

TECHWOMENFACTORY CHAMPION

Description du projet

Parcours : Data Science(DS)

Objectif du projet :

Notre projet a pour but de diagnostiquer 06 maladies dermatologiques complexes avec quasi-exactitude et en peu de temps. Notre réalisation s'intitule « Diseases Predict ».

Membres du Groupe :

- Djuigouing Simo Daina 20%
- Heude Dilane 20%
- Mafogang Borelle 20%
- Toukep Tchamba Minette 20%
- Minka Loyd Marcel 20%

Les technologies utilisées sont :

- **Microsoft Excel**, pour le traitement des valeurs manquantes (cours de fondement d'analyse de données)
- **Jupyter Notebook**, pour l'analyse de données et la construction du modèle.
- **IDLE de python** (éditeur de texte python), pour produire l'application.
- **Visual Studio Code** (éditeur de texte également), pour produire l'application
- **MYSQL**, pour la création de la base donnée de patients

Les outils utilisés :

- **Le module pandas**, utilisé pour la manipulation du jeu de données (Cours de « Framework »)
- **Le module matplotlib**, pour la visualisation des données (Cours de « Framework »).

- **Le module sklearn**, pour la construction et l'évaluation des modèles, **la metric d'évaluation de la performance du modèle utilisée est la F-mesure.** (Cours Evaluation des modèles)
- **Le Framework Streamlit**, pour la conception de l'application (front-end et back-end).
- **Le module mysql-connector-python**, pour la connexion de la base de données à l'application afin d'effectuer les requêtes (cours Requetage des bases de données relationnelles).
- **Le module pickle**, pour pouvoir sauvegarder le modèle afin de l'intégrer plus tard dans l'application.

Les techniques utilisées sont :

- La classification supervisée (cours Modèles de fouille de données)
- La programmation procédurale (cours Programmation procédurale en python)
- Le Hold-out, utilisé pour diviser le jeu de données (cours Evaluation des modèles).

Outils à installer :

1) MYSQL

Télécharger et installer Mysql workbench et Mysql community. Une fois installer, importer la base de données « derma_predict » dans mysql.

2) PYTHON (python 3.9)

Télécharger et installer python (pendant l'installation, cocher la deuxième case pour ajouter au « path »)

3) LES DIFFERENTS MODULES

→ Ouvrir l'invite de commande de Windows

→ Mettre à jour **pip** à travers la commande :

```
python -m install pip --upgrade pip
```

→ Installer **streamlit** à travers la commande :

```
pip install streamlit
```

→ Installer **openpyxl** à travers la commande :

Pip install openpyxl

→ Installer **sklearn** à travers la commande :

pip install sklearn

→ Installer **mysql-connector-python** à travers la commande :

pip install mysql-connector-python

4) Télécharger et installer anaconda (dans lequel se trouve le logiciel jupyter pour la lecture du notebook « twf_champs »).

Comment démarrer l'application ?

→ Ouvrir l'invite de commande de Windows

→ Ouvrir le dossier comp dans l'invite de commande

→ Saisir la commande « streamlit run multi.py »

→ Taper sur entrer