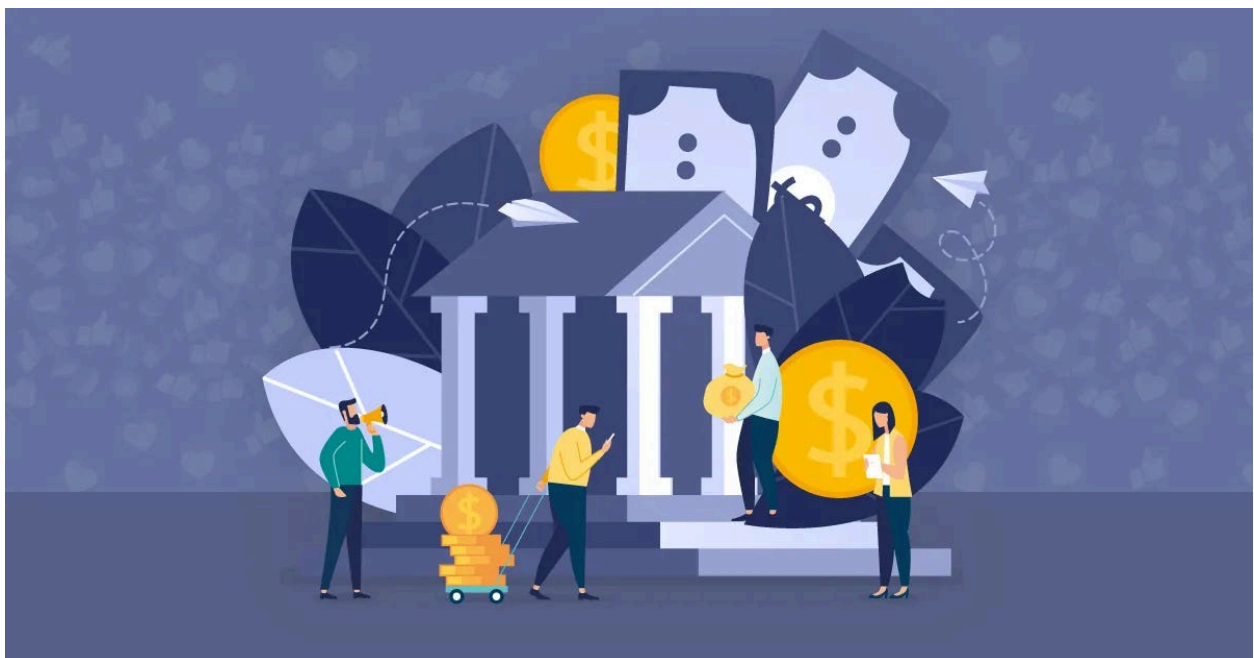


Bank Marketing

Why did the banker break up with their partner? Because they lost interest!

Bank Marketing UCI



Le dataset Bank Marketing UCI, provenant d'une banque portugaise, fournit des informations détaillées sur les campagnes de marketing **visant à promouvoir la souscription de dépôts à terme**.

L'ensemble de données contient des enregistrements d'interactions avec les clients **dans le cadre de campagnes de marketing direct menées par téléphone**. Chaque enregistrement comprend une série de caractéristiques qui décrivent les attributs démographiques et comportementaux du client,



ainsi que le résultat de la campagne de marketing, à savoir **si le client a souscrit à un dépôt à terme ou non.**

L'ensemble de données comprend divers attributs du client, tels que :

- **Âge** : L'âge du client.
- **Type d'emploi** : Le type d'emploi du client (par exemple, cadre, technicien, étudiant, etc.).
- **Formation** : Le niveau d'études le plus élevé atteint par le client.
- **État civil** : Si le client est célibataire, marié ou divorcé.
- **Type de communication** : Le moyen utilisé pour communiquer avec le client (par exemple, téléphone portable).
- **Durée du dernier contact** : Durée du dernier contact en secondes.
- **Nombre de contacts de la campagne** : Nombre de contacts effectués au cours de cette campagne de marketing spécifique.
- **Contacts des campagnes précédentes** : Nombre de contacts établis au cours des campagnes précédentes.
- **Résultat de la campagne précédente** : Le résultat de la dernière campagne de marketing (par exemple, succès, échec, inexistant).

Cet ensemble de données est beaucoup utilisé pour des tâches telles que la segmentation de la clientèle, la prédiction du taux de désabonnement et l'optimisation des stratégies de marketing.

