



# Speech Recognition

Instructor: Vision Wang

Email: [xinwang35314@gmail.com](mailto:xinwang35314@gmail.com)

# Speech Recognition

- History
- Definition
- How does it work?
- Applications
- Practice using Python



# History of Speech Recognition



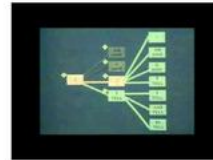
1784

**Thomas Edison** invents the first dictation machine



1952

**IBM Shoebox** can understand 16 English words



1971

**IBM Tangora**, using the Hidden Markov Model, predicts upcoming phonemes in speech



2006

**The National Security Agency (NSA)** starts using speech recognition to isolate key words in recorded speech

**Google** launches a voice search app, bringing speech recognition to mobile devices



2011

**Apple** announces Siri, ushering in the age of the voice-enabled digital assistant



2008

Wolfgang von Kempelen creates the **Acoustic-Mechanical Speech Machine** in Vienna



1879

Bell Labs releases **Audrey**, capable of recognizing spoken digits with **90% accuracy** - but only when spoken by its inventor



1962

**Harpy**, created at Carnegie Mellon University, can comprehend **1,011 words** - and some phrases



1986

# Definition of Speech Recognition

- Speech Recognition is a process in which a computer or device record the speech of humans and convert it into text format.
- It is also known as **Automatic Speech Recognition (ASR)**, **computer speech recognition** or **Speech To Text (STT)**.
- Linguistics, computer science, and electrical engineering are some fields that are associated with Speech Recognition.



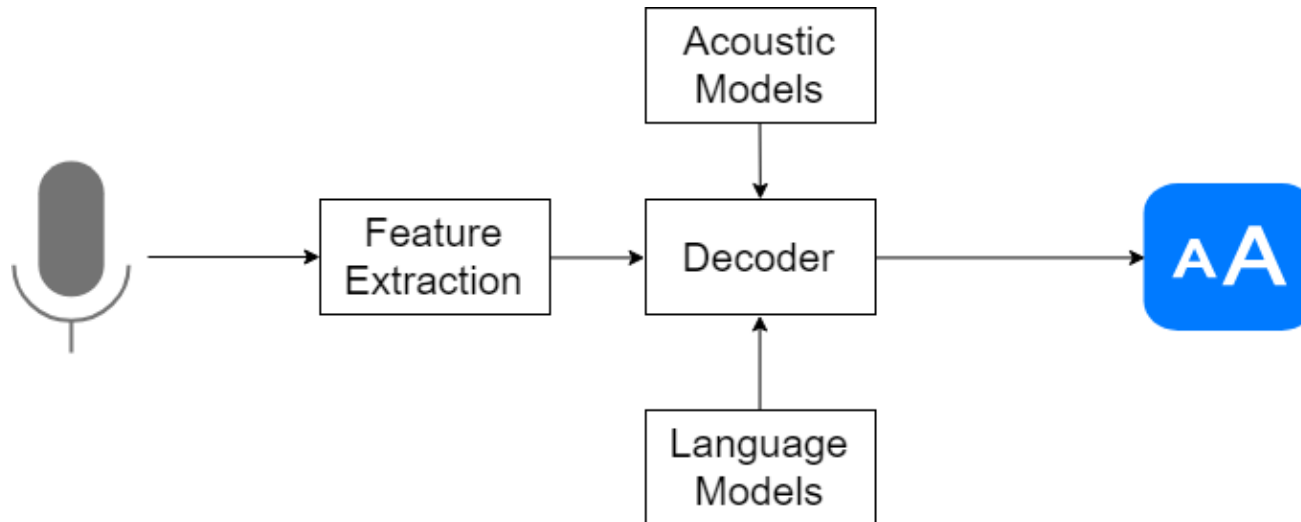
# How it works?

It is based on the algorithm of acoustic and language modeling.

- **Acoustic modeling** represents the relationship between linguistic units of speech and audio signals.
- **Language modeling** matches sounds with word sequences to help distinguish between words that sound similar.

**Any speech recognition program is evaluated using two factors:**

- **Accuracy** (percentage error in converting spoken words to digital data).
- **Speed** (extent to which the program can keep up with a human speaker).



## With IoT

### In Office



### In Healthcare



### Learning Languages



## Applications

The most frequent applications of speech recognition are as follow:

- In-car systems.
- Health care – Medical documentation and Therapeutic use.
- Military – High performance fighter aircraft, Helicopters, Training air traffic controllers.
- Usage in Education and Daily life.
- People with disabilities.

# Practice using Python

Let us create a new Speech Recognition project.

We have to install two modules for implement speech recognition.

- **SpeechRecognition**
- **PyAudio**

## Install Speech Recognition

- Go to terminal and type:
  - > pip install SpeechRecognition

SpeechRecognition is a library that helps in performing speech recognition in python. It support for several engines and APIs, online and offline e.g. Google Cloud Speech API, Microsoft Bing Voice Recognition, IBM Speech to Text etc.

## Install PyAudio

- Go to terminal and type:
  - > pip install pyaudio

PyAudio provides [Python](#) bindings for [PortAudio](#), the cross-platform audio I/O library. With PyAudio, you can easily use Python to play and record audio on a variety of platforms, such as GNU/Linux, Microsoft Windows, and Apple Mac OS X / macOS.

# Let's perform Speech Recognition!

```
import speech_recognition as sr

r = sr.Recognizer()

with sr.Microphone() as source:
    print("Please speak anything:")
    audio = r.listen(source)

    try:
        text = r.recognize_google(audio)
        print("You said: {}".format(text))
    except:
        print("Sorry could not recognize your voice.")
```



# Search Wikipedia using your voice

- **Wikipedia** is a Python module that makes it easy to access and parse data from Wikipedia.
- Search Wikipedia, get article summaries, get data like links and images from a page, and more.
- Wikipedia wraps the **MediaWiki API** so you can focus on using Wikipedia data.

Install Wikipedia

```
> pip install wikipedia
```

```
>>>import wikipedia  
>>>print(wikipedia.summary("Artificial Intelligence"))
```

**Q:** Try to use SpeechRecognition and Wikipedia modules to create a robot which can receive your voice and search wikipedia for you. Enjoy the fun!

# Translating the voice

We can use Google Translate Ajax API to detect and translate language.

Install Googletrans

> pip install googletrans

## Detect language

```
>>>from googletrans import Translator
>>>translator = Translator()
>>>s=translator.translate('안녕하세요.')
>>>print(s)
```

**Q:** Can you try to use Googletrans into last program to translate the language of the voice to a certain language?

## Translate your language to a certain language

```
>>>phrase = translator.translate('안녕하세요.', dest="zh-cn")
>>>print(phrase)
```