

Projet 3AI 2022

Commande vocale

Description

Les interfaces vocales sont de plus en plus présentes dans notre vie quotidienne et vont rapidement devenir notre interface privilégiée pour communiquer avec les objets/machines que nous utilisons.

L'objectif de ce projet est de construire un algorithme qui comprend des commandes vocales simples : `yes` , `no` , `up` , `down` , `left` , `right` , `on` , `off` , `stop` , `go` . Tout les autres mots doivent être considérés comme inconnus ou du bruit de fond/silence.

-0:00

Résultat attendu

Construire un modèle qui interprète ces commandes vocales simples à partir d'un enregistrement.
Proposer une application/démonstrateur.

Données

Le jeu de données (format wav) comprend 105 000 mots d'une demi-seconde, composés de 30 mots courts, prononcés par des milliers de personnes différentes.

Téléchargement des données :

https://storage.cloud.google.com/download.tensorflow.org/data/speech_commands_v0.02.tar.gz

Ressources :

https://github.com/tensorflow/docs/blob/master/site/en/r1/tutorials/sequences/audio_recognition.md