

# Projet data mining (STID 2A)

Envoi du rapport et du code : 05/12/2023 minuit

Date de soutenance : 07/12/2023

Titre du mini-projet : Absentéisme au travail

Remarque : vous devez faire en sorte d'utiliser TOUTES les connaissances pertinentes acquises pendant les cours de data mining du BUT Sciences des Données (2A et 3A). L'utilisation de connaissances pertinentes dans d'autres matières sera très appréciée.

## Contexte

Le jeu de données *Absenteeism\_at\_work.csv* contient des enregistrements sur l'absentéisme au travail concernant la période juillet 2007 – juillet 2010 dans une entreprise de messagerie au Brésil.

**Objectif de votre étude** : explorer et extraire de la connaissance de façon à améliorer la qualité de la prévention de l'absentéisme.

**Données disponibles** : *Absenteeism\_at\_work.csv*, *Attribute Information.docx*

**Langage de programmation** : Python

**Rendu** : Rapport (10 pages)

**Evaluation** : rendu + soutenance

## Bibliographie

Martiniano, A., Ferreira, R. P., Sassi, R. J., & Affonso, C. (2012). Application of a neuro fuzzy network in prediction of absenteeism at work. In *Information Systems and Technologies (CISTI)*, 7th Iberian Conference on (pp. 1-4). IEEE.