Documentation Technique Scrapper

Réalisé par Guillaume Tézenas du Montcel et Romain Blazevic

Le Scrapper a pour but de récupérer les données des meilleurs joueurs du jeu « League of Legend » sur le site « League of Graph ». Les données récupérées sont :

* Pseudo
* Rank
* Serveur
* Elo (niveau de grade)
* LP (score dans ce grade)
* Victoire (le nombre de victoire)
* %Victoire (le pourcentage de victoire)

Toute ces données seront envoyées dans une base de données MongoDB.

Le Scrapper est codé en Python. Il utilise comme librairie :

* Regex
* Csv
* Requests
* Pandas
* Html
* BeautifulSoup

Le Scrapper contient une Classe « Scrapper » qui possède des méthodes pour récupérer les données.

La classe est divisé en 3 groupes de méthodes distinctes :

* Connexion :

Les requêtes à répétition peuvent être détecter par le site est bloqué notre connexion. Des méthodes ont été créé pour palier ceci :

* + Get\_response, init\_user\_agents
* Nettoyage :

Les données récupérées doivent être nettoyer :

* + remove\_white\_spaces, remove\_white\_spaces\_split
* Récupération de la donnée :

Les données sont récupérées grâce a toutes les méthodes get. Les « extracts\_pages » permettent de gérer le défilement des pages.

* Classement :

Toutes les données sont ensuite stockées dans un tableau contenant un dictionnaire de joueur.