



A detailed satellite with two large solar panel arrays is shown in orbit above the Earth. The satellite has a central cylindrical body with various instruments and antennas. The Earth's surface is visible below, showing clouds and landmasses. The background is the blackness of space with distant stars.

Python for data analysis

Léa Youssef

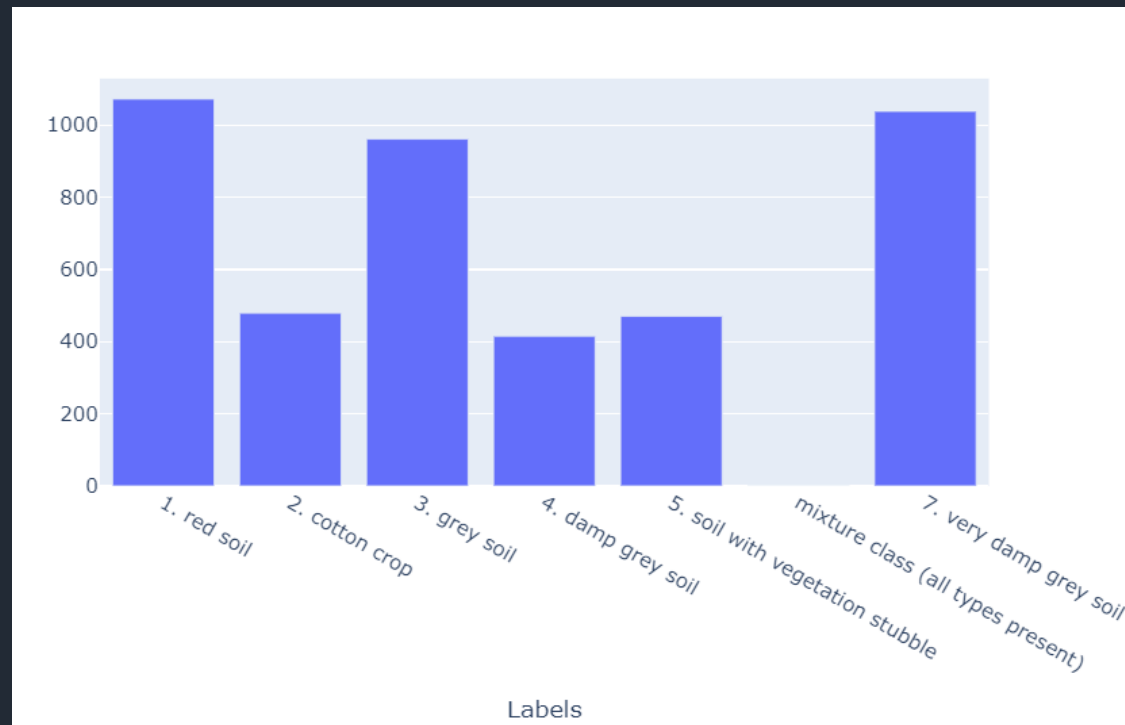
Nicolas Thomazo



I. Explication du dataset

Notre projet consiste à trouver le type de sol qui a été pris en photo par le satellite landsat. Vous pouvez voir ici les différents de sol du dataset. Nous avons seulement une partie de l'image prise par le satellite (82x100 pixels)

