

Abstract

Social media is a medium where millions of people can connect and utilize information through the internet network. The features provided by social media have facilitated the exchange of data on the internet. Social media has covered many things. Many people have more than one social media account. Day by day the number of social media users and has reached a very large number plus everyone can have more than one account. This enormous area makes social media into a data warehouse that can be used to load information. In this study, writers need to crawl to get information from several social media as well as directly. The process of crawling uses DOM Parser to parse the content on the site and SQL to insert data into the database. After that done the assessment process on crawling data on the user name. In this study obtained 90% accuracy rate for unverified account estimation on Facebook and 80% accuracy for verified account estimation on Instagram.

Keywords: crawling, DOM, SQL, social media