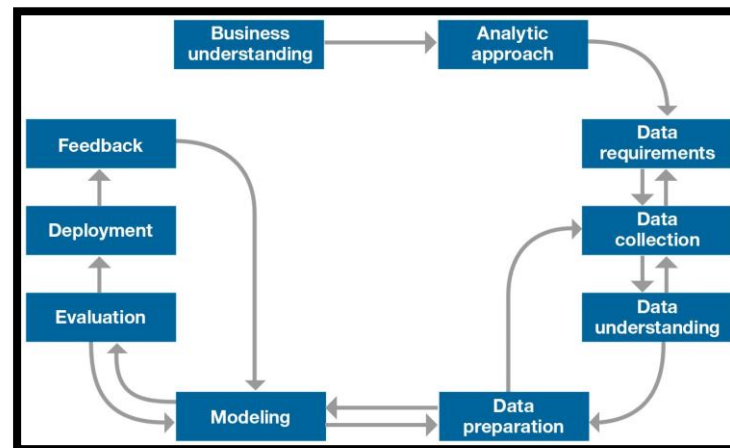


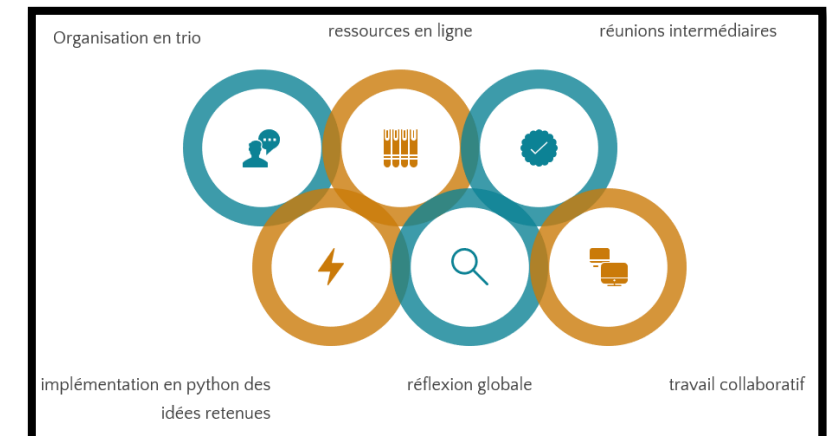
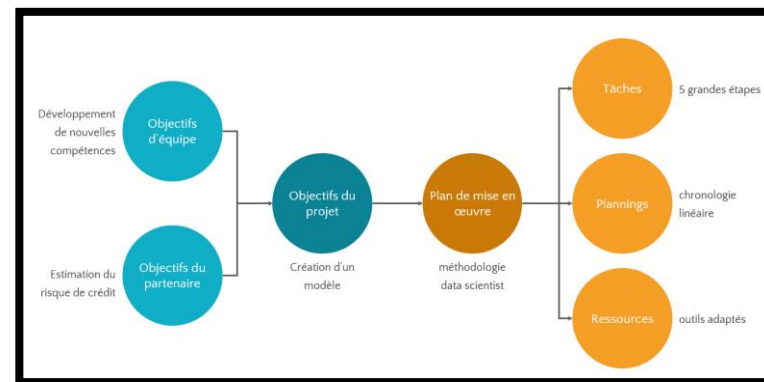
Présentation

Le but initial de ce projet était de répondre au problème de risque de crédit, c'est-à-dire prédire la solvabilité de clients d'une banque. Pour ce faire il nous a été demandé de réaliser un modèle de Machine Learning entraîné sur des données fournies par le client. Chaque ligne du dataset est compte bancaire, notre feature cible est un paramètre binaire (1 si il a déjà été en défaut), o y retrouve des indicateurs financiers et des données plus personnelles (âge, revenu moyen, ...).

Méthodes utilisées



Pour répondre à ce problème de machine learning nous avons découpés le projet en grandes étapes.



Résultats

Nous avons donc testé 8 modèles de classification différents. Nous avons choisis le modèle QDA pour car il a le meilleur AUC, puis avons optimisé les features et les paramètres.

| Modèle | Accuracy | AUC |
|--------|-----------------------|------------------------------|
| LR | 0.935 | 0.54 |
| LDA | 0.934 | 0.56 |
| QDA | 0.908 | 0.688 |
| KNN | 0.93 | 0.586 (2ème scalaire) |
| CART | 0.89 | 0.63 |
| NB | 0.935 | 0.646 (1er et 2ème scalaire) |
| SVM | 0.936 (2ème scalaire) | 0.555 (2ème scalaire) |
| RFC | 0.935 (1er scalaire) | 0.588 |
| XGB | 0.935 (2ème scalaire) | 0.59 |

QDA

