

Website Pencari Produk Terlaris untuk Membantu Penjualan Pada Toko Online

Audrey Setiono¹, Silvia Rostianingsih², Anita Nathania Purbowo³

Program Studi Teknik Informatika, Fakultas Teknologi Industri, Universitas Kristen Petra

Jl. Siwalankerto 121 – 131 Surabaya 60236

Telp. (031) – 2983455, Fax. (031) – 8417658

E-mail: seraphina.audrey@gmail.com¹, silvia@petra.ac.id², anitaforpetra@gmail.com³

ABSTRAK

Perkembangan teknologi menjadi dampak positif bagi orang yang ingin membuka peluang bisnis baru dengan menjual produknya secara online. Kendalanya adalah dalam pemilihan produk yang ingin dijual, masyarakat cenderung bingung untuk menentukan produk apa yang dibutuhkan karena ketidaktahuan produk apa yang sedang laris di pasaran.

Dalam rangka mempermudah pengguna yang ingin mengetahui produk terlaris maka dilakukan pembuatan website dengan memanfaatkan metode Web Scraping sebagai penunjang untuk mengumpulkan informasi mengenai produk terlaris yang terdapat dalam marketplace di antaranya Tokopedia, Shopee dan Bukalapak. Informasi tersebut berupa nama barang, harga barang, jumlah produk terjual, gambar produk dan asal situs. Secara keseluruhan website dapat membantu user dalam mencari produk terlaris sesuai dengan kategori yang dipilih oleh user.

Kata Kunci: Jurnal, Produk Terlaris, *Web Scraping*, *Online Shop*, *Website*.

ABSTRACT

Development of technologies become a positive impact for people who want to open new business opportunities by selling their products online. The obstacle is in the selection of products that will be sold, people tend to have a hard time deciding the product that they needed because they don't know what products are on demand in the market.

In order to facilitate users who want to know the current best-selling products, website by using Web Scraping method is made as a support to collect information about the best-selling products contained in the marketplace of Tokopedia, Shopee and Bukalapak. The informations include name of the goods, price of the goods, number of products sold, product image and origin of the site. In conclusion, the website can help users finding the best selling products according to the categories selected.

Keywords: *Journal, Best Selling Product, Web Scraping, Online Shop, Website.*

1. PENDAHULUAN

Berjualan dapat lebih mudah dilakukan dengan menjual produk yang dimiliki secara online. Dengan berkembangnya teknologi, telah diciptakan marketplace yang memungkinkan menjual produk secara online. Contoh marketplace yang telah ada tersebut adalah tokopedia, shopee, maupun bukalapak. Sudah menjadi suatu hal yang memberikan dampak baik kepada pebisnis yang ingin menjual produknya secara online dengan adanya kemudahan yang ditawarkan layanan internet. Saat ini bisnis online banyak dilakukan sehingga pebisnis juga harus tetap dapat mengikuti tren yang ada dalam melakukan penjualan barang. Tetapi kendala yang dihadapi oleh para pebisnis yaitu pebisnis toko online bingung dalam pemilihan produk yang akan dijual maupun dibeli karena saat melakukan transaksi, terkadang

pebisnis tidak mengetahui produk apa yang sedang laris di pasaran.

Idealnya dengan adanya suatu website yang menampilkan data produk yang sedang laris sesuai dengan kategori yang dipilih maka pebisnis dapat mengikuti tren yang ada dan dapat membuat keputusan produk seperti apa yang sebaiknya dijual maupun dibeli. Data produk yang perlu ditampilkan itu berupa nama produk, harga produk, jumlah produk yang terjual, deskripsi produk dan situs asal produk tersebut. Sehingga selain pebisnis yang ingin mengetahui produk apa yang sebaiknya dijual, pebisnis juga dapat melihat darimana situs asal produk tersebut dijual.

Oleh karena itu dibutuhkan website yang dapat membantu pebisnis toko online untuk dapat melihat apakah produk yang sedang laris terjual dan pebisnis toko online dapat menemukan tren yang sesuai yang dapat dijadikan peluang bisnis yaitu menentukan produk seperti apa yang akan dijual

2. TINJAUAN PUSTAKA

2.1 Web Scraping

Web Scraping adalah teknik untuk mendapatkan informasi dari *website* secara otomatis tanpa harus menyalinnya secara manual. Tujuan dari *web scaper* adalah untuk mencari informasi tertentu dan kemudian mengumpulkannya dalam *web* yang baru.

