Mise en contexte

La compagnie Mondrege vous demande de les aider à prévoir la consommation de certaines drogues d'un individu. Ils ont accumulé une très grande quantité de données sur plusieurs facteurs possible. À l'aide de celle-ci, vous devez créer un logiciel qui permet de prévoir la consommation de **Nicotine**.

Un fichier csv de données vous est fourni. Vous avez le choix de la séparation de ces données pour faire l'entrainement ou autre manipulation de votre choix, ils sont à vous pour vos tests. Vous devez penser à avoir au minimum 2 groupes. Le premier pour l'entrainement et le deuxième pour la validation. Il est laissé à votre discrétion de réduire le nombre d'attributs, si certains ont une moins grande importance.

Les paramètres des données sont décrits dans les tableaux suivant :

Paramètres	Description	
Id	Numéro d'identification d'une donnée	
Age	Étendu d'âge d'un individu	
Genre		
Education	Plus haut niveau d'éducation atteint	
Pays		
Ethnécité		
Neuroticisme	Score dans les traits de personnalité du modèle des big fives	
Extraversion	Score dans les traits de personnalité du modèle des big fives	
Ouverture	Score dans les traits de personnalité du modèle des big fives	
Agréabilité	Score dans les traits de personnalité du modèle des big fives	
Conscienciosité	Score dans les traits de personnalité du modèle des big fives	
Impulsivité	Score d'impulsivité calculé par BIS-11	
Recherche de sensations	Score de recherche de sensations calculé par ImpSS	

*Auteurs: Richard Rail et Lévis Thériault; Mise à jour: 2018-03-15

Paramètres	Description
Alcool	Dernière consommation d'une drogue
Amphetamines	0 : Jamais 1 : Il y a plus d'une décennie. 2 : Il y a une décennie. 3 : Il y a un an.
Amyl nitrite	
Benzodiazepine	
Caffeine	
Cannabis	
Chocolate	4: Il y a un mois.
Cocaine	5 : Il y a une semaine.
Crack	6: Il y a un jour.
Ecstasy	
Heroin	
Ketamine	
Legal highs	
LSD	
Methadone	
Magic mushroom	
Nicotine	
Semeron	
Volatile	
substance abuse	