# Projet 8 : Déployez un modèle dans le cloud

**Présenté par** : Bourama FANE Etudiant Data Scientist **Dirigé par** : Babou M'BAYE Mentor chez OpenClassrooms

08 Janvier 2024



#### Sommaire

- 1 Contexte & Problématique
- 2 Choix techniques généraux retenus
- Opploiement local
- 4 Déploiement Cloud
- Conclusion



## Plan de la présentation

- Contexte & Problématique
- Choix techniques généraux retenus
- 3 Deploiement local
- 4 Déploiement Cloud
- Conclusion



Vous êtes **Data Scientist** dans une très jeune **start-up de l'AgriTech**, nommée "**Fruits!**", qui cherche à proposer des solutions innovantes pour la récolte des fruits.

La volonté de l'entreprise est de préserver la biodiversité des fruits en permettant des traitements spécifiques pour chaque espèce de fruits en développant des robots cueilleurs intelligents.





Bourama FANE Soutenance P8 08 Janvier 2024 4 /

Votre start-up souhaite dans un premier temps se faire connaître en mettant à disposition du grand public une application mobile qui permettrait aux utilisateurs de prendre en photo un fruit et d'obtenir des informations sur ce fruit.



Pour la start-up, cette application permettrait de sensibiliser le grand public à la biodiversité des fruits et de mettre en place une première version du moteur de classification des images de fruits.



Bourama FANE Soutenance P8 08 Janvier 2024 5 / :

Développer une première chaîne de traitement des données qui comprendra le preprocessing et une étape de réduction de dimension.



- Tenir compte du fait que le volume de données va augmenter très rapidement après la livraison de ce projet, ce qui implique de :
  - ✓ Déployer le traitement des données dans un environnement Big Data
  - ✓ Développer les scripts en pyspark pour effectuer du calcul distribué



Bourama FANE Soutenance P8 08 Janvier 2024 6 /

### Différentes étapes du projet

- ✓ Liste des choix techniques généraux retenus
- ✓ Déploiement de la solution en local
- ✓ Déploiement de la solution dans le cloud





#### Données

- "Fruit 360" contient 90 483 images
  - ✓ Training: 67 692 images
  - ✓ Test : 22 688 images



Grape B

Grane W

- Répartition des images en 131 dossiers
  - ✓ 1 dossier = 1 fruit
  - ✓ Photos prises sous différents angles
  - ✓ Images de 100X100 pixels au format : JPG, RGB



# Plan de la présentation

- 1 Contexte & Problématique
- 2 Choix techniques généraux retenus
- 3 Deploiement local
- 4 Déploiement Cloud
- Conclusion



# Choix techniques généraux retenus

#### Calcul distribué

- Dans les deux environnements (Local et Cloud) nous utiliserons donc Spark et nous l'exploiterons à travers des scripts python grâce à **PySpark**.
- Dans la version cloud nous réaliserons les opérations sur un cluster de machine.

#### Transfert Learning

- Nous allons utilisons un MobileNetV2;
- L'avant dernière couche correspond à un vecteur réduit de dimension (1,1,1280);
- Rapidité d'exécution.



#### Plan de la présentation

- 1 Contexte & Problématique
- 2 Choix techniques généraux retenus
- 3 Deploiement local
- 4 Déploiement Cloud
- Conclusion



#### Deploiement local

- ✓ Préparer nos données
  - Importer les images dans un dataframe pandas UDF
  - Associer aux images leur label
  - Préprocessing
- ✓ Préparer notre modèle
  - Importer le modèle MobileNetV2;
  - Créer un nouveau modèle dépourvu de la dernière couche de **MobileNetV2** :
- ✓ Exécuter les actions d'extraction de features ;
- ✓ Enregistrer le résultat de nos actions;
- ✓ Tester le bon fonctionnement en chargeant les données enregistrées.