Web scraping berfokus dalam mendapatkan data dengan cara pengambilan dan ekstraksi[3].

Web scraping (panen web) adalah pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya berupa halaman-halaman web dalam bahasa markup seperti HTML atau XHTML, dan menganalisis dokumen tersebut untuk diambil data tertentu dari halaman tersebut. Istilah gampangnya yaitu pengambilan konten atau sebagian data dari suatu situs web. *Web scraping* mempunyai banyak kegunaan dan sangat membantu masalah pengambilan data. Salah satunya yaitu untuk konten artikel dimana isi kontennya langsung diambil dari situs lain misal konten dari wikipedia. Contohnya pada pesonainformatika.com misalkan kata kunci yang dicari tidak ditemukan, maka otomatis akan melakukan scraping dari situs wikipedia berdasarkan kata kunci atau keyword yang dicari[5].

2.2 E-Marketplace

E-Marketplace merupakan pasar virtual dimana penjual dan pembeli bertemu dan melakukan berbagai jenis transaksi. Di sini orang melakukan proses transaksi dengan pertukaran barang maupun jasa untuk menghasilkan uang. Setiap individu juga dapat membuka pasar yang menjual produk atau layanan secara online. Pasar elektronik terhubung ke penjual dan pembeli melalui Internet atau ke mitra dalam organisasi, intranet[6].

2.3 PHP Simple HTML DOM Parser

Ide asli didapat dari HTML Parser for PHP 4 yang dibuat oleh Jose Solorzano. Dimana pada halaman php-html.sourceforge.net

dijelaskan bahwa HTML Parser for PHP 4 adalah open source HTML parser yang ditulis menggunakan bahasa PHP. Dimana parser ini dilengkapi dengan alat yang mengkonversi HTML ke teks, sebagai contohnya. Parser ini dirancang untuk kecepatan dan fleksibilitas. Parser ini tidak membuat model objek yang baru, tetapi tidak mencegah apabila menggunakannya untuk membuat model objek jika diperlukan. Cara kerjanya adalah dengan meminta elemen atau node berikutnya di dokumen sesuai dengan kebutuhan[2].

2.4 Selenium Webdriver

Selenium Webdriver adalah kumpulan API open source yang digunakan untuk mengotomatisasi pengujian aplikasi web. Selenium Webdriver adalah platform independen karena kode yang sama dapat digunakan pada Sistem Operasi yang berbeda seperti Microsoft Windows, Apple OS dan Linux. Selenium Webdriver juga memiliki dukungan yang lebih baik karena halaman web dinamis seperti Ajax, di mana elemen-elemen halaman web dapat berubah tanpa halaman itu sendiri sedang dimuat ulang[1].

2.5 ChromeDriver

WebDriver adalah perangkat open source untuk pengujian otomatis dari aplikasi website di banyak browser. WebDriver menyediakan kemampuan untuk menavigasi ke halaman web, input pengguna, eksekusi JavaScript, dan banyak lagi. ChromeDriver adalah server mandiri yang menerapkan protokol WebDriver untuk Chromium. ChromeDriver adalah tool terpisah yang digunakan WebDriver untuk mengontrol Chrome. ChromeDriver dikelola oleh tim Chromium dengan bantuan dari kontributor WebDriver[4].

3. ANALISIS DAN DESAIN SISTEM

Alur *website* bekerja dimulai dengan jenis pengguna yang mana yang memasuki website, apabila pengguna merupakan admin, maka admin diarahkan ke halaman untuk mengolah data kategori maupun subkategori. Admin juga dapat mengatur pada pukul berapa download data dari Tokopedia, Shopee, dan Bukalapak akan dimulai.

