NLP Assignment

Problem Statement:

XYZ is a call audit company that audits the promotional call of their employees. They perform the call audits manually where a person listens all the recorded call and decide whether all the promotional offers are detailed correctly to the customer or not. But this process is time consuming and involved lots of manual intervention. So the company decide to use Natural Language Understanding and develops a product where they deal with processing audio calls, convert the speech into text, identifying intent and entities from a text and generate the summary for the audio calls.

The objective of this assignment is to perform following tasks:

Task 1: Download and transcribe the given the audio file using Speech-to-Text recognition.

https://baitrainingdataset.blob.core.windows.net/interviewdata/sales_call_telephone_marketers.wav

You may use open-sourced Automatic Speech Recognition (ASR) models such as Wav2Vec or a free-tier service from providers like Azure, GCP, AWS, etc.

Task 2: Train an NLU model to classify intents and recognize entities.

Refer following examples for intents and entities. Feel free to create your own intent and entities.

Intent	Example Sentence	Entities Names with values
Intro	My name is Jeff and I am calling from Amazon.	caller_name: Jeff company: Amazon
Intro	I am calling from Microsoft and my name is Satya.	caller_name: Satya company: Microsoft
Intro	I am Sundar and this is call from Google.	caller_name: Sundar company: Google
Purpose	I am calling about your Microsoft Azure subscription.	product: Microsoft Azure
Purpose	This is a call regarding your Google Cloud Platform account.	product: Google Cloud Platform

Purpose	I would like to talk about your Amazon Web Services account.	product: Amazon Web Services

The words in blue are the entities that should be identified and extracted.

(Note: Above intent and entities are for reference the actual values for entities and intent may vary as per the audio file input)

Task 3: Separate the sentences in the output of task 1. On each sentence, apply the model trained in task 2 to classify its intent and recognize the entities present in it.

Export the output in a JSON file called output.json in the following format:


```
{
  "task_1_output": "My name is Sundar and I am calling from Microsoft. This is a call about your Amazon Web Services account.",
  "task_3_output": [
    {
      "sentence": "My name is Sundar and I am calling from Microsoft.",
      "intent": "intro",
      "entities": [
        {
          "entity_name": "caller_name",
          "entity_value": "Sundar"
        },
        {
          "entity_name": "company",
          "entity_value": "Microsoft"
        }
      ]
    },
    {
      "sentence": "This is a call about your Amazon Web Services account.",
      "intent": "purpose",
```

```
        "entities": [
          {
            "entity_name": "product",
            "entity_value": "Amazon Web Services"
          }
        ]
      }
    ]
}
```


Task 4: For the whole text generated from the audio file generate a summary report.

Submission Steps:

Share the link of the GitHub repository and also create a Readme file mentioning detailed steps for each task including the details about the API used for speech to text and the NLU model used for intent and entity classification. The repo should also include list of all dependencies for running this application, and any other information if needed.