Classes à prédire

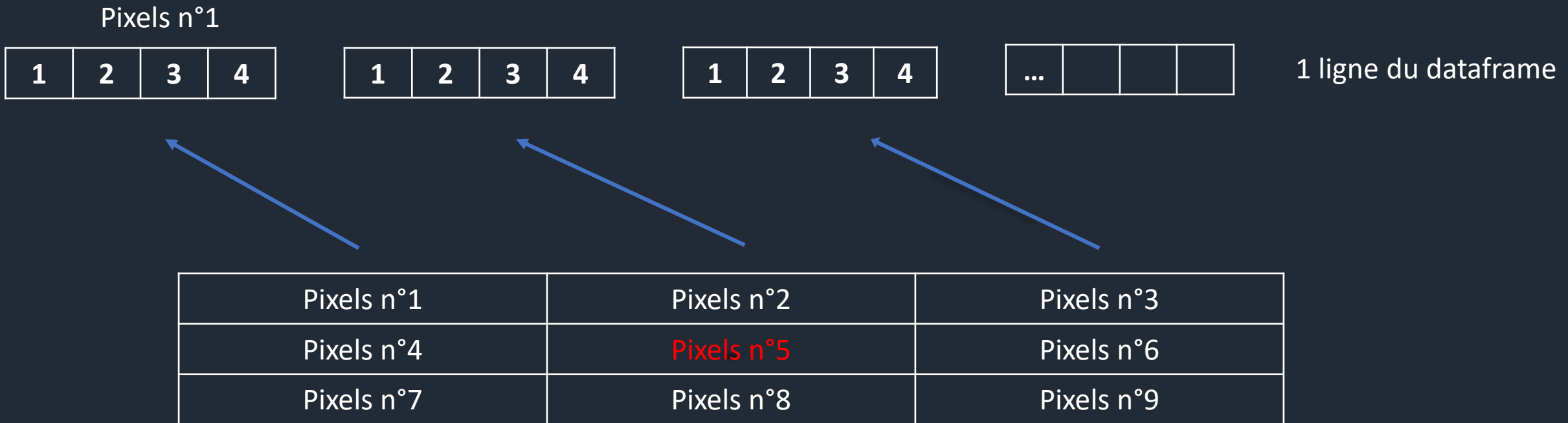




I. Explication du dataset

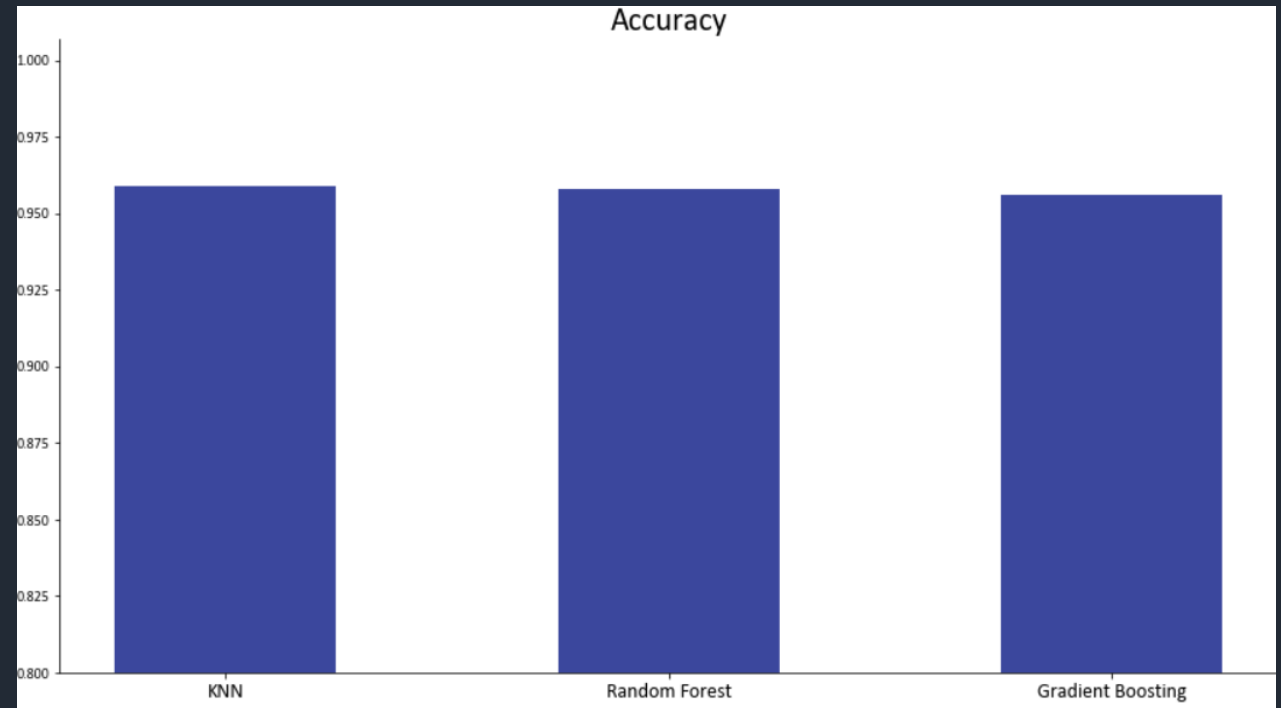
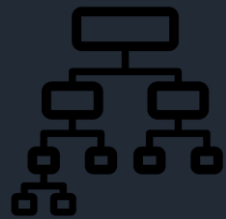
Les pixels ont été mis à la ligne les uns à la suite des autres. Chaque ligne du dataframe représente 9 pixels et le but est de prédire celui du milieu (le rouge).

Chaque image prise par le satellite consiste en 4 images prises dans des valeurs spectrales différentes. C'est pour cela que chaque case du tableau correspond à 4 valeurs de pixels et donc on a 36 valeurs de pixels dans chaque ligne.