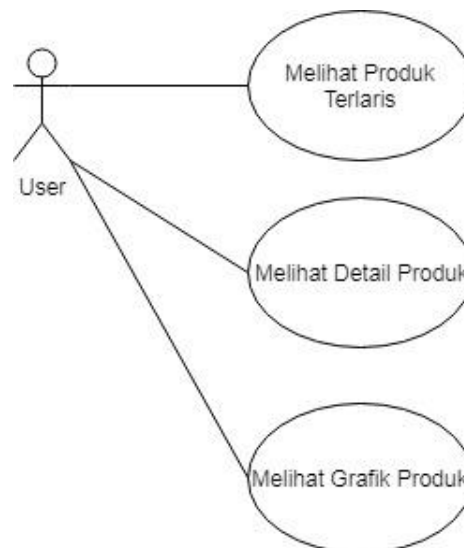
Apabila pengguna biasa masuk ke halaman utama *website*, maka pengguna akan memilih kategori. Kemudian *website* menampilkan produk terlaris sesuai dengan kategori yang dipilih. Pengguna biasa juga dapat melihat detail dari sebuah produk yaitu nama produk, harga produk, berapa produk yang terjual, link url produk, dan grafik terjualnya produk maupun harga produk.

Algoritma *web scraping* dalam *website* yang dibuat pertamanya adalah dengan mencari data yang diperlukan dalam sebuah halaman html. Setelah menentukan elemen atau tag yang akan diambil, data kemudian disimpan dalam variabel. Langkah selanjutnya adalah mengolah data menjadi informasi yang kemudian dapat ditampilkan kepada user melalui website yang telah dibangun.

3.1 Use Case Diagram

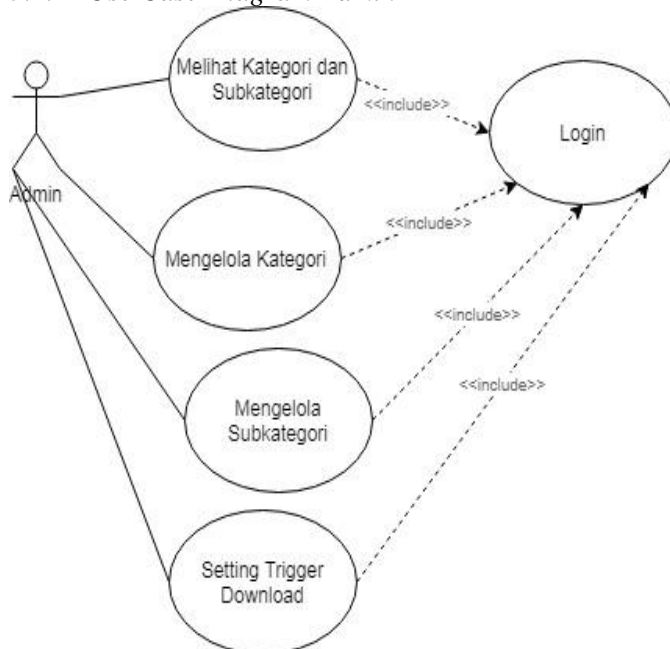
3.1.1 Use Case Diagram User

Gambar 1 merupakan Use Case Diagram User. Use Case Diagram User menjelaskan bahwa user yang menggunakan website dapat melihat produk-produk terlaris sesuai dengan kategori yang dipilih yang ditampilkan dari Tokopedia, Bukalapak, dan Shopee. Tanpa harus login, user dapat melihat detail dari sebuah produk apabila ingin mengetahui lebih rinci mengenai suatu produk. User juga dapat melihat riwayat suatu produk yang ditampilkan berupa grafik.



Gambar 1. Use Case Diagram User.

3.1.2 Use Case Diagram Admin

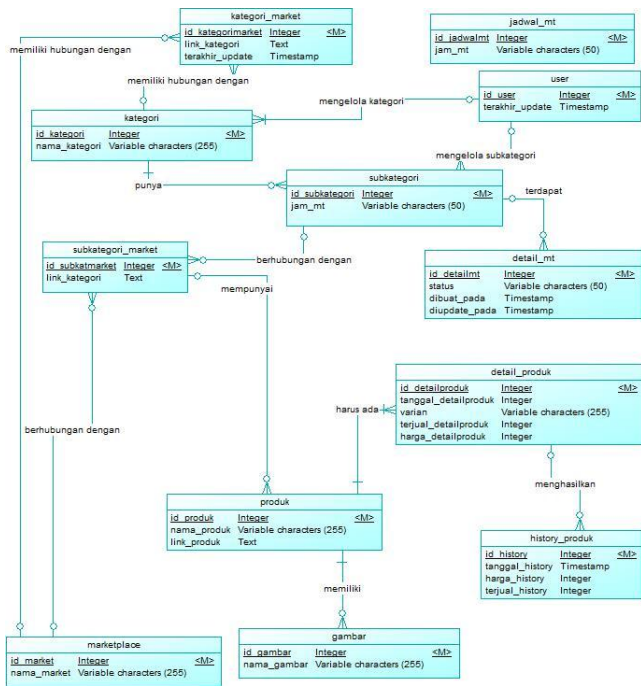


Gambar 2. Use Case Diagram Admin.

