**BANKING PROJECT**

**OBJECTIF:**

V1 : Analyser les demandes de prêts pour comprendre les facteurs qui influencent l’approbation ou le rejet

V2 : construire un modèle de machine learning pour prédire la probabilité d'acceptation d’un prêt.

V3 : automatisation

**DONNÉE:**

Dataset : <https://www.kaggle.com/datasets/architsharma01/loan-approval-prediction-dataset?utm_source=chatgpt.com>  
  
Fichier: loan\_approval\_dataset.csv

Taille : 385 ko

|  |  |
| --- | --- |
| **Nom de la colonne** | **Signification** |
| **loan\_id** | Identifiant unique du prêt (utilisé pour distinguer chaque ligne dans le dataset). |
| **no\_of\_dependents** | Nombre de personnes à charge du demandeur (enfants, parents, etc.). |
| **education** | Niveau d’éducation du demandeur (ex. : diplômé, non diplômé). |
| **self\_employed** | Indique si le demandeur est travailleur indépendant (Oui ou Non). |
| **income\_annum** | Revenu annuel brut du demandeur (souvent en devise locale, comme CAD ou USD). |
| **loan\_amount** | Montant total du prêt demandé. |
| **loan\_term** | Durée du prêt, généralement exprimée en mois. |
| **cibil\_score** | Cote de crédit (CIBIL) du demandeur, qui reflète sa solvabilité. Plus le score est élevé, mieux c’est. |
| **residential\_assets\_value** | Valeur des biens résidentiels détenus (maisons, appartements). |
| **commercial\_assets\_value** | Valeur des biens commerciaux détenus (bureaux, locaux professionnels). |
| **luxury\_assets\_value** | Valeur des biens de luxe (voitures de luxe, bijoux, œuvres d’art). |
| **bank\_asset\_value** | Valeur des actifs bancaires (épargne, dépôts à terme, investissements). |
| **loan\_status** | État du prêt (Approuvé, Rejeté, etc.). C’est souvent la variable cible pour un modèle prédictif. |

**ENVIRONEMENT ET SUIVI: github**

**CONNECTION AU DATASET: via python**

**TRANSFORMATION DES DONNÉES (ETL) : via python**

**Creation des colones :**

**total\_asset\_value** = residential\_assets\_value + commercial\_assets\_value + bank\_asset\_value + luxury\_assets\_value

**asset\_to\_loan\_ratio** = total\_asset\_value / loan\_amount

**risk\_category** :

Low : < 600

Medium : 600–750

High : > 750

### **is\_low\_income :**

### 1 si income\_annum < 300000, sinon 0

**CREATION DE MODELE : via SQl**

**MODELE EN ETOILE**

DIMENSION

DIMENSION(dim\_customer)

DIMENSION(dim\_credit\_score)

FACT TABLE(loan\_application)

DIMENSION(dim\_assets)

**VISUALISATION : via tableau**