

Description du problème

Vous êtes engagé par la directrice des ressources humaines d'une entreprise pour enquêter sur les raisons pour lesquelles les employés démissionnent. Elle vous a fourni une base de données sur les employés de l'année précédente. Elle veut comprendre quels sont les indicateurs les plus pertinents pour savoir qui envisage de quitter.

Vous disposez d'un jeu de données que vous devez analyser et sur lequel vous allez finalement construire des modèles qui prédisent si un employé va démissionner ou non.

Description des variables

- Id - numéro d'identification unique
- satisfaction_level - niveau de satisfaction de l'employé
- last_evaluation - score dans la dernière évaluation
- number_project - le nombre de projet dans lesquels l'employée participe
- average_monthly_hours - la moyenne des heures mensuelles
- time_spend_company - temps avec l'entreprise
- Work_accident - accident du travail
- left - quitter l'entreprise - variable cible
- promotion_last_5years - une promotion au cours de 5 dernières années
- department - le département
- salary - niveau de salaire