

Chemins spécifiques pour la classification dans les réseaux de neurones profonds

Bouzidi Belkassim - Elhouiti Chakib - Kezzoul Massili

Université de Montpellier

3 juin 2021

1 Introduction

Les réseaux de neurones profonds

Problématique

Solution proposée

2 Organisation

3 Analyse des données

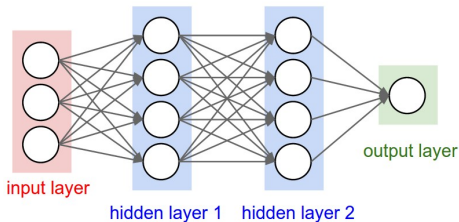
4 Développement de l'architecture

5 Analyse des résultats

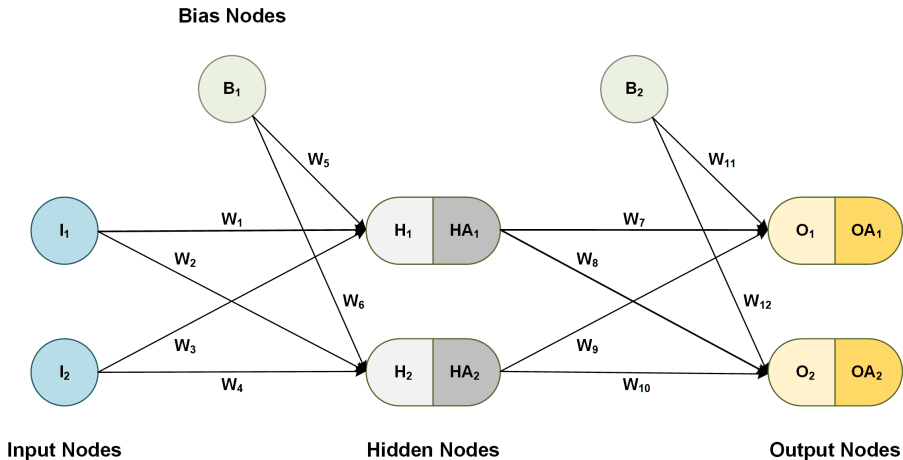
6 Conclusion

Présentation des réseaux de neurones

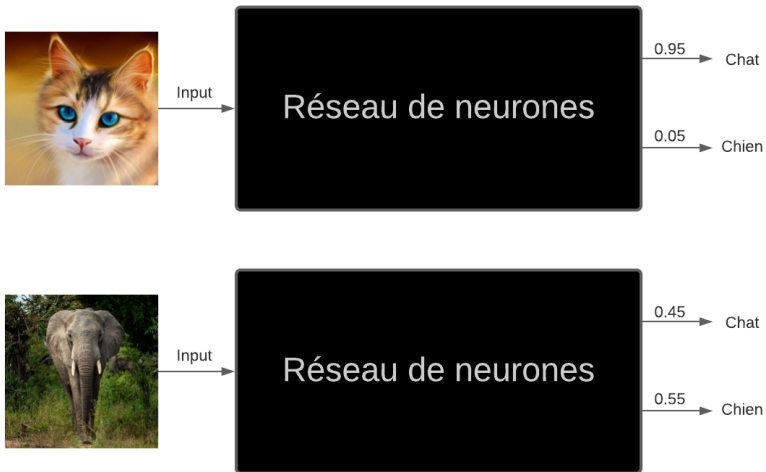
Les réseaux de neurones sont constitués de plusieurs couches consécutives de neurones interconnectées.



Fonctionnement



Boite noire



Problématique

Objectifs

L'objectif est de comprendre le fonctionnement interne d'un réseau de neurones et de repérer des signatures d'activations.

- ▶ À partir de quelle couche le modèle change de comportement pour reconnaître une image ?
- ▶ Les signatures des images de 7, sont-elles différentes de ceux des 1 ?
- ▶ Si on passe une image de 3 au modèle, à quoi va ressembler sa signature ?



UNIVERSITÉ
DE MONTPELLIER

Solution proposée

- ▶ Construire des réseaux de neurones.
- ▶ Récupérer, pour chaque donnée, la sortie des couches cachées.
- ▶ Extraire les signatures grâce à des algorithmes de *clustering*.
- ▶ Réaliser une interface de visualisation en utilisant différentes techniques.

