

## SIT 789

### Task 9.1: Speech recognition with IBM Watson

#### Task 1

Audio files:

arctic\_a0001.wav and p232\_014.wav

```
C:\Users\singh\Documents\MastersAppliedAI\Deakin\SIT789_CompVision_SpeechProcessing\Practicals\SpeechToTextData>curl -X POST -u "apikey:aRUa2wF6NEWP0IQjcu8YEi0QbZa-B2Km-20P9p2zo1E" --header "Content-Type: audio/wav" --data-binary @arctic_a0001.wav "https://api.au-syd.speech-to-text.watson.cloud.ibm.com/instances/cdd27c84-7f44-4e34-8a1d-7aa0e1f46c0c/v1/recognize"
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "author of the danger trail Phillips steals etcetera ",
          "confidence": 0.78
        }
      ]
    }
  ]
}

C:\Users\singh\Documents\MastersAppliedAI\Deakin\SIT789_CompVision_SpeechProcessing\Practicals\SpeechToTextData>curl -X POST -u "apikey:aRUa2wF6NEWP0IQjcu8YEi0QbZa-B2Km-20P9p2zo1E" --header "Content-Type: audio/wav" --data-binary @p232_014.wav "https://api.au-syd.speech-to-text.watson.cloud.ibm.com/instances/cdd27c84-7f44-4e34-8a1d-7aa0e1f46c0c/v1/recognize"
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "the Hebrews it was a token that there would be no more universal floods ",
          "confidence": 0.92
        }
      ]
    }
  ]
}
```

#### Task 2:

We can observe that using IBM speech-to-text service we can convert speech samples using command line as well as API in code to get JSON results.

# assignment\_9\_1

September 25, 2022

```
[1]: from ibm_watson import SpeechToTextV1
from ibm_cloud_sdk_core.authenticators import IAMAuthenticator
import json
import glob
import os
```

```
[2]: authenticator = IAMAuthenticator('aRUa2wvF6NEWP0lQjcu8YEi0QbZa-B2Km-20P9p2oqlE') #replace
    ↳ {APIkey} by your API key
speech_to_text = SpeechToTextV1(
    authenticator=authenticator
)
speech_to_text.set_service_url('https://api.au-syd.speech-to-text.watson.cloud.
    ↳ ibm.com/instances/cdd27c84-7f44-4e34-8a1d-7aa0e1f46c0c') #replace {url} by
    ↳ your URL
```

```
[3]: with open('SpeechtoTextData/arctic_a0005.wav', 'rb') as audio_file:
    speech_recognition_results = speech_to_text.recognize(
        audio = audio_file,
        content_type='audio/wav').get_result()
print(json.dumps(speech_recognition_results, indent = 2))
```

```
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "will we ever forget it ",
          "confidence": 0.95
        }
      ]
    }
  ]
}
```

```
[4]: with open('SpeechtoTextData/arctic_a0005.json', 'w') as outfile:
      json.dump(speech_recognition_results, outfile)
```

```
[5]: with open('SpeechtoTextData/arctic_a0005.json') as infile:
      data = json.load(infile) # load data from a json file
      print(data)
```

```
{'result_index': 0, 'results': [{'final': True, 'alternatives': [{'transcript':
'will we ever forget it ', 'confidence': 0.95}]]}]}
```

```
[10]: files = glob.glob("SpeechtoTextData/*")
      for file in files:
          filename, file_extension = os.path.splitext(file)
          print("Processing: ", filename, file_extension)
          if file_extension == ".wav" or file_extension == ".flac":
              audio_file = open(file, 'rb')
              speech_recognition_results = speech_to_text.recognize(
                  audio = audio_file,
                  content_type='audio/'+file_extension[1:]).get_result()
              print(json.dumps(speech_recognition_results, indent = 2))
          else:
              print("Unsupported extension")
```

Processing: SpeechtoTextData\367-130732-0000 .flac

```
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "lobsters and law officers ",
          "confidence": 0.74
        }
      ]
    }
  ]
}
```

Processing: SpeechtoTextData\367-130732-0001 .flac

```
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "when is a lobster not a lobster when it is a crayfish
",
```

```

        "confidence": 0.87
    }
]
}
Processing: SpeechtoTextData\367-130732-0004 .flac
{
    "result_index": 0,
    "results": [
        {
            "final": true,
            "alternatives": [
                {
                    "transcript": "a book could be written about this restaurant and then
all would not be told for all its secrets can never be known ",
                    "confidence": 0.92
                }
            ]
        }
    ]
}
Processing: SpeechtoTextData\arctic_a0001 .wav
{
    "result_index": 0,
    "results": [
        {
            "final": true,
            "alternatives": [
                {
                    "transcript": "author of the danger trail Phillips steals etcetera ",
                    "confidence": 0.78
                }
            ]
        }
    ]
}
Processing: SpeechtoTextData\arctic_a0003 .wav
{
    "result_index": 0,
    "results": [
        {
            "final": true,
            "alternatives": [
                {
                    "transcript": "for the twentieth time that evening the two men shook
hands ",
                    "confidence": 0.86
                }
            ]
        }
    ]
}

```

```

    }
  ]
}
]
}
Processing: SpeechtoTextData\arctic_a0005 .json
Unsupported extension
Processing: SpeechtoTextData\arctic_a0005 .wav
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "will we ever forget it ",
          "confidence": 0.95
        }
      ]
    }
  ]
}
]
}
Processing: SpeechtoTextData\p232_009 .wav
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "there is according to legend a boiling pot of gold at
one end ",
          "confidence": 0.98
        }
      ]
    }
  ]
}
]
}
Processing: SpeechtoTextData\p232_010 .wav
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "no one ever finds ",
          "confidence": 0.19
        }
      ]
    }
  ]
}

```

```

    }
  ]
}
]
}
Processing: SpeechtoTextData\p232_014 .wav
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "the Hebrews it was a token that there would be no more
universal floods ",
          "confidence": 0.92
        }
      ]
    }
  ]
}
Processing: SpeechtoTextData\p232_030 .wav
{
  "result_index": 0,
  "results": [
    {
      "final": true,
      "alternatives": [
        {
          "transcript": "they were described as being in a serious but stable
condition yesterday ",
          "confidence": 0.94
        }
      ]
    }
  ]
}
Processing: SpeechtoTextData\transcripts .txt
Unsupported extension

```

[ ]: