

Speech Recognition - Converting Spoken Language into Text

Bairi Kavya

Centurion University of Technology and Management, R. Sitapur, Odisha, India

Tangudu Swathi

Centurion University of Technology and Management, R. Sitapur, Odisha, India

Pyla Purnima

Centurion University of Technology and Management, R. Sitapur, Odisha, India

Revanth Jada

Centurion University of Technology and Management, R. Sitapur, Odisha, India

Dhawaleswar Rao CH

Centurion University of Technology and Management, R. Sitapur, Odisha, India

Abstract

This project explores the development and implementation of a speech recognition system designed to convert spoken language into text with high accuracy. Speech recognition technology has gained significant traction in various domains, including personal assistants, transcription services, accessibility tools, and voice-controlled applications. The core objective of this project is to create a robust framework capable of interpreting natural language from diverse speech inputs, accounting for variations in accent, tone, and background noise. The system employs advanced machine learning techniques, including deep neural networks and natural language processing algorithms, to process audio signals and extract meaningful text. The project leverages state-of-the-art models such as Hidden Markov Models (HMMs), Deep Neural Networks (DNNs), and Recurrent Neural Networks (RNNs) to train the system on large speech datasets, facilitating improved recognition performance. Additionally, techniques such as noise reduction, feature extraction, and language modeling are utilized to enhance the accuracy and efficiency of the system, especially in real-time applications. Challenges in speech recognition, such as handling multiple languages, dialects, homophones, and noisy environments, are addressed through custom solutions and data augmentation methods. The system is evaluated using a combination of standard metrics like Word Error Rate (WER) and user-based testing to assess its performance across different scenarios. Ultimately, this project contributes to the ongoing advancement of humancomputer interaction and has the potential to improve accessibility, automation, and user experience across various industries.