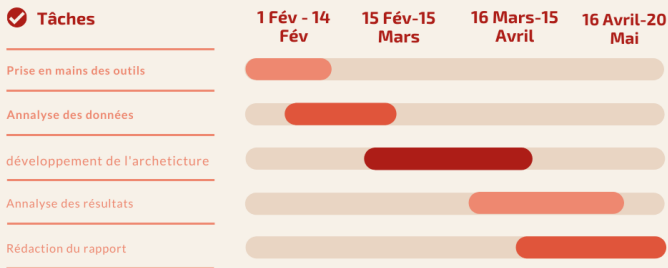


- ① Introduction
- ② Organisation**
- ③ Analyse des données
- ④ Développement de l'architecture
- ⑤ Analyse des résultats
- ⑥ Conclusion

Organisation du projet

Chemins spécifiques

✓ Tâches



1 Introduction

2 Organisation

3 Analyse des données

Sélection des données

Prétraitements

4 Développement de l'architecture

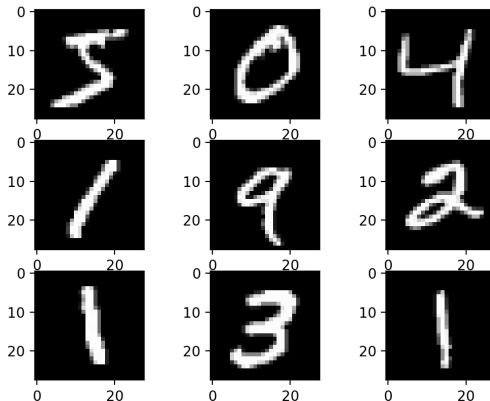
5 Analyse des résultats

6 Conclusion

Le jeu de données

Mixed National Institute of Standards and Technology

Base de données composée de 70000 images de chiffre manuscrit.



Sélection des données

- ▶ Garder un nombre précis d'images pour un ensemble de chiffres définis.
- ▶ Faciliter la phase de développement.
- ▶ Pouvoir mieux visualiser les résultats sur un petit ensemble de données.



Prétraitements

Scaling

Normalisation : Mettre les valeurs des images entre 0 et 1 au lieu de 0 et 255.

Flattening

Applatis les images pour avoir un tableau à une seule dimension au lieu d'une matrice.

One-hot encoding

Transformation des labels en un vecteur binaire contenant que des 0 et des 1.

- ▶ Pour un 1 \Rightarrow [1,0,0].
- ▶ Pour un 3 \Rightarrow [0,1,0].
- ▶ Pour un 7 \Rightarrow [0,0,1].

