음성인식(Speech Recognition)

1. 설치

"pip install SpeechRecognition", "conda install PyAudio"를 명령을 이용하여 SpeechRecongnition, PyAudio를 설치하였다. (pip install PyAudio 명령은 에러가 나서 conda install PyAudio 명령으로 해결하였다.)

```
C:₩Users₩KimMinyoung>pip install SpeechRecognition
Requirement already satisfied: SpeechRecognition in c:₩users₩kimminyoung₩anaconda3₩
```

```
C:₩Users₩KimMinyoung>conda install PyAudio
Collecting package metadata (current_repodata.json): done
Solving environment: done
```

2. 테스트

```
In [1]: import speech_recognition as sr
        sr.__version__
Out[1]: '3.8.1'

In [2]: r = sr.Recognizer()
        harvard = sr.AudioFile('c:/Users/KimMinyoung/python-speech-recognition-master/python-speech-recognition-master/audio_files/harvard.wav

In [3]: with harvard as source:
            audio = r.record(source)
        type(audio)
Out[3]: speech_recognition.AudioData

In [4]: r.recognize_google(audio)
Out[4]: 'the stale smell of old beer lingers it takes heat to bring out the odor a cold dip restores health and zest a salt pickle taste fine
        with ham tacos al Pastore are my favorite a zestful food is be hot cross bun'
```

참고사이트에서 안내한 https://github.com/realpython/python-speech-recognition에서 예제를 실행하기 위한 음성파일(harvard.wav, jackhammer.wav)을 다운받았다. 이후 Jupyter notebook 에서 recognize_google 메소드를 이용하여 테스트를 해보았다. 음성파일이 텍스트파일로 어느정도 잘 인식이 된 것을 확인할 수 있다.

Capturing Segments with offset and duration

```
In [5]: with harvard as source:
            audio = r.record(source,duration = 4)
        r.recognize_google(audio)
Out[5]: 'the stale smell of old beer lingers'

In [6]: with harvard as source:
            audio1 = r.record(source, duration = 4)
            audio2 = r.record(source, duration = 4)

In [7]: r.recognize_google(audio1)
Out[7]: 'the stale smell of old beer lingers'

In [8]: r.recognize_google(audio2)
Out[8]: 'it takes heat to bring out the odor a cold dip'

In [9]: with harvard as source:
            audio = r.record(source, offset=4, duration=3)
        r.recognize_google(audio)
Out[9]: 'it takes heat to bring out the odor'
```

The Effect of Noise on Speech Recognition

```
In [11]: jackhammer = sr.AudioFile('c:/Users/KimMinyoung/python-speech-recognition-master/python-speech-recognition-master/audio_files/jackhamm
         with jackhammer as source:
             audio = r.record(source)
         r.recognize_google(audio)

Out[11]: 'the stale smell of old gear vendors'

In [12]: with jackhammer as source:
             r.adjust_for_ambient_noise(source, duration=0.5)
             audio = r.record(source)
         r.recognize_google(audio)

Out[12]: 'the snail smell of old gear vendors'

In [13]: r.recognize_google(audio, show_all=True)

Out[13]: {'alternative': [{'transcript': 'the snail smell of old gear vendors',
             'confidence': 0.79132015},
            {'transcript': 'the snail smelly old gear vendors'},
            {'transcript': 'the snail smelled old gear vendors'},
            {'transcript': 'the snail smell old gear vendors'},
            {'transcript': 'the snails smell of old gear vendors'}],
          'final': True}
```

또한 마이크를 이용하여 직접 말한 영어를 텍스트로 변환하여 보기도 했다. 아래와 같이 마이크를 설정한 후 코드를 실행시켰다. Hello라고 발음했고 잘 인식이 된 것을 확인할 수 있다.

The Microphone Class

```
In [31]: import speech_recognition as sr
         r = sr.Recognizer()
         sr.Microphone.list_microphone_names()

Out[31]: ['Microsoft 사운드 매퍼 - Input',
          '마이크(Conexant HD Audio)',
          'Microsoft 사운드 매퍼 - Output',
          '스피커(Conexant HD Audio)',
          'Speakers (Conexant HD Audio output)',
          '마이크 (Conexant HD Audio capture)',
          '헤드폰 ()',
          '머리에 거는 수화기 (@System32₩₩drivers₩₩bthhfenum.sys,#2;%1 Hands-Free AG Audio%0₩r₩n;(퐁뎅이))',
          '머리에 거는 수화기 (@System32₩₩drivers₩₩bthhfenum.sys,#2;%1 Hands-Free AG Audio%0₩r₩n;(퐁뎅이))']

In [32]: mic = sr.Microphone(device_index=1) # 마이크 : device_index=1

In [33]: with mic as source:
             print("시작");
             audio = r.listen(source)
             print("끝")

         try:
             print("You said " + r.recognize_google(audio))
         except:
             print("Could not understand audio")

         시작
         끝
         You said hello

In [34]: r.recognize_google(audio)

Out[34]: 'hello'
```

Game 예제 또한 잘 실행이 된 것을 확인할 수 있었다. 다만 나는 graph라고 말했으나, gray라고 인식하는 것도 볼 수 있었다.

```
I'm thinking of one of these words:
apple, banana, grape, orange, mango, lemon
You have 3 tries to guess which one.

Guess 1. Speak!
You said: Apple
Incorrect. Try again.

Guess 2. Speak!
You said: Gray
Incorrect. Try again.

Guess 3. Speak!
You said: Orange
Sorry, you lose!
I was thinking of 'grape'.
```