

Enoncé - 1

mardi 30 mars 2021 09:45

Projet : reconnaissance d'activité humaine via un smartphone.

Il existe beaucoup d'applications permettant à un utilisateur de smartphone de connaître en direct (en temps réel) ses performances sportives (compter son nombre de pas, sa vitesse etc...). Pour cela, l'application a besoin d'identifier le type d'activité de l'utilisateur. L'idée du projet est de créer un programme qui permet de reconnaître l'activité de l'utilisateur (son mouvement : jogging - marche - monter les escaliers...) à partir de données captées sur son appareil mobile. Vous allez donc implémenter un algorithme "maison" de classification supervisée.

En quoi consiste la classification supervisée ?

Principe :

L'objectif de la classification consiste à attribuer une classe/une catégorie à chacune des observations d'un jeu de données.

On dispose au départ d'un ensemble de données d'apprentissage pour lesquelles le classement est connu (d'où le nom "supervisée"). Sur base de ces exemples, des règles vont pouvoir être établies pour permettre d'attribuer une catégorie à de nouvelles données du même type.

Exemple de classification supervisée :

De manière très schématique, supposons que l'objectif soit de déterminer un type de fleur (iris Sétosa - iris Versicolour - iris Virginica) en fonction de la longueur du pétale, sa largeur, la longueur du sépale, sa largeur. On aurait ainsi un ensemble de données ressemblant à ceci :

Long pétale	Larg pétale	Long sépale	Larg sépale	Type
...	Sétosa
...	Sétosa
...	Sétosa
...	Sétosa
..				Virginica
..				Virginica
..				Virginica
..				Virginica
...				
..				Versicolour
..				Versicolour

A partir de ces exemples, l'algorithme va établir des règles, trouver des patterns qui lui permettront par la suite lorsqu'une nouvelle donnée arrive

Long pétale	Larg pétale	Long sépale	Larg sépale
...

de déterminer de quel type d'iris il s'agit.

Quelles sont les différentes étapes d'une classification supervisée?

Revenons au projet :

Pour atteindre notre objectif de classification de mouvement , vous disposez de 15 fois 24 (+1) fichiers de données.



Un informaticien qui travaille sur des données ne peut évidemment pas faire fi de bien comprendre les données qu'il a à sa disposition et ce qu'il veut en faire. C'est une étape très importante pour éviter de travailler à l'aveugle et laisser passer des incohérences.

- **Prenez donc d'abord le temps de découvrir à quoi correspondent les données qui sont à votre disposition** en lisant attentivement les notes « context » et « content » du lien suivant : <https://www.kaggle.com/malekzadeh/motionsense-dataset> (Téléchargez ces données si vous ne les avez pas déjà)



Nous allons implémenter un algorithme de classification qui permettra de déterminer le mouvement de l'utilisateur sur base de ses données.