push the boundaries of voice-controlled assistant technology, ushering in a new era of efficiency, reliability, and user-centricity. By addressing the inherent limitations of existing solutions, this endeavor aspires to redefine the landscape of voice interaction, empowering users with an intuitive, versatile, and indispensable tool for navigating the digital realm.

## **Objectives**



**Chandigarh Engineering College Jhanjeri**  
**Mohali-140307**  
**Department of Computer Science & Engineering**

This project aims to create a sophisticated voice-controlled assistant using Python and specialized libraries for speech recognition and text-to-speech synthesis. The assistant will be equipped with diverse functionalities, including opening applications, retrieving information from the internet, and executing system commands, all driven by user voice commands.

A key focus of this project is to enhance the accuracy and robustness of command recognition through iterative refinement and optimization of speech recognition algorithms. By continually fine-tuning the algorithms, the assistant strives to accurately interpret user inputs, ensuring seamless execution of tasks with minimal errors.

Furthermore, the project prioritizes user interaction and usability by implementing intuitive feedback mechanisms and robust error handling. The assistant will provide clear and informative feedback to users, guiding them through the interaction process and addressing any errors or uncertainties encountered along the way. This proactive approach to user interaction aims to enhance user satisfaction and confidence in the assistant's capabilities.

To validate the effectiveness of the voice-controlled assistant, comprehensive performance evaluation will be conducted through user testing and feedback analysis. By soliciting feedback from users and assessing their interaction experiences, the project seeks to identify areas for improvement and refine the assistant's functionality to better align with user needs and expectations.

**Software which can be used for: a) Speech recognition  
b) Web searches c) Word guessing**



## **Methodology/Planning of work**

The methodology and planning of work for this project are structured into four key phases:

In the first month, the focus is on research and requirement analysis. A comprehensive literature review will be conducted to explore the latest advancements in speech recognition algorithms, natural language processing techniques, and existing voice-controlled assistants. This phase will also involve defining the functional requirements and designing the user interface for the voice-controlled assistant based on the insights gathered from the literature review.

The subsequent phase, spanning from the second to the fourth month, entails implementation and testing. During this period, the voice-controlled assistant will be developed using Python and relevant libraries. Core functionalities such as speech recognition, command interpretation, and task execution will be implemented. Rigorous testing will be conducted to evaluate the accuracy, robustness, and usability of the assistant, ensuring that it meets the defined requirements and performs effectively in real-world scenarios.

Following the implementation phase, months five and six will be dedicated to refinement and optimization. Based on feedback obtained from testing, the speech recognition algorithms and user interface will be refined to address any identified issues or areas for improvement. Performance and efficiency will be optimized through code optimization and enhancements to the underlying algorithms, with the aim of further enhancing the assistant's functionality and user experience.

The final phase, scheduled for the seventh month, involves evaluation and deployment. The performance and usability of the voice-controlled assistant will be assessed through user testing and feedback analysis. Any necessary adjustments will be made based on the feedback received, ensuring that the assistant meets user expectations and requirements. Finally, preparations will be made for the deployment and distribution of the finalized product, marking the culmination of the project.





### **Facilities required for proposed work**

For hardware requirements, a standard desktop or laptop computer equipped with a microphone and speaker is necessary to facilitate speech input and output for the voice-controlled assistant.

On the software front, the project necessitates a Python programming environment supplemented by essential libraries including pyttsx3 for text-to-speech synthesis, speech\_recognition for speech recognition capabilities, pyautogui for automation tasks, and wikipedia for retrieving information from online sources. Access to the internet is crucial for fetching data from online repositories.

Regarding the development environment, an integrated development environment (IDE) such as PyCharm or Jupyter Notebook is recommended for coding, debugging, and testing purposes. These platforms provide comprehensive tools and features to streamline the development process and ensure code efficiency and reliability.

Lastly, collaboration tools are indispensable for facilitating effective team coordination and project management. Communication tools such as Slack or Microsoft Teams, along with project management platforms like Trello or Asana, enable seamless collaboration, task assignment, and progress tracking among team members, fostering a cohesive and productive development environment.

## **REFERENCES**

### **BOOKS :**

- 1. G. L. Clapper, "Automatic word recognition", IEEE Spectrum, pp. 57-59, Aug. 1971.**
  - a. View Article Full Text: PDF (8868KB) Google Scholar**
  
- 2. M. B. Herscher, "Real-time interactive speech technology at threshold technology", Workshop Voice Technol. Interactive Real Time Command Control Syst. Appl., 1977-Dec.**
  - a. Google Scholar**
  
- 3. 3. J. W. Gleen, "Template estimation for word recognition", Proc. Conf. Pattern Recog. Image Processing, pp. 514-516, 1978-June.**
  - a. Google Scholar**

### **Internet:**

- 1. <https://pypi.org/project/SpeechRecognition>.**
- 2. [https://www.researchgate.net/publication/337155654\\_A\\_Study\\_on\\_Automatic\\_Speech\\_Recognition](https://www.researchgate.net/publication/337155654_A_Study_on_Automatic_Speech_Recognition).**
- 3. <https://www.ijedr.org/papers/IJEDR1404035.pdf>**