## Plan de la présentation

- 1 Contexte & Problématique
- 2 Choix techniques généraux retenus
- 3 Deploiement local
- 4 Déploiement Cloud
- Conclusion



## Qu'est ce que le BigData

## BigData

- « 3 caractéristiques »sont associées au Big Data nommés les « **3V** » :
  - **Volume** : Le Big Data implique des quantités massives de données. Datas ne tiennent pas en Ram (>million de gigaoctets)
  - Vélocité : Les Data sont générées et modifiées très rapidement.
  - Variété: Data hétérogènes en types et formats. (données structurées, données semi-structurées (comme le format JSON) et données non structurées (texte, vidéos, images, etc.).

#### Pourquoi le cloud

- Coût (Pay as You Go)
- Evite les problème lié à l'infrastructure et à sa maintenance
- Facilité pour passer à l'échelle



Bourama FANE Soutenance P8 08 Janvier 2024 14/2

#### Prestataire cloud: AWS

L'objectif premier est de pouvoir, grâce à AWS, louer de la puissance de calcul à la demande. L'idée étant de pouvoir, quel que soit la charge de travail, obtenir suffisamment de puissance de calcul pour pouvoir traiter nos images, même si le volume de données venait à fortement augmenter.

De plus, la capacité d'utiliser cette puissance de calcul à la demande permet de diminuer drastiquement les coûts si l'on compare les coûts d'une location de serveur complet sur une durée fixe (1 mois, 1 année par exemple).



Bourama FANE Soutenance P8 08 Janvier 2024 15 / 27

## Solution & Stockage

#### Service EMR

✓ la solution PAAS en choisissant d'utiliser le service EMR d'Amazon Web Services.

#### EC<sub>2</sub>

- X Espace limité;
- X Risque de saturation (ralentissements, disfonctionnements);
- X Pertes de données en cas de résiliation (Coûts);

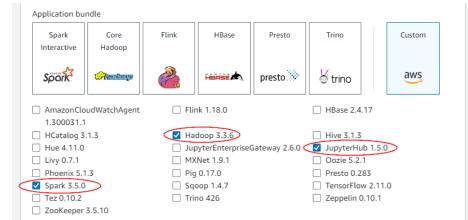
#### Amazon S3

✓ Utiliser Amazon S3 permet de s'affranchir de toutes ces problématiques. L'espace disque disponible est illimité, et il est indépendant de nos serveurs EC2. L'accès aux données est très rapide car nous restons dans l'environnement d'AWS



Bourama FANE Soutenance P8 08 Janvier 2024 16 / 2

# Configurations EMR : choix des logiciels à installer





4日 > 4周 > 4 3 > 4 3 >

# Configurations EMR : Persistance des notebooks

```
▼ Edit software settings – optional Info

    Enter configuration

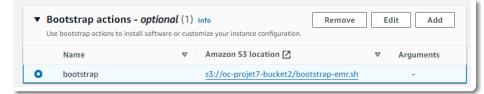
    Load JSON from Amazon S3

   2 ▼
           "classification": "jupyter-s3-conf",
           "properties": {
   4 ▼
              "s3.persistence.bucket": "oc-projet7-bucket",
   5
   6
              "s3.persistence.enabled": "true"
   7
   8
   9
```



Bourama FANE Soutenance P8 08 Janvier 2024 18 / 27

# Configurations EMR: Actions d'amorçage

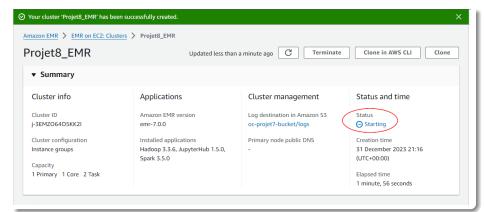


```
sudo python3 -m pip install -U setuptools pip
sudo python3 -m pip install wheel pillow pandas numpy pyarrow boto3 s3fs fsspec
sudo python3 -m pip install tensorflow
```



