

Projet Big Data Machine learning

December 22, 2022

Pour modéliser la prédiction du jeu de données “data” avec Gradio et MongoDB en Python, vous pouvez suivre les étapes suivantes :

1. Installez les bibliothèques nécessaires : Tout d’abord, assurez-vous d’avoir installé Gradio et pymongo. Vous pouvez les installer en utilisant ” pip install gradio et pip install pymongo”, respectivement.

2. Connectez-vous à la base de données MongoDB : Ensuite, vous aurez besoin de vous connecter à votre base de données MongoDB en utilisant la bibliothèque pymongo. Vous pouvez le faire en créant un objet MongoClient et en lui passant le nom d’hôte et le numéro de port de votre serveur MongoDB.

import des collections avec MongoDB

```
mongoimport -db data -collection train -type csv --headerline --file  
"/Users/mac/Desktop/Projet/data/train.csv"
```

```
mongoimport -db data -collection test -type csv --headerline --file  
"/Users/mac/Desktop/Projet/data/test.csv"
```

3. Chargez le jeu de données “data” : Une fois que vous êtes connecté à la base de données, vous pouvez charger le jeu de données “data” en utilisant la méthode find() de la bibliothèque pymongo. Cette méthode renvoie un objet curseur, que vous pouvez parcourir pour récupérer les documents de la collection.

4. Préparer les données : Avant de pouvoir commencer à modéliser, vous devrez prétraiter les données pour les préparer à l’entrée de votre modèle. Cela peut inclure des tâches telles que la gestion des valeurs manquantes, l’encodage des variables catégorielles et l’échelle des variables numériques.

5. Entraînez un modèle : Une fois que vous avez prétraité les données, vous pouvez entraîner un modèle en utilisant n’importe quelle bibliothèque d’apprentissage automatique, comme scikit-learn ou pytorch Vous pouvez utiliser les données d’entraînement pour ajuster le modèle et l’utiliser pour faire des prédictions sur les données de test.

6. Utilisez Gradio pour créer une interface : Enfin, vous pouvez utiliser Gradio pour créer une interface interactive pour votre modèle. Cela vous permettra de saisir des données et de recevoir des prédictions de votre modèle en temps réel

Voici un exemple de code qui montre comment mettre en œuvre ces étapes :

```
[ ]: import numpy as np
import pandas as pd

import gradio as gr
from pymongo import MongoClient

# Connect to the MongoDB database
client = MongoClient(host='localhost', port=27017)
db = client['data']
collection_train = db['train']
collection_test = db['test']

# Load the dataset
data_train = list(collection_train.find())
data_test = list(collection_test.find())

# Preprocess the data
# ...

# Train a model
# ...

# Use Gradio to create an interface
def predict(inputs):
    # Use the trained model to make predictions on the input data
    prediction = model.predict(inputs)
    return prediction

# Create the interface
interface = gr.Interface(fn=predict, inputs=inputs, outputs=outputs)

# Launch the interface
interface.launch()
```

7 liste des projets :