

---

# Artificial Intelligence-Speech Recognition

AI is the study of the abilities of computers to perform tasks, which currently are better done by humans. AI has an interdisciplinary field where computer science intersects with philosophy, psychology, engineering and other fields.

(AI) is defined as intelligence exhibited by an artificial entity to solve complex problems and such a system is generally assumed to be a computer or machine. Artificial Intelligence is an integration of computer science and physiology Intelligence in simple language is the computational part of the ability to achieve goals in the world. Artificial intelligence concerned with making computers behave like humans more human-like fashion and in much less time than a human takes.

Pamela McCorduck writes, AI began with "an ancient wish to forge the gods. The seeds of modern AI were planted by classical philosophers who attempted to describe the process of human thinking as the mechanical manipulation of symbols. This work culminated in the invention of the programmable digital computer in the 1940s, the machine based on the abstract essence of mathematical reasoning.

The field of AI research was founded at a workshop held on the campus of Dartmouth College during the summer of 1956. Many of them predicted that a machine as intelligent as a human being would exist in no more than a generation and they were given millions of dollars to make this vision come true.

In 1973, in response to the criticism of James Lighthill and ongoing pressure from Congress, the U.S. and British Governments stopped funding undirected research into artificial intelligence, and the difficult years that followed would later be known as an "AI winter". Japanese Government inspired governments and industry to provide AI with billions of dollars, but by the late 80s the investors became disillusioned by the absence of the needed computer power (hardware) and withdrew funding again. Investment and interest in AI boomed in the first decades of the 21st century when machine learning was successfully applied to many problems in academia and industry due to the presence of powerful computer hardware. As predicted the imminent arrival of artificial general intelligence: a machine with intellectual capabilities that exceed the abilities of human beings.

Speech recognition is the process of extracting text transcriptions or some form of meaning from speech input. Speech analytics can be considered as the part of the voice processing, which converts human speech into digital forms suitable for storage or transmission computers.

Models :

Acoustic Model

An acoustic model is used in Automatic Speech Recognition to represent the relationship between an audio signal and the phonemes or other linguistic units that make up speech. The model is learned from a set of audio recordings and their corresponding transcripts. It is created

---

by taking audio recordings of speech, and their text transcriptions, and using software to create statistical representations of the sounds that make up each word.

## Language Model

A statistical language model is a probability distribution over sequences of words. Given such a sequence, say of length  $m$ , it assigns a probability  $\{displaystyle P(w_{\{1\}}, \dots, w_{\{m\}})\}$  to the whole sequence. Having a way to estimate the relative likelihood of different phrases is useful in many natural language processing applications, especially ones that generate text as an output.

In speech recognition, the computer tries to match sounds with word sequences. The language model provides context to distinguish between words and phrases that sound similar.

## Types of Recognition :

There are two types of speech recognition. One is called speaker-dependent and the other speaker-independent.

Speaker-dependent solutions are found in specialized use cases where there a limited number of words that need to be recognized with high accuracy.

Speaker-independent software is designed to recognize anyone's voice, so it requires no training. This means it is the only concrete option for applications such as interactive voice response systems — where businesses can't ask callers to read pages of text before using the system.

## Advantages :

delivers a great customer experience while improving self-service system's containment rate

encourages natural, human-like conversations that create more satisfying self-service interactions with customers

automates what touchtone cannot by collecting dynamic data such as names and addresses

enables organizations to save agents for more important tasks

## Disadvantages :

Lack of Accuracy and Misinterpretation.

Time Costs and Productivity

Accents and Speech Recognition

Background Noise Interference

Physical Side Effects

---

Applications :

In-car systems

Medical documentation

High-performance fighter aircraft

Usage in education and daily life

Speech recognition will revolutionize the way people conduct business over the Web and will, ultimately, differentiate world-class e-businesses. It is important to consider the environment in which the speaker system has to work. The grammar used by the speaker and accepted by the system, noise level, noise type, position of the microphone, and speed and manner of the user's speech are some factors that may affect the quality of speech recognition.

gradesfixer.com