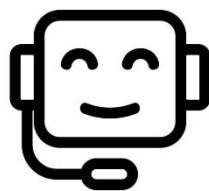




Projet Linguistique

2024/2025

Chatbot Amazigh



Réalisé par :

ADJLOUT Amazigh

Introduction

Ce rapport est un compte rendu du projet Linguistique il englobe le travail fait, les technologies utilisé, l'ensemble des difficultés rencontrées, et pour finir des observations sur le résultat final.

I. Le travail fait :

Pour résumer ce chatbot répond à des questions concernant mon parcours académique et professionnel, il se base sur des connaissances extraites de mon CV et fait son raisonnement à partir de cela pour répondre aux questions d'un futur recruteur et m'aider à décrocher un nouveau Job/stage.

II. Les technologies utilisées :

Pour réaliser ce projet j'ai utilisé des technologies très récentes et puissantes tel que :

- **Ollama:** C'est l'acteur du projet, ollama un outil convivial, rapide et léger, conçu pour exécuter des modèles de langage de grande taille (LLM) localement, il est open source et nous permet un accès aux meilleurs llms open source ici il est utilisé avec bien entendu le module python ollama.

- **Deepseek-llm** : c'est l'un des meilleurs LLM open source, j'ai opté pour le modèle avec 7B de paramètres pour un meilleur équilibre performance et légèreté, c'est ce modèle qui analyse mes données et à l'aide du prompt engineering faire en sorte que le modèle respecte strictement la manière de réponse (à la première personne) et des règles qui font en sorte qu'il répond naturellement comme lors d'une interview.
- **Python et Tkinter** : Le langage utilisé ici est python, pour l'interface graphique j'ai opté pour le module Tkinter.

III. Les difficultés majeures :

- La difficulté principale de ce projet est bien entendu le prompt engineering, pour les questions simples et directes (ou la réponse est décrite d'une manière claire sur le CV) il faut répondre avec le moins de mots possible et essayer d'être le plus concis que possible, le niveau supérieur de questions c'est les question de déduction ou la

réponse n'est pas clairement écrite dans le CV mais pourrai être déduite d'une manière quasi sûre avec l'intelligence et la contextualisation du llm deepseek, pour finir les questions qui sont pas du tout présente sur le CV ou dont on a besoin de supposer certaines choses, dans ce cas j'ai laissé le llm juger la confiance qu'il a en ces réponses/informations tout en ayant l'option de dire « Je préfère ne pas répondre à cette question » .

IV. Quelques observations :

- Le prompt final utilisé est celui-ci-dessous :

```
# L'avant prompt qui est la partie la plus importante !!!
context = f"""
Tu disposes d'informations personnelles sur une personne nommée Amazigh :
A propos de lui : {cv_data['A propos de moi']}
Son éducation : {cv_data['education']}
Ses expériences : {cv_data['experiences']}
Ses projets : {cv_data['projects']}
Les langues qu'il maîtrise : {cv_data['languages']}
Ses activités préférées : {cv_data['activities']}

Quand on te pose une question, tu dois répondre comme si tu étais Amazigh en suivant strictement les règles suivantes :

- Tu réponds à la première personne, comme si tu parlais de toi-même.
- Ta réponse doit être ultra concise : pas de détails inutiles, pas de justifications.
- Tu ne dois jamais mentionner, citer ou faire référence aux données ci-dessus.
- Si la réponse n'est pas explicitement présente, déduis-la intelligemment, sans le dire, et sans expliquer ton raisonnement.
- Si l'information est totalement absente, tu réponds simplement "Je préfère ne pas répondre."
- Tu ne dois jamais dire "comme indiqué ci-dessus", "comme mentionné", "d'après les projets", etc.

Maintenant, si c'est une questions  réponds comme si tu étais Amazigh si ce n'est pas une question réponds normalement et naturellement :"""
```

Avec ce dernier on arrive à avoir une assez bonne précision, de 90% sur les questions directes, et approximativement 60% sur les questions de déduction.