

Ilyes Boukaffal  
Pedro Alves  
L1-INT4  
2023-2024



# PYCHATBOT

Boukaffal-Alves-int4



# TABLE DES MATIÈRES :

- INTRODUCTION
- PRÉSENTATION FONCTIONNELLE DU PROJET
- PRÉSENTATION TECHNIQUE DU PROJET
- PRESENTATION DES RESULTATS
- CONCLUSION

## INTRODUCTION :

Ce projet python a pour but de créer un chatbot, qui lorsque qu'on lui donne à disposition des textes, doit être capable d'analyser les textes, et de pouvoir nous renvoyer des réponses, soit par des questions déjà enregistrer dans le code, soit par des questions que l'utilisateur pose et que le chatbot doit à leurs tours les analyser.

Les principaux objectifs de ce projet sont de nous aider à approfondir notre utilisation et nos compétences dans le langage python, de résoudre des problèmes, en apportant les solutions les plus claires et efficaces possible. Ce projet nous a permis d'approfondir nos compétences dans l'utilisation des fonctions, des listes, des tableaux, ainsi que des dictionnaires. Ce projet a aussi pour but de nous pousser à rechercher des informations sur nos problèmes, que ce soit dans des livres, des cours, des blogs etc...

## PRÉSENTATION FONCTIONNELLE DU PROJET :

- pouvoir lui poser des questions déjà enregistré dans le code.
- Nous pouvons lui indiquer la question que l'on souhaite lui poser grâce à l'interface de la console python.

- La deuxième fonctionnalité: poser n'importe quelle question au chatbot, et il nous renvoie le fichier qui correspond le mieux à notre question, ainsi que la phrase qui correspond le mieux à votre question dans ce fichier.
- Malheureusement, cette fonctionnalité n'a pas été aboutie, suite à de nombreux problèmes survenus lors du développement de cette partie.

## PRÉSENTATION TECHNIQUE DU PROJET :

Le programme réalise les différentes fonctionnalités proposées grâce à :

- un calcul de matrice TF-IDF, qui permet de calculer entre autres, la fréquence d'un mot par rapport à la fréquence de document donné.
- Un calcul de similarité vectorielle, qui permet à l'utilisateur de pouvoir poser la question qu'il souhaite au chatbot.
- Le choix de l'utilisation des dictionnaires dans notre code est important, car il nous permet de mieux se rapprocher de l'architecture d'une matrice, pour pouvoir ainsi utiliser pleinement le fonctionnement de notre calcul TF-IDF.
- Le plus gros problème technique était l'utilisation de GitHub car il y avait de nombreux conflits de fichiers entre les participants.
- Le deuxième problème est l'affichage des caractères spéciaux dans les matrices, qui n'est malheureusement pas réglé ainsi qu'un bug d'affichage de la réponse pour la question 2.

## PRÉSENTATION DES RÉSULTATS :

- Interface utilisateur :

```
--- Menu ---

- To display the least important words, enter 1
- To display the most important words, enter 2
- To display the most important word by Chirac, enter 3
- To display the presidents who talk about the Nation, as well as the one who talks about it the most, enter 4
- To display the president who speaks most about ecology, enter 5
- Sorry, this command is not available at the moment : (To access the bot, enter 0.)

Enter a number between 0 and 5 :
```

- Affichage des matrices :

['efforcerais', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['exprimés', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['fondée', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['fortement', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['raels', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['durant', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['irraçprochable', 0, 0, 0, 0, 0, 0, 0, 0]	['achever', 0, 0, 0, 0, 0, 0, 0, 0]
['battrai', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['confiés', 0, 0, 0, 0, 0, 0, 0, 1.8061799739838869]
['méditerranée', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['dàcevoir', 0, 0, 0, 0, 0, 0, 0, 0]
['afrique', 0, 0, 0, 0, 0, 0, 0, 0]	['lorsqu', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['lutte', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['auxquels', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['raçchauffement', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['confrontés', 0, 0, 0, 0, 0, 0, 0, 0]
['diplomatie', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['respecter', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['inscrire', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['donnés', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['conviction', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['tenir', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['camp', 0, 0, 0, 0, 0, 0, 0, 0]	['àbranlé', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['bonnes', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['fragile', 0, 0, 0, 0, 0, 0, 0, 0]
['volontés', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['morale', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['aient', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['fort', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['animés', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['raçhabilitier', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['passion', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['tolérance', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['veulent', 0, 0, 0, 0, 0, 0, 0, 1.8061799739838869]	['intolérance', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['prât', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['sectarisme', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]
['travailler', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]	['destructeurs', 0, 0, 0, 0, 0, 0, 0, 0.9030899869919434]

- Affichage des réponses pour la question 1 :

```
Enter a number between 0 and 5 : 1

- Voici les mots les moins importants

messieurs
les
mesdames
en
ce
je
la
de
le
une
a
france
qui
se
et
dans
peuple
```

- Affichage des réponses à la question 2 :

```
Enter a number between 0 and 5 : 2  
  
- Voici le mot le plus important du dossier :
```

- Affichage des réponses à la question 3 :

```
Enter a number between 0 and 5 : 3  
  
- Voici le mot le plus important des discours de Chirac : doit
```

- Affichage des réponses à la question 4 :

```
Enter a number between 0 and 5 : 4  
  
- Voici les présidents parlant de Nation  
  
Chirac  
Hollande  
Macron  
Mitterrand  
  
- Mitterrand parle le plus de Nation
```

- Affichage des réponses à la question 5 :

```
Enter a number between 0 and 5 : 5  
  
- Macron est le président qui parle le plus du climat
```

## CONCLUSION :

Ainsi, sur le soft skills, ce projet nous aura permis de découvrir le travail de groupe, de savoir s'organiser et de répartir les tâches, bien gérer son temps entre les différents projet académiques, ainsi que de savoir rechercher des bonnes informations, là où il faut : sur internet, dans des livres, des cours...

Sur le plan des hard skills, ce projet nous a permis de renforcer nos compétences en python tel que sur l'utilisation des fonctions, des listes 1D, 2D, ainsi que des dictionnaires. Il nous a aussi permis d'approfondir le paradigme procédural car au lycée, nous nous focalisons plus sur la programmation orientée objet.