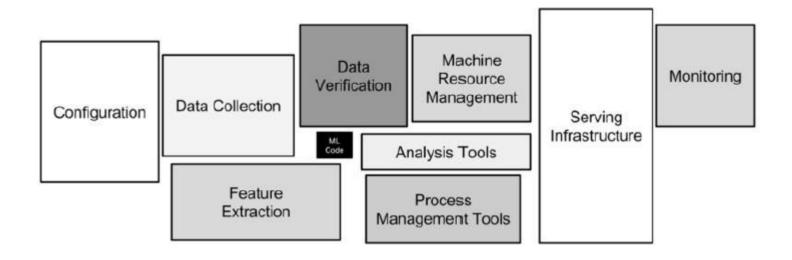
Introduction Déploiement de modèles en production

Agenda

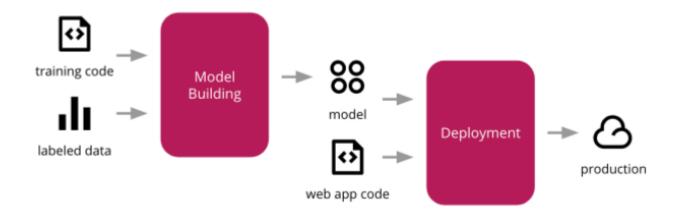
- ☐ Dans cette leçon, on passe en revue les points suivants:
 - ➤ Déploiement en production
 - > Cycle de vie

Dette technique cachée dans les systèmes ML

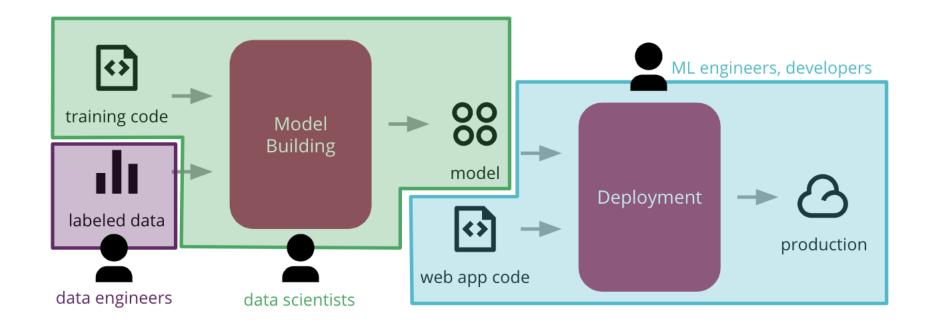


Processus

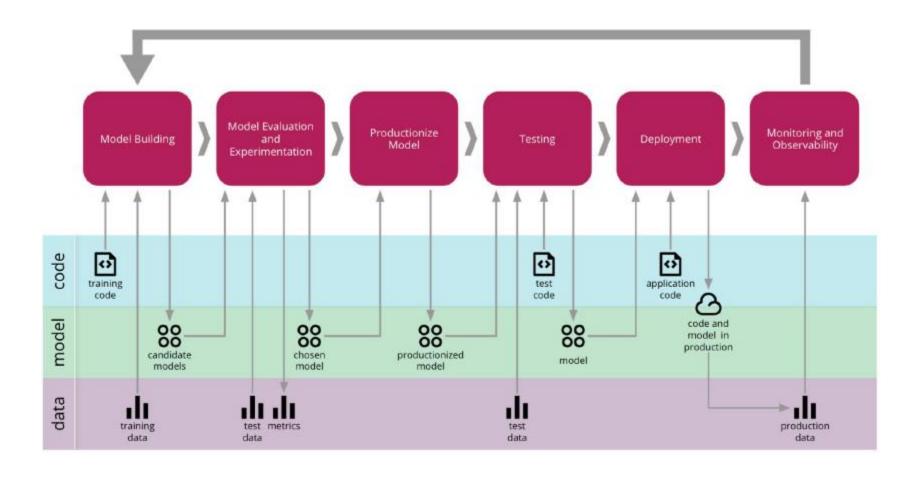
➤ Processus initial pour former notre modèle ML, l'intégrer à une application Web et le déployer en production