① Introduction

② Organisation

③ Analyse des données

④ Développement de l'architecture

Technologies utilisées

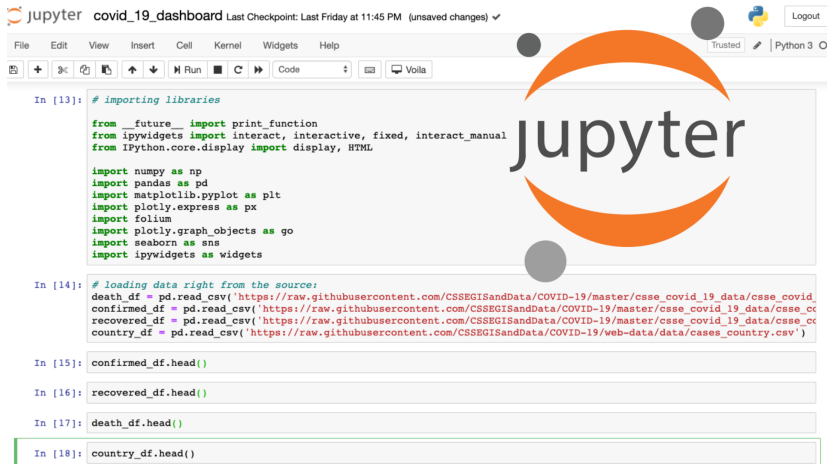
Modèle d'apprentissage

Interface de visualisation

⑤ Analyse des résultats

⑥ Conclusion

Jupyter notebook



```
In [13]: # importing libraries

from __future__ import print_function
from ipywidgets import interact, interactive, fixed, interact_manual
from IPython.core.display import HTML

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import plotly.express as px
import folium
import plotly.graph_objects as go
import seaborn as sns
import ipywidgets as widgets

In [14]: # loading data right from the source:
death_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data/csse_covid_19_data/confirmed_df.csv')
confirmed_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data/csse_covid_19_data/recovered_df.csv')
recovered_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data/csse_covid_19_data/country_df.csv')
country_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/web-data/data/cases_country.csv')

In [15]: confirmed_df.head()

In [16]: recovered_df.head()

In [17]: death_df.head()

In [18]: country_df.head()
```

Tensorflow, Keras



UNIVERSITÉ
DE MONTPELLIER

Voilà

jupyter Untitled

Se déconnecter

Python 3

Fichier Édition Affichage Insérer Cellule Noyau Widgets Aide

Exécuter Code

illustration Voilà

Voilà permet de transformer un notebook jupyter en une application web autonome (standalone).

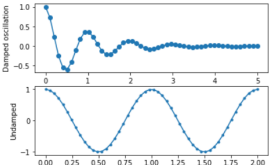
Entrée [4]:

```
import numpy as np
import matplotlib.pyplot as plt
```

Entrée [5]:

```
plot()
```

A tale of 2 subplots

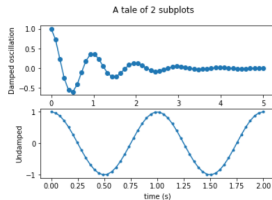


The figure displays two subplots. The top subplot, titled 'Damped oscillation', shows a blue line with circular markers that starts at 1.0 and decays towards 0.0 over a time interval from 0.00 to 5.00. The bottom subplot, titled 'Undamped', shows a blue line with circular markers that oscillates between 1.0 and -1.0 over a time interval from 0.00 to 2.00.



illustration Voilà

Voilà permet de transformer un notebook jupyter en une application web autonome (standalone).



Création du modèle

Modèle

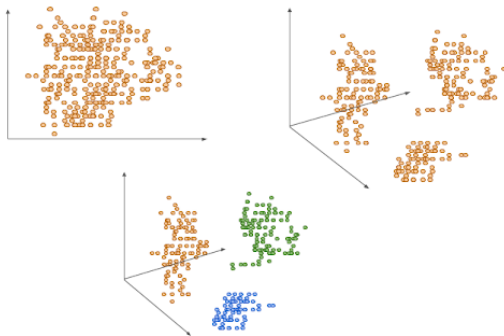
- ▶ 2 couches cachées ;
- ▶ 32 neurones pour la première ;
- ▶ 64 pour la deuxième ;
- ▶ fonction d'activation *relu* pour les couches internes ;
- ▶ et *softmax* pour la dernière couche.

```
1 self.model = Sequential([Dense(32, input_shape=input_shape, activation='relu'),
2   .....Dense(64, activation='relu'),
3   .....Dense(2, activation='softmax')])
4
5 self.model.compile(optimizer=Adam(
6   .....learning_rate=learning_rate), loss='mean_squared_error', metrics=['accuracy'])
```

Clustering

K-means

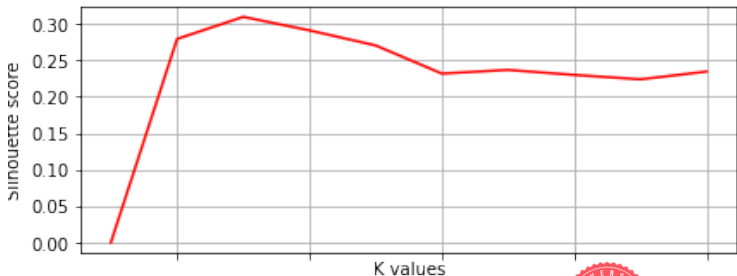
K-means prend en paramètres les données et un certain K donnée par l'utilisateur, puis construit K clusters qui regroupent les données qui sont proches (en terme de distance euclidienne).



Choix du K

Choix du K

Concrètement, cette méthode consiste à calculer pour un clustering, la moyenne du score *Silhouette* de chaque point.