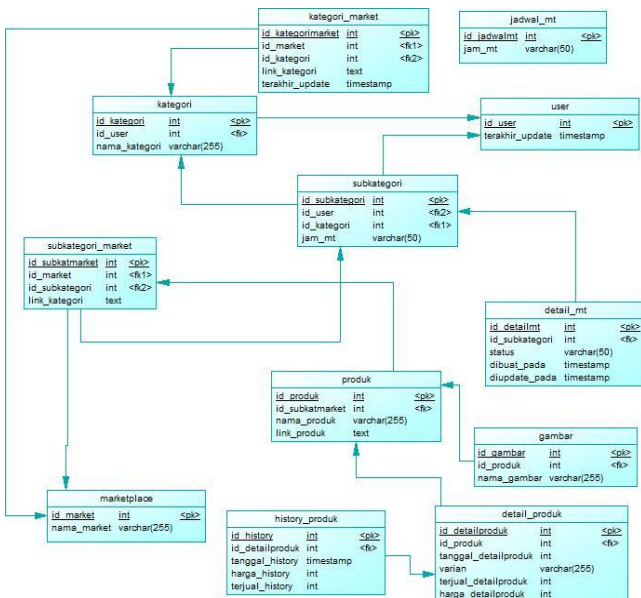
Gambar 2 merupakan Use Case Diagram Admin. Use Case Diagram Admin menjelaskan bahwa admin yang menggunakan website dapat melihat table kategori dan subkategori. Admin juga dapat melihat mengelola (*create, update, delete*) kategori dan subkategori. Admin juga diberi akses untuk mengatur pada jam berapa *website* akan melakukan proses *download* untuk memperbarui isi *database*.

3.2 Desain Entity Relationship Diagram (ERD)

ERD memiliki 2 macam jenis, yaitu *conceptual data model* dan *physical data model*. CDM berguna untuk dapat menunjukkan relasi antar tabel yang ada. Sedangkan PDM berguna untuk menunjukkan struktu tabel secara fisik. Gambar ERD jenis CDM dapat dilihat pada Gambar 3. Gambar ERD jenis PDM dapat dilihat pada Gambar 4.



Gambar 3. Conceptual Data Model.



Gambar 4. Physical Data Model.

4. IMPLEMENTASI SISTEM

4.1 Implementasi Website Untuk Koneksi Database

Koneksi ke *database* diperlukan untuk dapat mengakses, menambah dan mengelola data yang tersimpan dalam *database*. Koneksi *database* menghubungkan antara program dengan database MySQL.

4.2 Implementasi Program

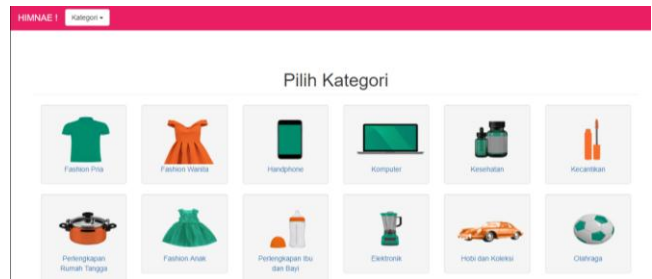
Implementasi program berupa potongan / segmen *source code* aplikasi dalam bahasa pemrograman Python, PHP dan HTML dengan *database* menggunakan *database* MySQL. Implementasi ini berguna agar *user* dapat menggunakan website sesuai dengan hak akses yang dimiliki, yaitu admin atau *user* biasa.

