

Compte rendu N°1

Énoncé

- L'objectif de ce compte rendu est de réaliser la phase de la préparation de données en interagissant avec l'application Web, telles que :
 - ✓ Nettoyage de données
 - ✓ Transformation de données (Normalisation de données)
 - ✓ Encodage des variables catégoriques
 - ✓ Réduction des attributs
 - ✓ Interaction avec une application Web via le framework **streamlit**
- Les comptes rendus doivent être répartis entre **les binômes d'étudiants**
- Le dernier délai à remettre les comptes rendus est **18 février**.

Frameworks Web Python

Streamlit, **Flask** et **Django** sont des frameworks populaires en Python utilisés pour le développement d'applications web. Voici une brève description de chacun d'entre eux :

1. **Streamlit** : C'est un framework open source qui permet de créer rapidement des applications web interactives pour la visualisation et l'exploration de données. Il offre une syntaxe simple et intuitive pour créer des interfaces utilisateur réactives et dynamiques. Streamlit est souvent utilisé pour développer des tableaux de bord interactifs, des démonstrations de modèles d'apprentissage automatique et des applications de visualisation de données.
2. **Flask** : C'est un micro-framework web léger et flexible en Python. Il offre les fonctionnalités de base pour créer des applications web, telles que le routage des URL, la gestion des requêtes et des réponses, et la gestion des sessions. Flask est souvent utilisé pour développer des applications web simples et légères, avec une flexibilité et une extensibilité élevées.
3. **Django** : C'est un framework web complet en Python, conçu pour faciliter le développement d'applications web complexes. Il offre un ensemble complet de fonctionnalités, y compris la gestion des bases de données, l'authentification des utilisateurs, la gestion des formulaires, la gestion des URL, etc. Django suit le principe du "batteries included", ce qui signifie qu'il inclut de nombreuses

fonctionnalités prêtes à l'emploi. Django est souvent utilisé pour développer des applications web robustes et évolutives.

Chacun de ces frameworks a ses propres caractéristiques et avantages, et le choix dépendra des besoins spécifiques du projet. Streamlit est idéal pour les applications de visualisation de données interactives, Flask est adapté aux applications web légères et personnalisées, tandis que Django convient aux applications web plus complexes et structurées.

Dans ce compte rendu, vous utiliserez le framework Streamlit.

Tâches à réaliser

Mettre les variables indépendantes (*Country*, *Age*, *Salary*) dans X et la variable dépendante (*Purchased*) dans Y de la base de données « *achat* »

En utilisant les données X, réalisez les phases de la préparation de données suivantes :

- ✓ **Nettoyage de données**
- ✓ **Transformation de données (Normalisation de données)**
- ✓ **Encodage des variables catégoriques**
- ✓ **Réduction des attributs**

Avec l'interaction Web, implémentez les options suivantes :

- ✓ Un upload de fichier de la base de données à partir du PC
- ✓ Un affichage du nombre des instances et des attributs de la base de données « *achat* »
- ✓ Un affichage des premières lignes de la base de données « *achat* »
- ✓ Un bouton pour **nettoyer** les données X
- ✓ Un affichage des premières lignes de X après la phase de **nettoyage**
- ✓ Un bouton pour **normaliser** X (les attributs *Age* et *Salary*)
- ✓ Un affichage des premières lignes de X après la phase de **normalisation**
- ✓ Un bouton pour **encoder** la colonne catégorique *country*
- ✓ Un affichage des premières lignes de X après la phase de **l'encodage**
- ✓ Un bouton pour **réduire** les attributs *Age* et *Salary* en seul attribut
- ✓ Un affichage des premières lignes de X après la phase de **réduction**