Bourama FANE Soutenance P8 08 Janvier 2024 19 / 27

### Configurations EMR : Instanciation du serveur





4 D F 4 B F 4 B F

Bourama FANE Soutenance P8 08 Janvier 2024 20 / 27

# Configurations EMR : Tunnel SSH

```
hadoop@ip-172-31-22-72:~
PS C:\Users\Fane0763\OpenClassroom\OC Projet 8> ssh -i ./Projet8KeyEC2Naby.pem -D 5555 hadoop@ec2-54-167-31-13.compute-1.amazonaws.com
Warning: Identity file ./Projet8KeyEC2Naby.pem not accessible: No such file or directory.
hadoop@ec2-54-167-31-13.compute-1.amazonaws.com: Permission denied (publickey,qssapi-keyex,qssapi-with-mic).
PS C:\Users\Fane0763\OpenClassroom\OC Projet 8> ssh -i ./Projet8keyNaby.pem -D 5555 hadoop@ec2-54-167-31-13.compute-1.amazonaws.com
A newer release of "Amazon Linux" is available.
  Version 2823.3.28231218:
Run "/usr/bin/dnf check-release-update" for full release and version update info
                    Amazon Linux 2023
     \ #####\
                    https://aws.amazon.com/linux/amazon-linux-2023
 ast login: Tue Jan 2 88:46:81 2824
EEEEEEEEEEEEEEEEEE MMMMMMM
                                       E:::::EEEEEEEEE:::E M:::::::M
                                     M:::::::: M R:::::RRRRRR:::::R
 E::::E
  E:::::EEEEEEEEE M:::::M M::::M M::::M M:::::M
 Economic Monocom
                    M:::::M
                                                 R:::R
                                                            R::::R
              FFFFF M::::M
                                        Marata M
                                                            R . . . . R
 E:::::EEEEEEEE::::E M:::::M
                                                            R::::R
                                        M:::::M RR::::R
                                                            R . . . . R
EEEEEEEEEEEEEEEEEE MMMMMM
                                        мимимим рарарара
                                                            PPPPPP
[hadoop@ip-172-31-22-72 ~]$
```

Bourama FANE Soutenance P8 08 Janvier 2024 21 / 27

# Configurations EMR : SwitchyOmega





4日 > 4 周 > 4 目 > 4 目

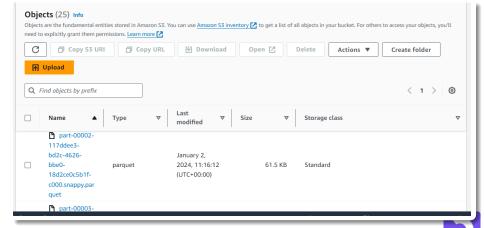
# Configurations EMR : Connexion à Jupyter

de Enable an SSH connection
UI URL [7]
☐ <u>http://ec2-54-167-31-13.compute-1.amazonaws.com:9870/</u>
₱ https://ec2-54-167-31-13.compute-1.amazonaws.com:9443/
☐ http://ec2-54-167-31-13.compute-1.amazonaws.com:8088/
☐ http://ec2-54-167-31-13.compute-1.amazonaws.com:18080/
sk nodes
UI URL
http://ec2-000-000-000-000.compute-1.amazonaws.com:9864/



Bourama FANE Soutenance P8 08 Janvier 2024

## On peut également constater la présence des fichiers



24 / 27

## Plan de la présentation

- 1 Contexte & Problématique
- 2 Choix techniques généraux retenus
- 3 Deploiement local
- 4 Déploiement Cloud
- Conclusion





- La solution a parfaitement fonctionné en mode local.
- La deuxième phase a consisté à créer un réel cluster de calculs (AWS & EMR).
- Amazon S3 pour stocker les données de notre projet
- Il nous sera facile de faire face à une monté de la charge de travail en redimensionnant simplement notre cluster de machines.





# MERCI POUR VOTRE AIMABLE ATTENTION

