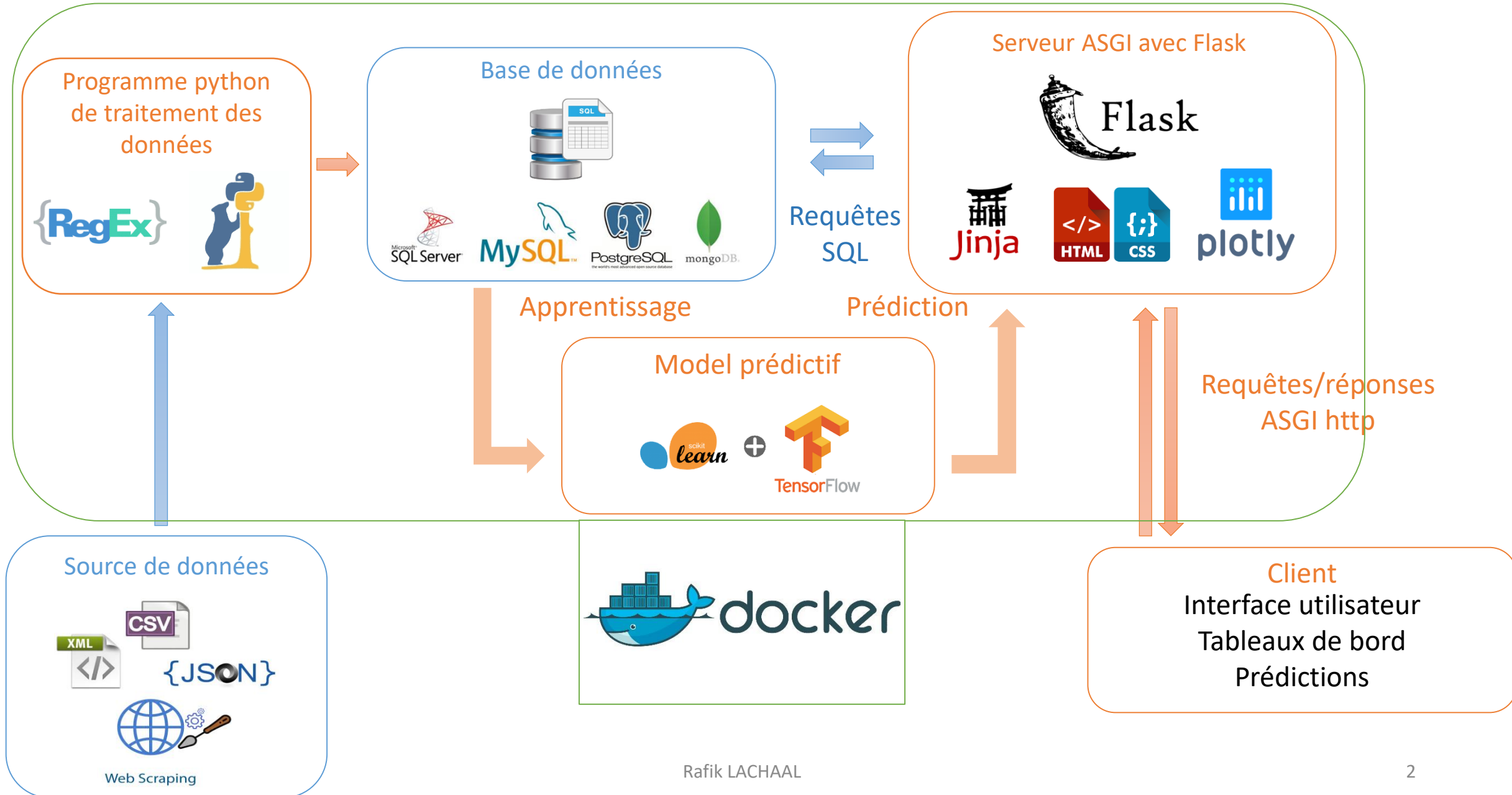


Créer et déployer une API de Machine Learning

Objectif de la Dev IA



Schéma fonctionnel d'une API d'IA



Un projet de machine learning de bout en bout

1. Cerner le problème et rechercher une vision d'ensemble :
 - Etablir un cahier des charges
 - un planning
 - définir les outils nécessaires
 - définir les critères de performances, ...
2. Récupérer les données :
 - Lister les données et la quantité nécessaire
 - Créer un environnement de travail : Docker
 - Modéliser une base de données,...
3. Explorer les données pour les comprendre :
 - Pandas, Matplotlib
 - Visualiser les données
 - statistiques descriptives
 - Etudier les corrélations entre variables
 - données manquantes et aléas,...

Un projet de machine learning de bout en bout

4. Préparer les données pour mieux exposer leurs structures sous-jacentes aux algorithmes d'apprentissage automatique
 - Nettoyer les données
 - Sélectionner les variables
 - Transformer les variables
 - Ecrire des fonction pour toutes transformation de données,...
5. Essayer beaucoup de modèles différents et constituer une liste restreinte des meilleurs d'entre eux.
 - Entraîner avec les paramètres usuels de nombreux modèles
 - Mesurer et comparer les résultats : Validation croisée, métriques de performances
 - Ingénierie sur les variables,...

Un projet de machine learning de bout en bout

6. Régler finement les modèles et les combiner en une solution performante :
 - Régler les hyperparamètres
 - Combiner les meilleurs modèles
 - Mesure de performance avec le jeu de test
7. Présenter votre solution
 - Créer une présentation agréable
 - Communiquer les résultats clé de manière visuel
8. Lancer surveiller et maintenir votre système
 - Raccorder sa solution au source de données
 - Ecrire le code qui vérifiera à intervalles régulier les résultats fournie par votre système en fonctionnement et déclenchera des alertes
 - Réentraîner régulièrement le modèles sur des données actualisées

Les modèles abordé durant la formation

- La régression linéaire :
 - Multiple
 - Polynomiale
 - Régularisé
- La régression logistique
- Machine à vecteur de support
 - Linéaire
 - Non linéaire
- Clustering :
 - K-means
 - Centre mobile
 - Ascendant/descendant hiérarchique
- Traitement du langage naturel : model TF-IDF
- Réseaux de neurones profonds
- Réseaux de neurones convolutifs

Autres modèles

- Naïve bayesien
- Arbres de décision
- Ensemble de forêts aléatoires
- Réduction de dimension
- Réseaux de neurones récurrent
- Apprentissage de représentations et apprentissage génératif avec autoencodeurs et des GAN
- Apprentissage par renforcement

Ce qui a été vus

- Les bases python : Boucles et conditions
- Analyse de données avec Numpy et Pandas
- La méthode MERISE
- Requêtes SQL :
 - simples et approfondi
 - Les jointures
 - Optimisation des requêtes