Équipe-Rôles



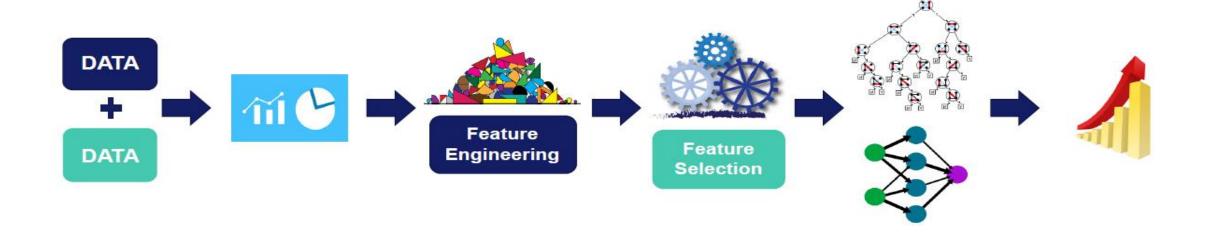
Cycle de vie d'un système ML



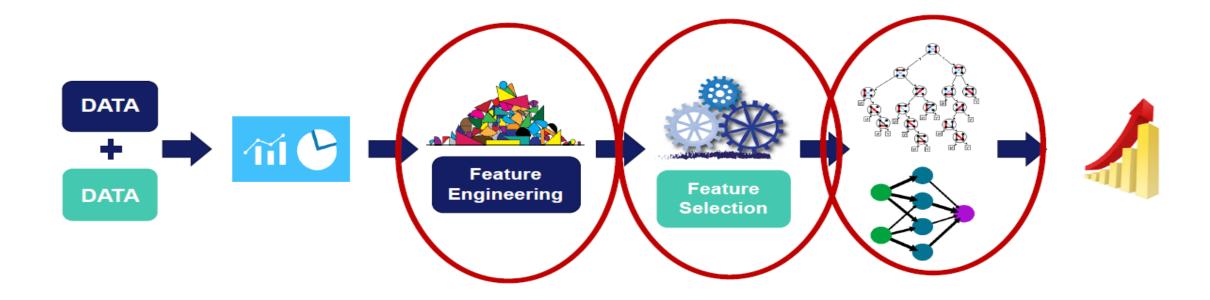
Modèle prédictif

- Les étapes sont les suivantes
 - Créer un modele d'apprentissage sur la base d'une approche train-test
 - Une fois l'étape d'apprentissage et de test est terminé, on prepare le déploiement
 - Dépend du modele
 - Effectuer le déploiement comme un service web
 - Peut être la solution idéale
 - Tenir compte d'autres approaches de déploiement.

Workflow-Pipeline



Mise en production- Quoi?



code de déploiement



```
File Edit Selection Find View Gato Table Project Profesences Hely Lacused
                                                                      (% extends "Layout.html" %)
    from flask import Flask, render template
                                                                      (# endblock #)
                                                                      (# block header #7
         | Following are the application routes which
                                                                      cheader class="jumbotron-fluid text-center">
          control all of the routing inside of the web
                                                                             chilo A Simple Quote coro
                                                                                conally A Personc/smally
         application.
                                                                         <div class="wrapper">
                                                                             class="breadcrumb-item">Home
        return render_template("index.html")
    @app.route("/portfolio")
                                                                      (# endblock #)
        return render_template("portfolio.html")
    @app.route("/projects")
                                                                             cmain class="col-md-8">
        return render template("projects.html")
                                                                                             ca href-"#" class-"label
                                                                                             label-primary">Category</
                                                                                          <hd class >Post Title</hd>
```

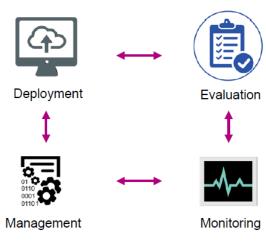
Code de déploiement

```
import pandas as pd
4 Import numpy as np
   from sklearn.model selection import train test split
10 from sklearn.preprocessing import StandardScaler
   from sklearn.linear_model import LinearRegression, Lasso
   from sklearn.metrics import mean_squared_error
   def load data(df path):
        return pd.read_csv(df_path)
   def divide_train_test(df, target):
       X_train, X_test, y_train, y_test = train_test_split(df, df[target],
       return X train, X test, y train, y test
       return df[var].fillna(mean val)
   def remove_categorical_na(df, var):
       return df[var].fillna('Missing')
   def cap_outliers(df, var, cap, bigger_than=False):
           capped_var = np.where(df[var])cap, cap, df[var])
           capped_var = np.where(df[var]ccap, cap, df[var])
        return capped var
```

```
_def remove_numerical_na(df, var, mean_val);
        return df[var].fillna(mean_val)
    def remove categorical na(df, var):
        return df[var].fillna('Missing')
   def cap outliers(df, var, cap, bigger than=False):
           capped_var = np.where(df[var] cap, cap, df[var])
           capped_var = np.where(df[var] cap, cap, df[var])
       meturn capped var
   def transform skewed variables(df, var):
        return np.log(df[var])
        return np.where(df[var].isin(frequent_labels, df[var], 'Rare'))
    def train_scaler(df, output_path):
        scaler StandardScaler()
        scaler.fit(df)
        joblib.save(scaler, output_path)
        return scaler
   def scale_features(df, scaler):
        scaler = load(scaler) # with joblib probably
        return scaler.transform(df)
        lin_model Lasso(random_state 2909)
        lin_model.fit(scaler.transform(df[features]), target)
        joblib.save(lin model, output path)
        return lin model
74 def predict(df, model, features, scaler):
              model.predict_proba(scaler.transform(df[features]))
```

