

Tugas Pekan ke-12 *Intro to Speech Processing*

Batas pengumpulan: **Senin 14 Desember 2020, pukul 10.59 pagi**, melalui LMS

Deskripsi

Pengamatan eksperimen pengenalan suara berdasarkan tutorial di https://www.tensorflow.org/tutorials/audio/simple_audio.

Pengamatan yang dilakukan

Amati hasil tes pengenalan suara untuk **setiap** perintah *command*: “down”, “go”, “left”, “no”, “right”, “stop”, “up” dan “yes”. Pada tutorial tersebut diberikan contoh pengujian untuk suara kata “no”.

```
sample_file = data_dir, 'no/01bb6a2a_nohash_0.wav'

sample_ds = preprocess_dataset([str(sample_file)])

for spectrogram, label in sample_ds.batch(1):
    prediction = model(spectrogram)
    plt.bar(commands, tf.nn.softmax(prediction[0]))
    plt.title(f'Predictions for "{commands[label[0]]}"')
    plt.show()
```

Path file menunjukkan direktori tipe command dan nama file. Pilih salah satu nama file yang ada di tiap direktori, bebas file yang mana saja. Cara untuk mendapatkan nama file di sebuah folder command dapat dilihat di tutorial berikut: https://colab.research.google.com/drive/1yY-I50GofTjQqpVbdVfjGFzccjNUz_U7?usp=sharing

Cek nama file di folder sebuah command

```
!ls 'data/mini_speech_commands/yes'
```

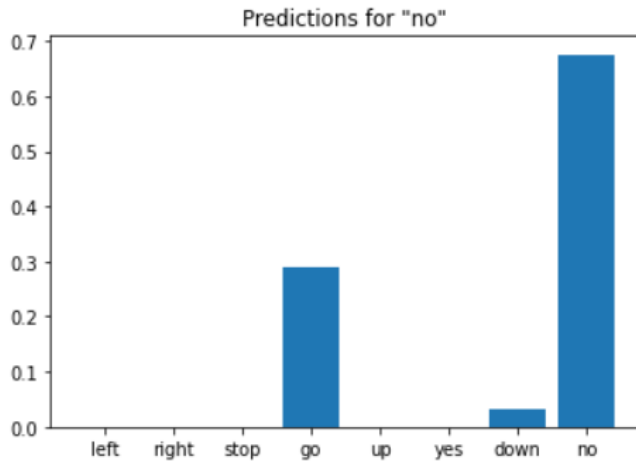
264f471d_nohash_3.wav	7e1054e7_nohash_1.wav	ccf418a5_nohash_1.wav
2796ac50_nohash_1.wav	7e843da8_nohash_1.wav	cd7f8c1b_nohash_1.wav
27c30960_nohash_0.wav	7ea032f3_nohash_1.wav	cd85758f_nohash_4.wav
28ce0c58_nohash_3.wav	7ea032f3_nohash_2.wav	ceaadb24_nohash_2.wav

Misal saya akan melakukan pengujian untuk *command* yes, lalu saya akan memilih file ‘00f0204f_nohash_0.wav’, maka saya akan mengganti baris sample file dengan:

```
sample_file = data_dir/'yes/00f0204f_nohash_0.wav'
```

1. Amati visualisasi hasil prediksi pada diagram batang. Apakah terdapat tipe *command* lain yang mempunyai nilai prediksi > 0? Jika iya, menurut Anda, mengapa hal tersebut terjadi?

Sebagai contoh, di file tutorial tersebut, *command* no dikenali sebagai *command* go dan down dengan nilai prediksi > 0.



Lakukan untuk **semua** tipe command (**total 8 command**), masing-masing diwakili **1 file** saja.

2. File audio yang digunakan berada dalam format .wav. Apakah format audio tersebut dapat merepresentasikan suara dengan fitur huruf vokal dan huruf konsonan dengan cukup baik?
3. Pada proses klasifikasi, dilakukan transformasi fitur suara menjadi spectrogram. Mengapa hal tersebut perlu dilakukan? Apakah dengan fitur spectrogram fitur huruf vokal dan konsonan dapat direpresentasikan dengan lebih baik?

File yang harus dikumpulkan:

1. Laporan: 1 file pdf, maksimum panjang laporan adalah 5 halaman. Laporan berisi jawaban pertanyaan nomor 1, 2 dan 3 (tidak hanya berupa print screen hasil pengenalan suara), dan referensi jawaban nomor 2 dan 3.

Detail penilaian:

Laporan:

- kelengkapan jawaban nomor 1, 2 dan 3 [60 poin]
- ketepatan analisis jawaban nomor 1 [20 poin]
- ketepatan analisis jawaban nomor 2 dan 3 [20 poin]

Jika ada pertanyaan, silakan disampaikan melalui *channel* pekan_12_tugas_intro_speech_processing di slack, <https://pemrosesanbah-c5k6846.slack.com/archives/C01G9HETDL2>