Forestier Quentin & Herzig Melvyn

MLG – 13.04.2022

Voice recognition

Labo 3



# Introduction

Ce laboratoire demande d’appliquer la méthodologie apprise durant les derniers travaux pratiques pour évaluer les performances d’un réseau de neurones entrainé.

Pour ce faire, il faut réaliser 4 expériences qui comparent des voix d’hommes, de femmes et d’enfants, naturelles ou synthétiques.

Au terme de chaque expérience, il faut pouvoir définir le model final à sélectionner pour traiter le problème associé.

La sélection du modèle consiste à trouver le modèle dont la complexité et la configuration sont appropriées.

## Contenu commun

Les caractéristiques suivantes, seront communes à chaque expérience :

* Une couche d’entrée de 13 neurones.
  + Pour chaque expérience, nous récupérons 13 valeurs MFCC par fenêtre.
  + Nous travaillions avec soit la moyenne, soit l’écart type.
* Une couche cachée de longueur variable que nous souhaitons optimiser.
  + [2, 4, 6, 8, 10, 15, 20, 25, 30]
* Une couche de sortie variable
  + Si deux classe, 1 neurone.
  + Sinon, autant de neurones que de classes.
* Les poids synaptiques sont attribués aléatoirement.

# Expérience 1 - Hommes Vs Femmes avec une voix naturelle

Dans cet exemple, il s’agit d’entrainer le modèle à différencier des voix d’hommes naturelles à des voix de femmes naturelles.

Dans un premier temps, nous avons travaillé avec la moyenne des 13 valeurs MFCC.

### Nombres d’observations

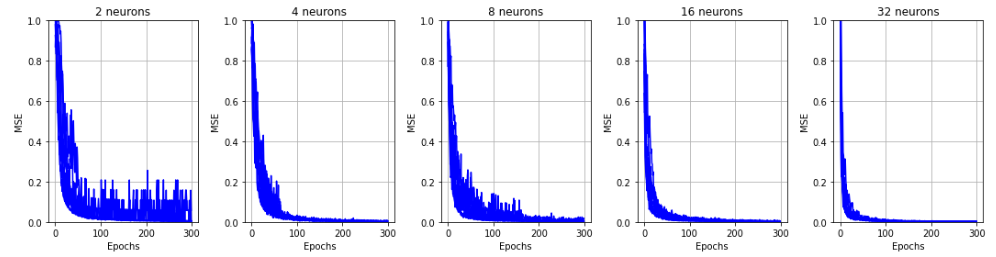
|  |  |  |
| --- | --- | --- |
| Total | Homme | Femme |
| 72 (100%) | 36 (50%) | 36 (50%) |
|  |  |  |

### Procédure

Dans un premier temps, nous avons estimé combien d’époques étaient nécessaires pour trouver un faible taux d’erreur d’entrainement.

**Initialisation**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **N\_INITS** | **EPOCHS** | **N\_NEURONS** | **LEARNING\_RATE** | **MOMENTUM** |
| 10 | 300 | 2,4,8,16,32 | 0.001 | 0.5 |



Dans chaque configuration de 2 à 32 neurones, nous