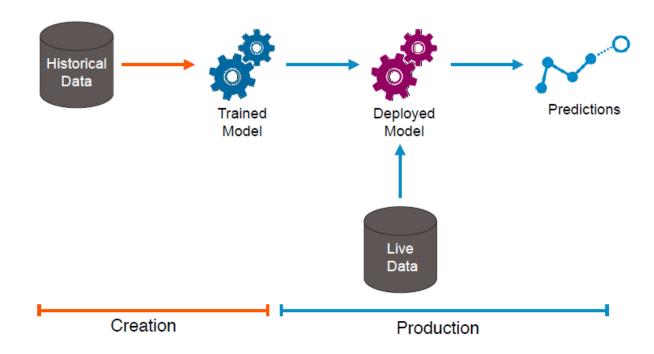
Production

- Deploiement
 - Mettre à disposition une prediction facile en term
- Evaluation
 - Mesure de la qualité des models déployés
- Monitoring
 - Suivi de la qualité des modèles
- Management
 - Améliorer la qualité des modèles avec une retro-action (continue)



Déploiement

- En phase de développement, on travaille sur des données d'apprentissage/test disponibles au niveau d'un référentiel
- En phase de production, on travaille sur des données live



Déploiement API Rest

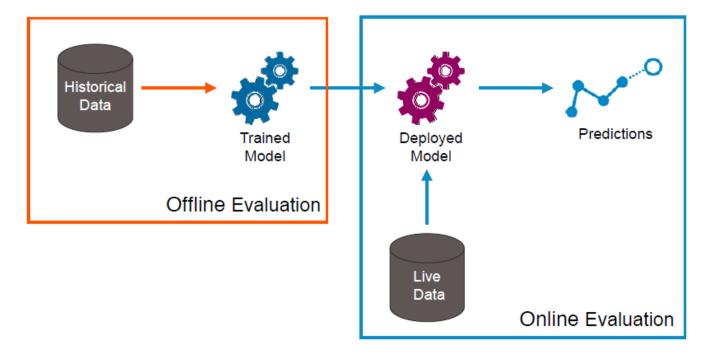
- ☐ Permettre le déploiement du modèle en faisant abstraction du langage utilisé pour le développement du modèle
 - N'est pas important qu'on ait utilisé R, python ou un autre langage
- ☐ Voir les solutions disponibles
 - Azure ML
 - Amazon AWS
 - Déploiement REST classique

Évaluation

- ☐ En phase de développement, on a évalué les métriques en relation avec le modèle considéré
- Après déploiement, les gestionnaires seront très probablement intéressé par un autre type de métrique en relation avec le modèle: métrique d'affaire ou business métrique.
- ☐ Essentiellement: est ce que ce modèle joue son role pour les besoins d'affaires
 - ➤ Prédire sur la base des données fournies par le client s'il va faire un achat
 - ➤ Modèle indique OUI mais le client ne va pas acheter. Combien vont faire la même chose?

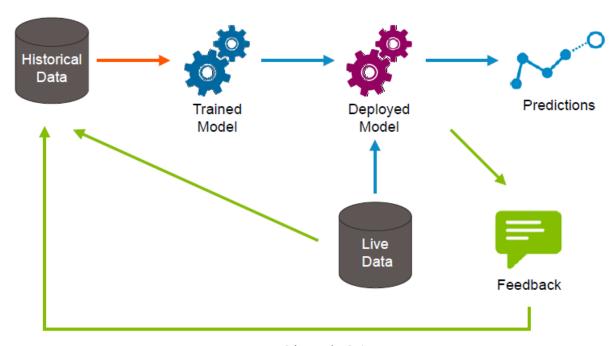
Évaluation | à quel moment

- ☐ Évaluation des métriques du modèle: se produit offline lors du développement
- ☐ Évaluation des métriques: peut se produire offline et online



