

Etude de cas projet fil rouge

Auteur de l'étude de cas : Mohamed EL AFRIT (2025)

Remarque : pour des raisons de confidentialité, ce document réadapté a été réadapté pour le travail demandé dans le cadre de l'UE FAD11 du CNAM de Paris

Objectifs pédagogiques

A l'issue de la formation, l'apprenant sera capable de :

- **(Technique)** → Concevoir et déployer un modèle de Machine Learning en ligne (via API), de produire un dashboard et une note méthodologique pour expliquer sa démarche, et de présenter oralement son travail,
- **(Gestion de projet)** → Planifier, organiser et piloter le projet de déploiement d'un modèle de Machine Learning en appliquant la méthode Scrum, en structurant un Product Backlog, en animant les rituels agiles et en assurant le suivi des itérations jusqu'à la livraison des livrables attendus

Mise en situation

Vous êtes *Consultant Data* au sein d'une société financière, nommée "[home Credit](#)", qui propose des *crédits* pour des personnes ayant peu ou pas du tout d'historique de prêt

Fraîchement embauché depuis une semaine avec [ce salaire annuel](#), vous avez fait connaissance avec vos collègues et votre nouveau bureau. Mais revenons à vos missions : il est temps de mettre les mains dans le cambouis ! Le **DSI** vous a donné l'accès à [la base de données](#). L'entreprise souhaite **développer un modèle de credit scoring** et de le **mettre en production**. Les données à disposition sont variées : *données comportementales, données provenant d'autres institutions financières*, etc. (**à vous de vous familiariser avec cette data !!!**).

Mission 1 : technique : conception et développement d'un modèle de Machine Learning

Créer la base de données de la Banque

- MCD, tables, jointures, PK, FK, connexion sécurisée, ..., backup (création dans l'art en respectant les compétences **RNCP** à valider, voir ci-dessous).

Créer le modèle de scoring et le mettre en production

- Construire un modèle de scoring et le déployer via une **API** sur le Web où en saisissant l'identifiant client (ou num de dossier de crédit), l'API renvoie bien la prédiction correspondante.
- Pensez à utiliser un outil gratuit et disponible plusieurs mois *en vue de vos entretiens techniques, jury,....* par exemple [Heroku](#) ([netlify](#) : alternative gratuite) où vous déployez [Flask](#) ou [Django](#) : [help1](#) et [help2](#) ; en passant par un fichier pickle contenant votre modèle sérialisé.
- Le focus pour le modèle de machine learning sera mis sur :
 - La conception du modèle, son évaluation et son interprétation compréhensible pour les métiers
 - La systématisation de la création de **features**, via des *jointures, groupby, LabelEncoder, OneHotEncoder,...* ou via la combinaison de features (rapport de 2 features, notamment montants, ...)
 - Dans le cadre de l'optimisation du modèle, penser à utiliser **SMOTE** (génération de

lignes pour ré-équilibrer le nombre de valeur cible à 1 par rapport à 0) et **Hyperopt** (optimisation des hyperparamètres).

- N'oublier pas de mettre en œuvre une **matrice de coût** adaptée au contexte de crédit afin de proposer une optimisation orientée métier et non pas technique : par exemple : le coût d'un *faux positif* (bon crédit considéré comme mauvais constitue un manque à gagner modéré pour la banque, une perte de marge) est différent d'un faux négatif (mauvais crédit considéré comme bon = constitue une perte importante pour la banque, un défaut de paiement et/ou une perte de capital non remboursé). Idéalement, vous montrez que l'optimum « métier » est différent de l'optimum du **fscore** ou autres mesures purement « techniques ».

Créer un dashboard interactif

Les **chargés de relation client** ont fait remonter le fait que les clients sont de plus en plus demandeurs de transparence vis-à-vis des décisions d'octroi de crédit. Cette demande va tout à fait dans le sens des valeurs que l'entreprise veut incarner. Votre **manager** décide donc de **développer un dashboard interactif** pour que les conseillers puissent expliquer de façon la plus transparente possible les décisions

Cahier des charges rédigé par le manager pour le dashboard :

Les spécifications du Dashboard devront a minima contenir les fonctionnalités suivantes :

- *Permettre de visualiser des informations descriptives relatives à un client (via un système de filtre).*
- *Permettre de visualiser le score et l'interprétation de ce score pour chaque client de façon intelligible pour une personne non experte en data science.*
- *Permettre de comparer les informations descriptives relatives à un client à l'ensemble des clients ou à un groupe de clients similaires.*

Livrables attendus pour la partie technique

1. Une organisation [DevOps Azure](#) qui contient :
 - Un **board** où vous répartissiez les tâches entre les membres de votre groupe suivant votre méthode de travail [Agile](#), [Scrum](#)
 - Les **pipelines** pour le CI/CD
2. Un **repo** [Github](#) contenant :
 - Un fichier README (où vous expliquez comment lancer les scripts, ...),
 - Un notebook Python (non cleané, pour comprendre votre démarche :
 - Les problèmes rencontrés sur le jeu de données
 - Comment vous avez nettoyé les données
 - Votre modélisation (du preprocessing à la prédiction).
 - Le code générant le dashboard et permettant de déployer le modèle sous forme d'API
- L'URL de la WebApp mise en ligne et répondant au cahier des charges précisé ci- dessus.

