

Créer un modèle

La commande `ollama create` est utilisée pour créer un modèle à partir d'un fichier **Modelfile**.

```
ollama create mymodel -f ./Modelfile
```

Cette commande crée un modèle nommé **mymodel** en utilisant les instructions définies dans le fichier `Modelfile`.

Télécharger un modèle :

Pour télécharger un modèle, utilisez la commande suivante :

```
ollama pull llama3.2
```

Cette commande télécharge le modèle **llama3.2** depuis la bibliothèque Ollama. Elle peut également être utilisée pour **mettre à jour un modèle local**. Seules les différences (diff) entre le modèle local et la version la plus récente seront téléchargées.

Supprimer un modèle :

Si vous souhaitez supprimer un modèle, vous pouvez utiliser la commande suivante :

```
ollama rm llama3.2
```

Cela supprimera le modèle **llama3.2** de votre machine.

Copier un modèle :

Pour copier un modèle, utilisez la commande :

```
ollama cp llama3.2 my-model
```

Cela crée une copie du modèle **llama3.2** sous le nom **my-model**.

Entrée multilignes :

Pour entrer plusieurs lignes de texte, vous pouvez entourer le texte avec des guillemets triples (`"""`), comme suit :

```
>>> """Hello, ... world! ... """
```

Cela permet de saisir un texte réparti sur plusieurs lignes. Par exemple, cette entrée pourrait produire la sortie suivante :

```
I'm a basic program that prints the famous "Hello, world!" message to the console.
```

Modèles multimodaux :

Les modèles multimodaux permettent d'interagir avec des fichiers autres que du texte, comme des images. Par exemple, pour analyser une image avec un modèle multimodal, utilisez la commande :

```
ollama run llava "What's in this image? /Users/jmorgan/Desktop/smile.png"
```

Cela pourrait donner la réponse suivante :

```
The image features a yellow smiley face, which is likely the central focus of the picture.
```

Passer le prompt en argument :

Vous pouvez aussi passer un prompt directement en argument à la commande `ollama run`. Par exemple :

```
$ ollama run llama3.2 "Summarize this file: $(cat README.md)"
```

Cela permet de résumer le contenu d'un fichier, comme le fichier `README.md`. Le modèle peut générer un résumé tel que :

```
Ollama is a lightweight, extensible framework for building and running language models on the local machine. It provides a simple API for creating, running, and managing models, as well as a library of pre-built models that can be easily used in a variety of applications.
```

Afficher les informations sur un modèle :

Pour afficher les informations détaillées d'un modèle, utilisez la commande suivante :

```
ollama show llama3.2
```

Cela affiche des informations sur le modèle **llama3.2**, telles que sa version et ses paramètres.

Lister les modèles sur votre ordinateur :

Pour voir la liste de tous les modèles installés sur votre machine, utilisez cette commande :

```
ollama list
```

Lister les modèles actuellement chargés :

Pour voir quels modèles sont actuellement chargés en mémoire, utilisez la commande :

```
ollama ps
```

Arrêter un modèle en cours d'exécution :

Si vous souhaitez arrêter un modèle qui est en cours d'exécution, utilisez la commande suivante :

```
ollama stop llama3.2
```

Cela arrêtera le modèle **llama3.2** en cours d'exécution.

Démarrer Ollama :

Si vous voulez démarrer Ollama sans utiliser l'application de bureau, vous pouvez utiliser la commande suivante :

```
ollama serve
```

Cela démarre Ollama en mode serveur, permettant ainsi d'interagir avec les modèles via une API sans l'interface graphique.

Revision #3

Created 27 November 2024 13:05:48 by Maxime

Updated 3 December 2024 14:06:29 by Maxime