Suivi et Management

- ☐ Permet de suivre la qualité du modèle à travers le temps
- ☐ Rétro-action sera importante pour améliorer la qualité du modèle
 - Prévoir le chargement des données live vers le référentiel pour l'apprentissage continu

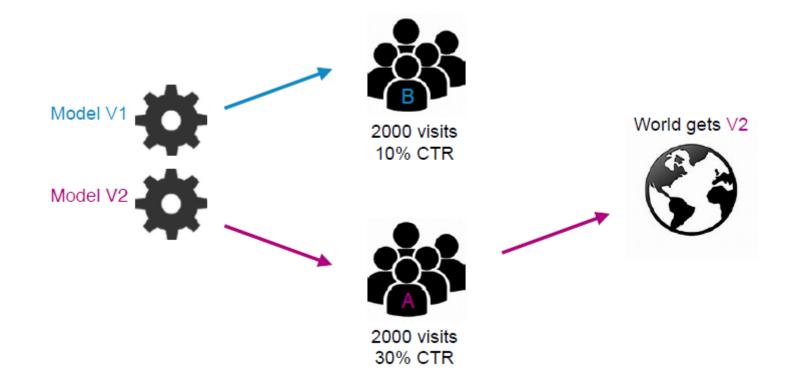


Apprentissage | Mise à jour modèle

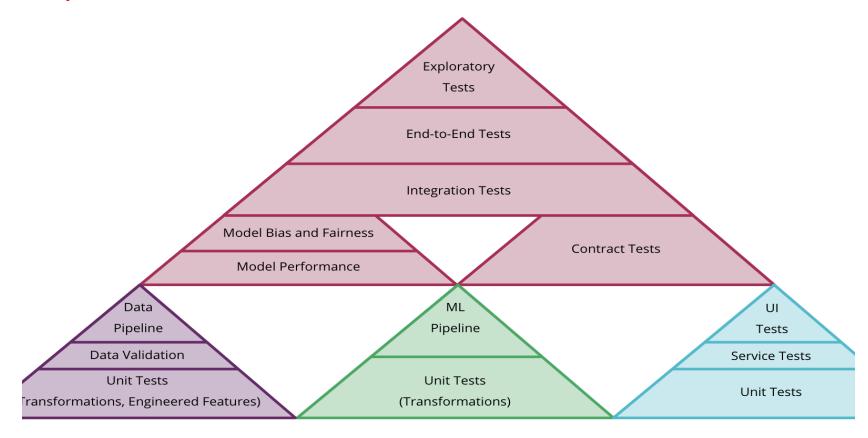
- ☐ Quand faut-il faire la mise à jour et pourqoui?
 - ➤ Selon le domain d'affaires, il peut y avoir des saisons et tendances particulières
 - Exemple: Halloween, Noel, saison de la grippe, vacances d'été, etc.
- ☐ Que faut-il faire pour le suivi?
 - Faire un suivi des métriques online et offline
- ☐ Decision pour update
 - ➤ A/B testing
 - ➤ Multi-armed bandits

A/B testing

☐ Il faut prendre une decision sur le nouveau modèle par rapport au modèle précédent

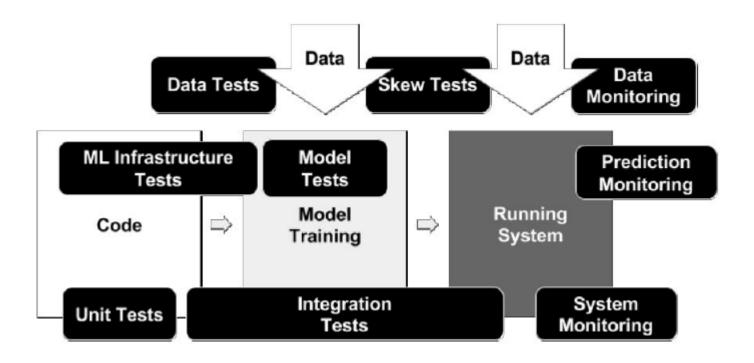


Exemple-Les tests



Exemple de combinaison de différentes pyramides de test pour les données, le modèle et le code

Tests du système ML/ surveillance



Références

- https://docs.microsoft.com/en-us/azure/machine-learning/machine-learning-publish-a-machine-learning-web-service
- https://tech.zalando.com/blog/scalable-fraud-detection-fashion-platform/?gh_src=4n3gxh1
- https://loads.pickle.me.uk/2016/04/04/deploying-a-scikit-learn-classifier-to-production/
- https://www.analyticsvidhya.com/blog/2016/05/8-reasons-analytics-machine-learning-models-fail-deployed/
- Note: les figures de cette presentation sont tirées des liens ci-dessus.