Ressources complémentaires

- Un [article](#) donnant quelques bonnes pratiques pour le design de dashboard.
- Un [document](#) décrivant les bonnes pratiques pour réaliser des graphiques clairs et pertinents.
- Des informations sur deux bibliothèques permettant de construire des dashboards interactifs en Python : [Dash](#) et [Bokeh](#).
- [Ce lien](#), et [celui-ci](#) pour aider à construire l'interprétabilité du modèle.

Mission 2 : gestion de projet : piloter le projet avec la méthode Scrum

Livrables pendant la phase d'initialisation du projet

Etude d'opportunité (Business Case)

Décrire les justifications économiques du projet (coûts, bénéfices, risques), les opportunités identifiées, et la manière dont le projet générera de la valeur. Expliquer le modèle de revenus et la proposition de valeur.

- [Lecture 1](#)
- [Lecture 2](#)
- [Modèle de document](#)

Charte de projet

Document officiel en une page qui définit les objectifs, le périmètre, les parties prenantes, et autorise le chef de projet à utiliser les ressources. Il permet de formaliser l'engagement de la direction.

- [Lecture 1](#)
- [Lecture 2](#)
- [Modèles de documents](#)

Registre et Mapping des Parties Prenantes

Identification et analyse des parties prenantes influençant le projet. Ce document permet d'élaborer une stratégie d'engagement et de communication adaptée à chaque acteur.

- [Lecture 1](#)
- [Lecture 2](#)
- [Modèles de documents](#)

Livrable de la phase de planification du projet

Matrice des exigences (Requirement Traceability Matrix)

Définition des exigences métiers et matrice de traçabilité qui assure que toutes les exigences sont couvertes tout au long du cycle de vie du projet.

- [Lecture 1](#)
- [Lecture 2](#)
- [Modèle de document](#)

Spécification fonctionnelle et technique

Description précise des fonctionnalités attendues du produit. Ce document est le point de référence pour garantir que les besoins des utilisateurs sont bien pris en compte.

Spécifications techniques nécessaires pour la réalisation du projet, incluant l'infrastructure, le réseau, et les outils technologiques. Ce document assure la cohérence de la solution.

- [Lecture 1](#)
- [Lecture 2](#)
- [Modèle de document](#)

Project Management Plan

Plan global de gestion incluant le périmètre, le calendrier, le budget, les risques, et les stratégies de communication. Ce document sert de guide pour l'exécution et le suivi du projet.

- [Lecture 1](#)
- [Lecture 2](#)
- [Modèle de document](#)

Livrables des phases d'exécution et de monitoring du projet

Plan de Test Détaillé

Définition des stratégies, types, et objectifs des tests à réaliser pour garantir la qualité du

produit. Le plan de test inclut le calendrier des tests, les scénarios, et les critères de validation.

- [Lecture 1 : Stratégie de test](#)
- [Lecture 2 : Plan de test](#)
- [Lecture 3 : Cahier de recette](#)
- Modèles de documents :
 - [Cahier de recette](#)
 - [Plan de test](#)

Reporting des Tests

Document résumant les résultats des tests réalisés, les anomalies détectées et les actions correctives mises en place. Il permet de valider la qualité du livrable.

- [Lecture : Rapport Flash](#)
- [Modèles de documents](#)

Reporting régulier (Project Status Report)

Rapport mensuel présentant l'avancement du projet, les risques, les blocages, et les prochains jalons. Ce reporting permet de suivre le bon déroulement du projet. Afficher et interpréter les indicateurs de performance de la méthode EVM (Earned Value Analysis)

- [Lecture 1 : Suivi du budget](#)
- [Lecture 2 : Comité de pilotage](#)
- [Lecture 3 : EVM](#)
- [Vidéo 1 à consulter](#)
- [Vidéo 2 à consulter](#)
- Modèles de documents :
 - [Suivi de budget](#)
 - [Comité de pilotage](#)

Livrable de la phase de clôture du projet

Bilan du Projet

Synthèse finale du projet présentant les réussites, les échecs, et les leçons apprises. Sert de référence pour améliorer les pratiques futures.

- [Lecture 1 : Bilan de projet](#)
- [Lecture 2 : Explication du modèle](#)
- [Modèle de document](#)

Documentation d'Exploitation et de Maintenance

Instructions pour la gestion opérationnelle et la maintenance du produit. Permet aux équipes techniques de garantir la continuité des opérations.

Guide Utilisateur

Documentation destinée aux utilisateurs finaux expliquant comment utiliser le produit. Il facilite l'adoption du produit par les utilisateurs.