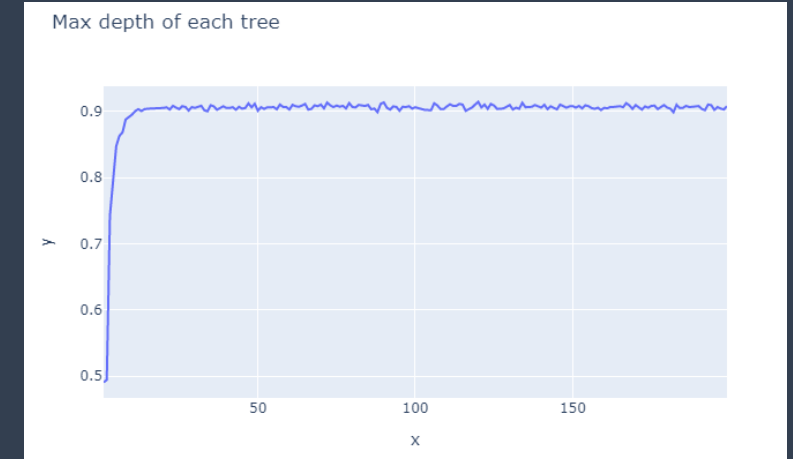
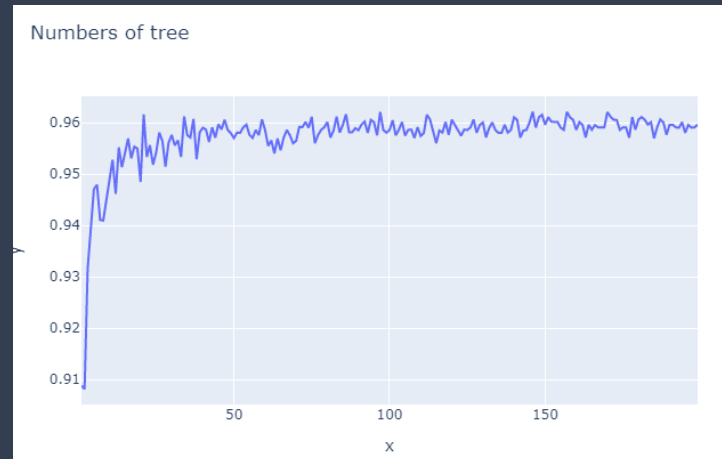
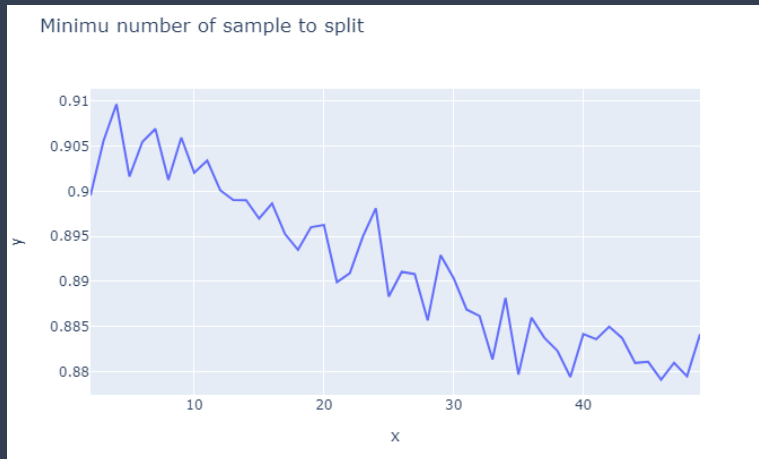
II. Modèles utilisés

- k-nearest neighbors
- Random Forest
- Gradient Bosting



II. Modèles utilisés

Paramètres testés sur Random Forest



II. Modèles utilisés

Pour chacun des modèles on a obtenu un score f1 d'environ 0,96. Nous avons trouvé ce résultat assez étonnant pour 2 raisons.

- Un score f1 de 0,96 sur le test d'évaluation est très élevé pour un modèle de machine learning
- Les modèles ont quasiment tous prédit avec la même précision

On peut donc penser que la relation entre les pixels et leur classes est assez évidente, c'est pour cela que les modèles prédisent tous très bien.



III. API Flask

- Rentrer la valeur des 4 pixels dont on veut prédire la classe
- Résultat de la classe prédite, ainsi que de sa probabilité
- Modèle utilisé : KNN

Prédiction de la classe d'un pixel

Predict

Le pixel correspond à grey soil (classe 3)
avec un probabilité de 0.6