5. PENGUJIAN SISTEM

5.1 Pengujian User

5.1.1 Pengujian Fitur Menampilkan Kategori

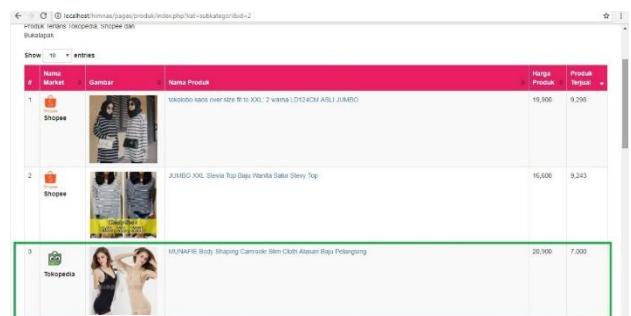
Dalam halaman utama *user* ditampilkan semua kategori yang terdapat dalam database sehingga *user* dapat memilih kategori mana yang ingin dilihat produk terlarisnya. Tampilan halaman utama *user* dapat dilihat pada Gambar 5. Dari tampilan Gambar 5 dapat dilihat bahwa kategori berhasil ditarik dari database.



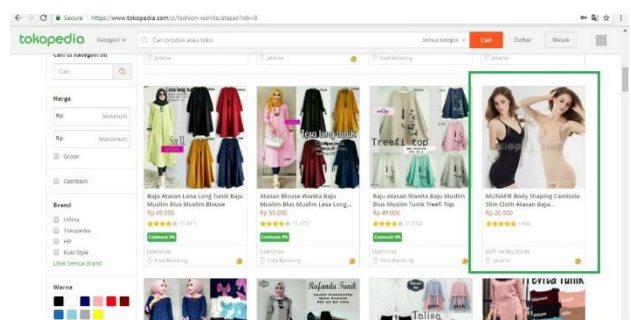
Gambar 5. Tampilan Halaman Utama User

5.1.2 Pengujian Fitur Menampilkan Produk Terlaris

Ketika *user* memilih kategori yang terdapat pada Gambar 5, *user* akan ditampilkan gambar produk-produk terlaris yang ada pada Tokopedia, Shopee dan Bukalapak. Tampilan halaman produk terlaris Tokopedia dapat dilihat pada Gambar 6. Dari Gambar 6, dapat dilihat bahwa produk Tokopedia yang ditampilkan pada website yang dibuat sama dengan produk yang ada terdapat dalam website Tokopedia sendiri yang terdapat pada Gambar 7. Dari Gambar 8, ditampilkan pada *website* yang dibuat produk Shopee sama dengan produk yang ada dari *website* Shopee yang terdapat pada Gambar 9. Pada Gambar 10, ditampilkan produk Bukalapak pada *website* yang dibuat dan produk yang sama terdapat pada *website* Bukalapak yang ada pada Gambar 11. Tampilan-tampilan ini menunjukkan bahwa produk-produk terlaris dari Tokopedia, Shopee dan Bukalapak berhasil diambil.



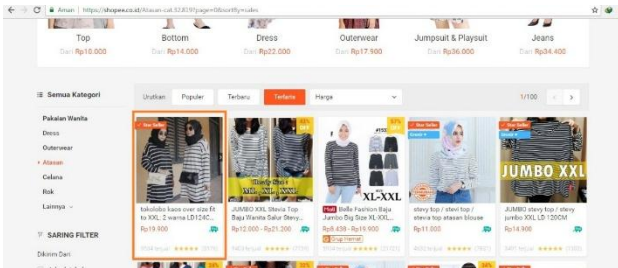
Gambar 6. Halaman Produk Terlaris Tokopedia