UNIVERSITÉ
DE MONTPELLIER

UMAP

UMAP

(Uniform Manifold Approximation and Projection) Utilise des algorithmes de mise en page graphique pour organiser les données dans un espace de faible dimension.

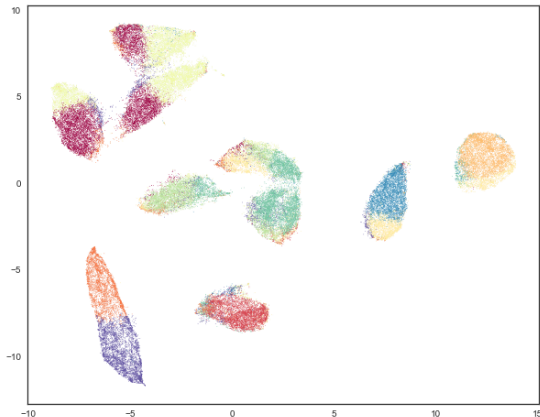
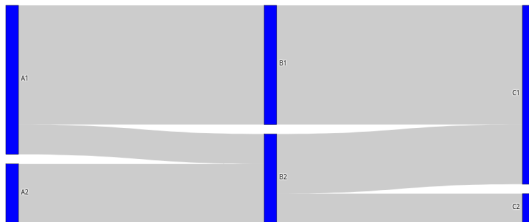


Diagramme de Sankey

Un diagramme de Sankey est un type de diagramme de flux dans lequel la largeur des flèches est proportionnelle au flux représenté.



UNIVERSITÉ
DE MONTPELLIER

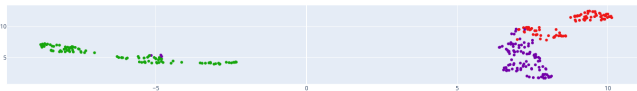
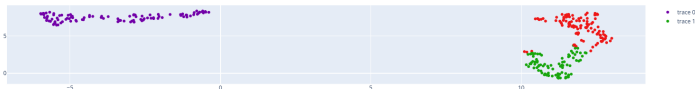
Application web

Page web

Transformation d'un Jupyter notebook contenant les différentes visualisations et faisant le lien entre eux.

Color:

None	Red	Purple	Green	Blue	Pink	Yellow	Orange
Black	Aqua	Brown					



Application web

Page d'accueil

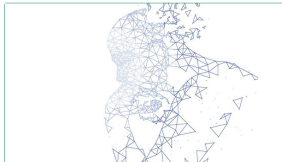
Création d'une page d'accueil personnalisée présentant nos différentes expérimentations.

MTQ Signature[Github](#)[À propos](#)

Ce site web présente les différents réseaux de neurones étudiés dans le cadre du projet *mtq-neural-networks*. L'objectif ici est de mieux comprendre comment s'exécute un réseau de neurones profonds. Il s'agit de repérer des signatures d'activation au sein des couches cachées du réseau en fonction des données d'entrées.

Afin de repérer les signatures, nous avons entraîné des modèles sur plusieurs partitions de données. Vous trouverez ci-dessous les liens qui présentent chacune des partitions.

Les liens sont de la forme *signature ABCD N* tel que *A*, *B*, *C* et *D* sont les chiffres sur lesquels s'est entraîné le modèle et *N* le nombre d'exemplaires de chacun de ces chiffres.



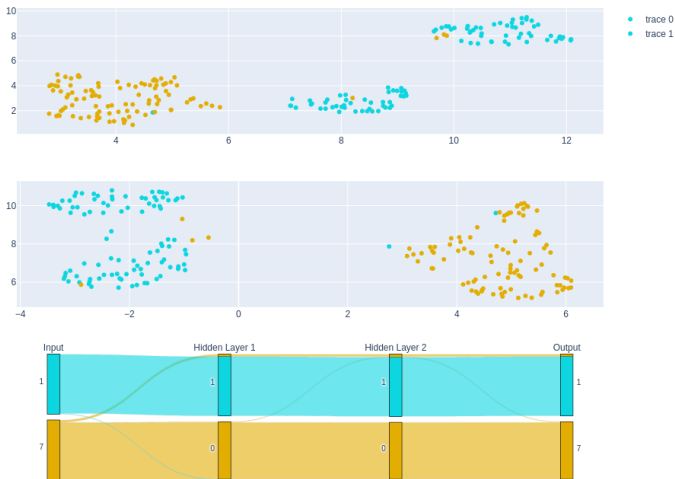
Choisir un modèle à ouvrir:

[signature_01_100.ipynb](#)[signature_137_100.ipynb](#)[signature_14_100.ipynb](#)[signature_28_100.ipynb](#)[signature_56_100.ipynb](#)[signature_69_100.ipynb](#)[signature_all_100.ipynb](#)

Ce site web a été réalisé par l'équipe [MTQ](#). Plus de détails sur le projet sont disponibles sur [Github](#).

