Mawdoo3 AI Task

Introduction

Speech is the basic form of human-to-human communication, whereas the speech signals hold a lot of information about the speaker's identity and characteristics, such as; gender, emotion and dialect. Hence, the classification of the speaker's features gained a great importance in Natural Language Understanding (NLU) and Processing (NLP) applications.

The task

This task includes building a machine learning model/solution to detect the gender from speech. In other words, given an utterance that is spoken by any person, detect the gender of the speaker.

In this task, you will have to do the following steps:

- **Specify the required data; size, type, resources:**
  Write a brief description of the required data to train and test your solution, mention the required size in terms of hours and number of utterances and some possible resources.

- **Crawl the speech data from any open source resources:**
  Extract/crawl the data from any possible open source resources mentioned in the previous step. The speech files can be in any language, English would be the easiest choice as there are many well known speech datasets available.

- **Provide some exploratory analysis on the data with visualizations:**
  Do some exploratory analysis on the crawled data; e.g. utterances per class, lengths of the utterances, languages, age of the speakers, etc. Provide some visualizations of the data using any tool.

- **Extract features from the crawled data:**
  Extract features from the speech files using any features extraction technique and store them in a seperate file to be accessible at any time.

- **Build multiple models/solutions to detect the gender of the speaker:**
  A lot of networks can be used to detect the gender of the speaker; implement at least two networks with hyperparameter tuning and justify your choice for these networks.

- **Compare the implemented models/solutions:**
  Compare the results of the implemented models with different measurements and justify the results of each model with some error analysis.

- **Serve the resulted model in a REST API (Optional):**
  Choose the best model and wrap it with a REST API using any web framework.

- **Create a Docker image for your application (Optional).**

Deliverables:

- Github repository that includes all the codes used to crawl the data, extract features, train and test the models, etc.
- Readme file that briefly describes every step of the task with the required explanations.

Useful Resources:

- Python libraries: beautifulsoup, librosa, scipy, flask, tensorflow, Pytorch
- Google colab.

The time given for the task is two weeks since the day received.
After finishing the task, send the github repo link to the following emails:
hisham.bataineh@mawdoo3.com
muhy.zater@mawdoo3.com
qusai.abuobaida@mawdoo3.com
ayat.almalahmih@mawdoo3.com

Please do not hesitate to contact us at any time in case any explanation is needed.