Gambar 7. Halaman Produk Terlaris dari Tokopedia



Gambar 8. Halaman Produk Terlaris Shopee



Gambar 9. Halaman Produk Terlaris dari Shopee



Gambar 10. Halaman Produk Terlaris Bukalapak

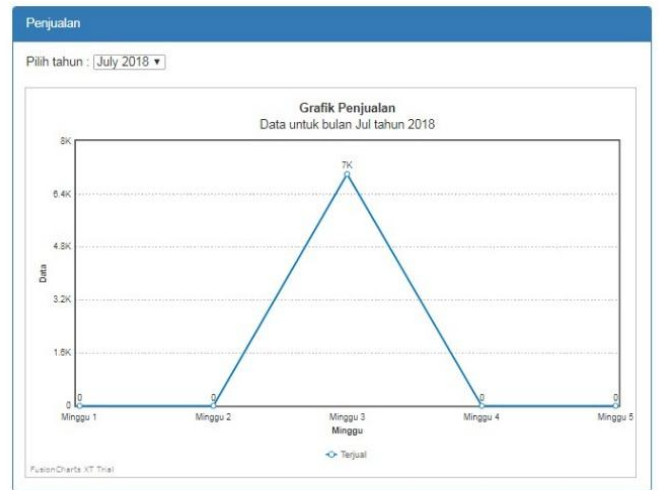


Gambar 11. Halaman Produk Terlaris dari Bukalapak

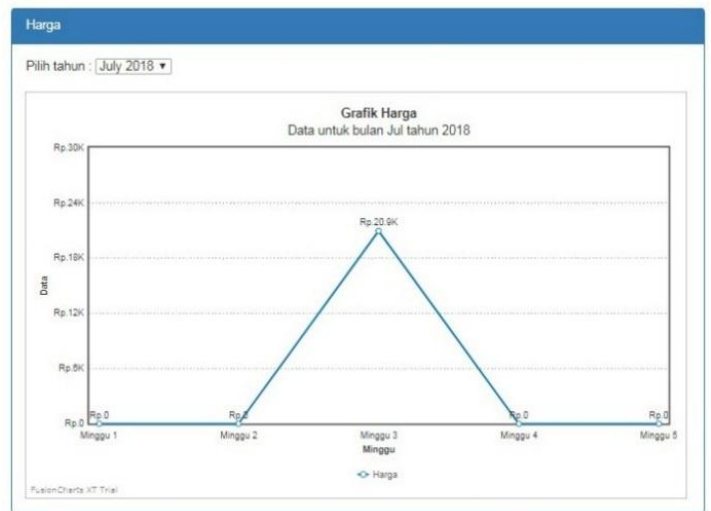
5.1.3 Pengujian Fitur Menampilkan Detail Produk
 Ketika user memilih salah satu produk terlaris yang terdapat pada halaman produk, user akan ditampilkan detail dari suatu produk terlaris yang melingkupi nama produk yang dapat mengarahkan ke link darimana produk itu berasal, jumlah terjual produk dan harga produk. Tampilan halaman detail produk dapat dilihat pada Gambar 12. Dalam halaman detail produk juga ditampilkan grafik terjual produk dan grafik harga produk yang dapat dilihat pada Gambar 13 dan Gambar 14.



Gambar 12. Tampilan Halaman Detail Produk



Gambar 13. Grafik Penjualan Produk



Gambar 14. Grafik Harga Produk

5.2 Pengujian Waktu

Dalam proses pengambilan data produk (download halaman) dan data dimasukkan ke database maka waktu yang diperlukan untuk 6 kategori adalah 1 jam 16 menit 41 detik, dengan bandwidth download sebesar 10Mbps dan bandwidth upload sebesar 1Mbps dan dengan menggunakan processor Intel Core i7 dengan RAM sebesar 8GB. Sesuai dengan waktu yang terdapat pada Gambar 15, website pada halaman setting menunjukkan bahwa pada kolom "Diupdate Terakhir" jam 12 siang proses dimulai. Apabila melakukan refresh pada halaman setting ketika produk-produk terlaris telah berhasil diambil maka jam akan berubah sesuai pada Gambar 16 dimana pada kolom "Diupdate Terakhir" berubah menjadi jam 1 siang. Semua proses pengambilan data produk (download halaman) berjalan secara background untuk proses download per 6 kategori dapat dilihat pada Gambar 17.