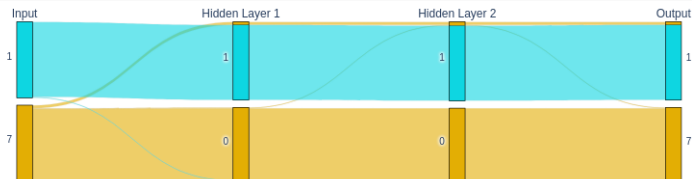
- ① Introduction
- ② Organisation
- ③ Analyse des données
- ④ Développement de l'architecture
- ⑤ Analyse des résultats**
Réponses aux questions
- ⑥ Conclusion

Résultats



Changement de comportement

Notre modèle arrive, dès la première couche cachée, à reconnaître une image.



Différence de signatures

On observe que les signatures des 1 sont majoritairement différentes de celles des 7. Sauf pour quelques rares exceptions.



UNIVERSITÉ
DE MONTPELLIER

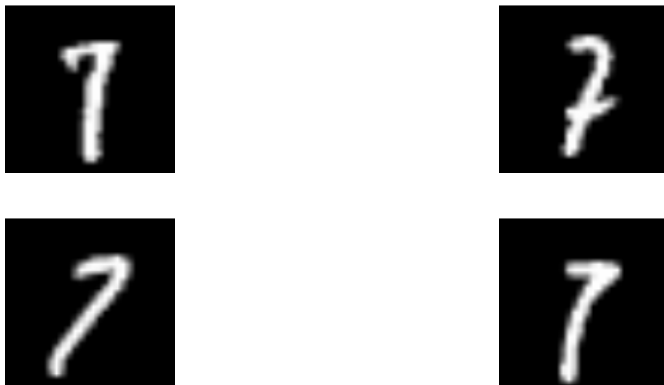


Figure – Les images de 7 ressemblant à des 1



UNIVERSITÉ
DE MONTPELLIER

Insertion d'anomalies

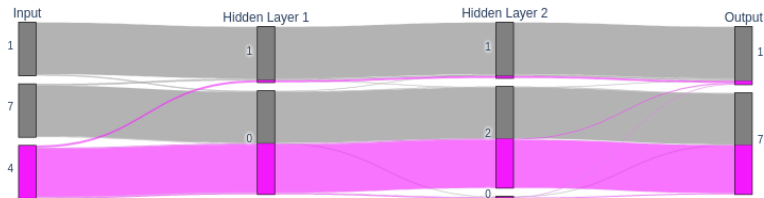


Figure – Insertion d'images de 4



UNIVERSITÉ
DE MONTPELLIER

Insertion d'anomalies

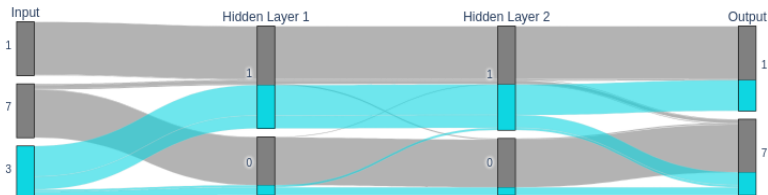


Figure – Insertion d'images de 3



UNIVERSITÉ
DE MONTPELLIER

- ① Introduction
- ② Organisation
- ③ Analyse des données
- ④ Développement de l'architecture
- ⑤ Analyse des résultats
- ⑥ **Conclusion**

Conclusion

- ▶ Un outil de visualisation ;
- ▶ Apports du projet ;
- ▶ Perspective.



UNIVERSITÉ
DE MONTPELLIER

Merci pour votre attention.



UNIVERSITÉ
DE MONTPELLIER