

PROJET 04 : Système d'Alerte Précoce de Décrochage Scolaire

OBJECTIF DU PROJET

Les écoles perdent des financements et les étudiants perdent leur avenir lorsqu'ils décrochent. Identifier tôt les étudiants à risque permet une intervention ciblée.

FICHIER DE DONNÉES

- decrochage_scolaire.csv
- 800 lignes
- Inclut des valeurs manquantes, valeurs aberrantes et déséquilibres pour un apprentissage réaliste

TÂCHES À RÉALISER

■ TÂCHE PRINCIPALE

Classer l'étudiant comme À Risque ou Sûr

■ TÂCHES SECONDAIRES (OPTIONNELLES)

1. Identifier les 3 principaux facteurs favorisant le décrochage
2. Regrouper les étudiants en "High Achievers", "Strugglers", "Coasters"
3. Prédire la moyenne finale pour les étudiants restants
4. Recommander des interventions spécifiques (ex: "Pass Bus" pour les longs trajets)

LIVRABLES ATTENDUS

1. ****Exploration des Données (EDA)****
 - Analyse statistique descriptive
 - Visualisations des distributions
 - Analyse des corrélations
 - Identification des valeurs manquantes et aberrantes
2. ****Ingénierie des Fonctionnalités****

- Création de nouvelles variables pertinentes
- Transformation des variables existantes
- Encodage des variables catégorielles
- Normalisation/Standardisation si nécessaire

3. ****Modélisation****

- Choix et justification du/des modèle(s)
- Entraînement et validation
- Optimisation des hyperparamètres
- Évaluation avec métriques appropriées

4. ****Résultats et Interprétation****

- Performance du modèle principal
- Résultats des tâches secondaires (si réalisées)
- Insights et recommandations
- Visualisations des résultats

CONSEILS

- Commencez par une exploration approfondie des données
- L'ingénierie des fonctionnalités est CRUCIALE pour ce projet
- Documentez votre démarche et vos choix
- Comparez plusieurs approches si possible
- Pensez à l'aspect pratique et utilisable de vos résultats

Bon travail ! ■