En analysant ces points de données, les entreprises peuvent comprendre quelles caractéristiques des clients sont prédictives d'une souscription réussie à un dépôt à terme, **ce qui leur permet d'adapter plus efficacement leurs stratégies de marketing.**



AutoML



PyCaret une librairie de Machine Learning open source et low code sur **Python**. Cette solution automatise le **workflow** de l'apprentissage automatique de bout en bout. Grâce à l'automatisation des tâches et la gestion des modèles ML, PyCaret accélère le cycle d'expérimentation. Les data scientists sont ainsi beaucoup plus productifs et capables de développer des modèles d'apprentissage automatiques encore plus performants.

Avant même d'ouvrir votre éditeur de code, vous réalisez une veille concise sur les éléments suivants :

- 1) Qu'est ce que l'**apprentissage automatique supervisé** ?
- 2) Qu'est ce que les **données étiquetées** (ou labellisées) ?
- 3) En particulier, qu'est ce que la **classification supervisée** (ou méthode de classement) ?
- 4) Qu'est ce qu'on appelle **un modèle** d'apprentissage automatique ?
- 5) Qu'est ce que les **données d'entraînement** ? Qu'est ce que l'entraînement d'un modèle de classification supervisée ?
- 6) Qu'est ce que la donnée **cible** (ou le target) ?



- 7) A quoi sert **la phase d'entraînement** d'un modèle d'apprentissage automatique ?
- 8) A quoi sert **la phase de prédiction** d'un modèle d'apprentissage automatique ?
- 9) Qu'est ce que le **prétraitement des données** et pourquoi est-il **important à réaliser avant** l'entraînement d'un modèle d'apprentissage automatique ?
- 10) Comment l'**évaluation d'un modèle d'apprentissage automatique** pour une tâche de classification peut-elle se faire ? Etudiez en particulier et seulement, la métrique d'**accuracy**. En quoi est-elle un indicateur de performance d'un modèle ?
- 11) Qu'est ce que l'**AutoML** et pourquoi est-il pertinent pour vous en tant que **débutant** dans le domaine de l'intelligence artificielle ?
- 12) Qu'est ce qu'un **arbre de décision** pour la classification supervisée ? Comment fonctionne-t-il ? (félicitations, vous avez appris votre premier modèle d'intelligence artificielle !).

Vous suivez les étapes ci-dessous pour **la construction de votre outil de prédiction** :

- 1) Vous récupérez les données marketing [ici](#), ainsi que le descriptif du dataset [ici](#).
- 2) Vous importez vos données à l'aide de Pandas et vous les passez en revue. **Quelles données vont-elles servir à l'apprentissage ?** Quelle variable va-t-elle servir de **target** ?



- 3) Vous identifiez les **anomalies** présentes au sein de vos données et **vous les traitez**.
- 4) Vous réalisez **une analyse exploratoire complète**.
- 5) Vous vous familiarisez avec la documentation de la **librairie Pycaret**, vous l'**installez** et vous l'**importez** sur votre machine.
- 6) Vous définissez l'**environnement de Pycaret** pour une expérience de classement.
- 7) Vous vous **entraînez et comparez des modèles de classification** (c'est automatique vous avez juste besoin d'une seule commande !).
- 8) Votre dernière manipulation résulte en un **tableau de performances** des modèles entraînés. **Interprétez** les résultats de la colonne "**Accuracy**".
- 9) Affichez le **meilleur modèle entraîné et identifié par Pycaret**. Quel est ce modèle ?





Compétences visées

- Analyse de données
- Apprentissage automatique

Rendu

L'évaluation de ce projet se fera sur deux aspects :

1. Une présentation explicative de votre travail sous forme de diapositives.
2. Un repository github public nommé **bank-marketing**, contenant les éléments suivants :
 - a. Un **notebook Python propre et commenté** (introduction, conclusion, etc) contenant le procédé de développement de votre outil de prédiction, du nettoyage à la modélisation des données, en passant par l'**analyse exploratoire**. **Pensez à répondre à la problématique**. Vous pouvez avoir au maximum deux notebooks, un pour l'exploration et l'autre pour la modélisation de données.
 - b. Un fichier **README.md** présentant le contexte du projet, les données et leur analyse, les algorithmes utilisés et une conclusion sur votre travail. Pensez à inclure la veille réalisée.

Base de connaissances

- [PyCaret 3.0 | Documentation](#)
- [StatQuest | A Gentle Introduction to Machine Learning](#)
- [Classification in Machine Learning: An Introduction](#)