# Aide mémoire Projet Jedha Lead

A voir à remplacer la bdd OLTP mettre une bdd OLAP.

4 parties pour la présentation :

1. La problématique et la proposition de valeur.
2. Présentations de nos données
3. Architecture et justification de nos choix
4. Conclusion et perspective de developpement

Dans le projet pour la certification :

1. Cahier des charges
2. CI/CD
3. Script de réentrainement
4. Pilotage des performances et solution d emonitoring

Architecture

3 Blocs :

1 - Bloc de Workflow Orchestration via Airflow

2 - Bloc cloud computing pour l entrainement du modèle

3 - Web Application pour l affichage des données

De Gauche à droite pour suivre la donnée

* On va récupérer les différentes données via des sources externes API ou scraping de sites. Ces données sont mises à jour tous les ans et vont nous servir de sources de données pour entrainer notre modèle.
* Toutes ces données sont stockées de façon brut dans un Bucket S3.
* Une tâche va ensuite s’occuper de nettoyer et agréger ces données pour les structurer dans dans une base de données OLTP.
* Ces données sont ensuite testées via jenkins avant d’être envoyé dans une machine qu on va popper sur EC2 pour faire un entraînement et créer un modèle qui sera loggé sur MLFLOW
* On a ensuite des metadonnées pour le moment les predictions meteos qui vont enrichir la base de données tous les jours
* Ces nouvelles données vont permettre de faire une prédiction. Et une tâche airflow va se lancer quotidennement pour faire une prédiction et l enregistrer dans la base de données.
* Coté Web app, nous avons une API FastAPI qui va permettre d’aller récupérer les prédictions dans la base de données mais aussi d’autres infos permettant de données des indications diverses pour les pompiers.
* Ensuite il y a un dashboard fait sur streamlit

Pistes d’améliorations

L’idée c’est d’améliorer la qualité des données et d’ajouter différentes métadonnées qui permettrons d’améliorer notre modèle.

Il faut aussi ajouter plus de tests pour accroître la robustesse de l’application

Aider les équipes à organiser leurs équipes.

Déployer sur toute la France

Comment améliorer l application pour aider au mieux les équipes.