

Fiche de proposition d'un projet tutoré

Proposé par : Abdessamad Imine

Email : abdessamad.imine@univ-lorraine.fr (bureau 106)

Date : Septembre 2019

Type : Classique () Exploratoire (X)

Description du projet

Développement d'un robot pour extraire des données à partir de Twitter

Twitter est un réseau social de microblogage qui permet à un usager d'envoyer et transférer des messages brefs et instantanés. Un profil peut être suivi par plusieurs utilisateurs. Le moteur de recherche proposé sur Twitter est souvent limité et aléatoire. Ce qui nécessite plusieurs requêtes de recherche ainsi qu'un temps important de connexion pour croiser et affiner les résultats de recherche.

Objectifs à atteindre

Concevoir et réaliser un robot pour charger un profil de Twitter, analyser ses données (qui sont représentées en général dans un format HTML ou XML), et les mémoriser dans une base de données. Ensuite, le robot doit charger tous les suiveurs et les suivis qui sont en relation avec le profil et les stocker dans une file d'attente afin d'analyser par la suite leurs profils. L'usage d'une base de données va nous permettre d'avoir plus de souplesse pour manipuler les données des données provenant de Twitter.

Outils à utiliser

Le langage java.

L'outil sélénium (www.seleniumhq.org)