

# Prédiction des faillites d'entreprises



Notre objectif dans ce projet est d'utiliser les modèles ensemblistes pour nous aider à détecter les entreprises proches de la faillite.



## Livrables

- créer un repo git et un espace sur github/gitlab pour le projet
- présenter un planning de travail
- faire une EDA des données (notebook\_EDA)
- analyser et nettoyer cette base de données (procédure expliquée dans le notebook EDA)
- présenter la procédure suivie pour trouver un modèle adapté (notebook\_model)
- sélectionner les données (features) pertinents pour le modèle (dans le notebook model)
- créer un modèle d'IA entraîné et adapté au problème (format pickle)
- proposer un protocole d'utilisation du modèle pour faire des prédictions
- faire une présentation qui explique votre démarche et les résultats obtenus. Vous devez expliquer (10 slides maximum)



## Membres du groupe

Zohra

Matthis

# Jeudi 21/10

## Matin

### ▼ Notebook EDA

- Exploration des données
- Balancing
- Visualisations

## Après-midi

### ▼ Notebook Model

- Feature Selection
- Standardisation
- Régression Logistique
- KNN Classifier
- Gradient Boosting
- Random Forest
- Decision Trees

### ▼ Méthodes ensemblistes

- Bagging
- Boosting
- Stacking

# Vendredi 22/10

### ▼ PCA

- Variance expliquée
- Nombre de composants

Analyse des dimensions

Cercle de corrélation

### ▼ Analyse approfondie de l'EDA

Visualisations

Interprétations

Correlation



## Mercredi 27/10

### Après-midi

Feature Selection

Modélisation des modèles

Roc curve

Learning curves

Pipeline

Multi Layer Perceptron

## Jeudi 28/10

### Matin

Performance des modèles

Sélection du Model et Evaluation

Mise en commun

### Après-midi

Protocole d'utilisation

Pickle

Git

## Vendredi 29/10

- Répartition des tâches
- Présentation technique