Setting

#	Nama Subkategori	Jenis	Diupdate Terakhir
1	Fashion Wanita	Kategori	2018-07-12 12:00:00
2	Atasan	Sub Kategori Fashion Wanita	2018-07-12 12:00:00
3	Celana	Sub Kategori Fashion Wanita	2018-07-12 12:00:00
4	Rok	Sub Kategori Fashion Wanita	2018-07-12 12:00:00
5	Tas	Sub Kategori Fashion Wanita	2018-07-12 12:00:00
6	Sepatu	Sub Kategori Fashion Wanita	2018-07-12 12:00:00

Gambar 15. Halaman *Setting Download Jam Mulai*

Setting

#	Nama Subkategori	Jenis	Diupdate Terakhir
1	Fashion Wanita	Kategori	2018-07-12 13:14:51
2	Atasan	Sub Kategori Fashion Wanita	2018-07-12 13:15:41
3	Celana	Sub Kategori Fashion Wanita	2018-07-12 13:16:35
4	Rok	Sub Kategori Fashion Wanita	2018-07-12 13:15:12
5	Tas	Sub Kategori Fashion Wanita	2018-07-12 13:15:48
6	Sepatu	Sub Kategori Fashion Wanita	2018-07-12 13:16:41

Gambar 16. Halaman *Setting Download Jam Selesai*

Process Name	User	PID	Private Bytes	Working Set	Session ID
php.exe	Audrey	00	7.364 K	CLI	
php.exe	Audrey	00	7.364 K	CLI	
php.exe	Audrey	00	7.352 K	CLI	
php.exe	Audrey	00	27.816 K	CLI	
php.exe	Audrey	00	28.060 K	CLI	
php.exe	Audrey	00	7.376 K	CLI	

Gambar 17. *Background Process Download*

6. KESIMPULAN DAN SARAN

6.1. Kesimpulan

Dari hasil pengujian sistem yang telah dilakukan, dapat diambil beberapa kesimpulan antara lain :

- Dibutuhkan scheduler dalam hal ini menggunakan task scheduler/cronjob untuk dapat melakukan trigger kepada website dengan cara melakukan pengecekan jam agar proses web scraping berjalan.

- Waktu yang dibutuhkan untuk pengambilan dan pemasukkan data produk terlaris ke dalam database untuk 6 kategori/subkategori adalah 1 jam 16 menit 41 detik, dengan *bandwidth download* sebesar 10Mbps dan *bandwidth upload* sebesar 1Mbps dan dengan menggunakan *processor* Intel Core i7 dengan RAM sebesar 8GB.

- Website menampilkan data produk terlaris beserta detail produk yang sama sesuai dengan yang ada pada marketplace Tokopedia, Shopee dan Bukalapak.

- Proses pengambilan memperberat server karena untuk mengambil data produk terlaris perlu mengunduh terlebih dahulu halaman produk terlaris dan hal tersebut dilakukan dengan cara membuka halaman baru Chrome untuk dapat diunduh secara background menggunakan chromedriver.

Saran yang dapat diberikan untuk penyempurnaan dan pengembangan program lebih lanjut antara lain :

- Waktu untuk pengambilan dari website marketplace Tokopedia, Shopee dan Bukalapak dipersingkat.

- Website dapat mengambil semua produk dari website marketplace Tokopedia, Shopee dan Bukalapak dengan lebih ringan/tidak terlalu memperberat server yang melakukan pengambilan data produk terlaris.

- Tampilan dibuat lebih bagus lagi agar lebih mudah dalam penggunaan user.

7. DAFTAR PUSTAKA

- [1] Bennett, Coleman.2018. Definition of 'Selenium Web Driver'. Retrieved Juni 9, 2018, from <https://economictimes.indiatimes.com/definition/selenium-web-driver>
- [2] Chen.2014. PHP Simple HTML DOM Parser. Retrieved April 23, 2018, from <http://simplehtmldom.sourceforge.net/>.
- [3] Josi, A., Abdillah, L. A., & Suryayusra.2014. Penerapan Teknik Web Scraping pada Mesin Pencari Artikel Ilmiah. Jurnal Sistem Informasi. Retrieved November 23, 2017, from <https://arxiv.org/ftp/arxiv/papers/1410/1410.5777>.
- [4] Khachatryan, Artur.2018. ChromeDriver. Retrieved Juni 9, 2018, from <http://chromedriver.chromium.org/>
- [5] Syabab, Akhi.2018. Apa itu web scaping?. Retrieved Januari 24, 2018, from <http://pesonainformatika.com/other-notes/apa-itu-web-scraping/>.
- [6] Turban, E.2012. Electronic Commerce (8th ed). Pearson. Retrieved April 23, 2018, from <https://sabraz.files.wordpress.com/2017/08/e-commerce-turban.pdf>.