

## BAB 1. PENDAHULUAN

### 1. 1. Latar Belakang

Pada era milenial seperti saat ini, aktivitas kegiatan sehari-hari manusia tidak lepas dari teknologi, mulai dari berinteraksi atau berkomunikasi, berpolitik, berniaga dan lainnya. Dalam kegiatan berniaga masyarakat sudah tidak lepas dari teknologi informasi, baik televisi maupun website (toko online).

Setiap hari ribuan transaksi penjualan berlangsung melalui e-commerce, hal ini dibuktikan dengan jumlah pembeli online (online shopper) yang mencapai angka 11,9% dari populasi penduduk yang ada di Indonesia (kompas.com, 2018). Dengan meningkatnya jumlah konsumen toko online, hal ini tentunya juga akan mendorong para penjual untuk melakukan promosi barang salah satunya *diskon*. *Diskon* adalah pengurangan harga secara langsung terhadap produk tertentu pada periode waktu tertentu, serta pembelian potongan harga yang diberikan penjual kepada pembeli (Putra, Kumadji, & Yulianto, 2016). Namun, dengan adanya banyak diskon yang terdapat pada toko online membuat konsumen bingung mencari barang kebutuhannya yang mendapatkan diskon di toko online. Masalah lainnya juga konsumen tidak dapat mengetahui toko online mana yang memiliki diskon paling besar.

*Web Scraping* adalah proses 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 untuk digunakan bagi kepentingan lain (Josi, Abdillah, & Suryayusra, 2014).

*Web scraping* sering dikenal sebagai *screen scraping*. *Web Scraping* tidak dapat dimasukkan dalam bidang *data mining* karena *data mining* menyiratkan upaya untuk memahami pola semantik atau tren dari sejumlah besar data yang telah diperoleh. Aplikasi *web scraping* (juga disebut *intelligent, automated, or autonomous agents*) hanya fokus pada cara memperoleh data melalui pengambilan dan ekstraksi data dengan ukuran data yang bervariasi (Josi dkk., 2014).

Manfaat dari *web scraping* ialah agar informasi yang dikeruk lebih terfokus sehingga memudahkan dalam melakukan pencarian sesuatu, adapun cara mengembangkan teknik *web scraping* yaitu dengan cara pertama Pembuat program mempelajari dokumen HTML dari *website* yang akan diambil informasinya untuk di tag HTML tujuannya ialah untuk mengagit informasi yang diambil (*Create Scraping Template*), setelah itu pembuat program mempelajari teknik navigasi pada *website* yang akan diambil informasinya untuk ditirukan pada aplikasi *web scraping* yang akan dibuat (*Explore Site Navigation*), kemudian aplikasi *web scraping* akan mengotomatisasi informasi yang didapat dari *website* yang telah ditentukan (*Automate Navigation and Extraction*), informasi yang didapat tersebut akan disimpan ke dalam tabel basisdata (*Extracted Data and Package History*) (Josi et al., 2014).

*Xpath* (XML Path Language), adalah bahasa *query* untuk memilih node dari dokumen XML. *Xpath* dapat digunakan untuk menghitung nilai (misalnya, string, angka, atau nilai Boolean) dari isi dokumen XML. *Xpath* didefinisikan oleh *World Wide Web Consortium* (W3C) (Jumadi, 2016). Bahasa *Xpath* didasarkan pada representasi pohon dokumen XML, dan menyediakan kemampuan untuk menavigasi di sekitar pohon XML, memilih node dengan berbagai kriteria.

Berdasarkan pemaparan di atas maka dibangunlah sebuah aplikasi web scraping berbasis desktop untuk menampung informasi dari berbagai toko online yang akan mengumpulkan informasi mengenai promo atau diskon suatu produk dari beberapa toko online sesuai kebutuhan konsumen. Sehingga konsumen dapat mengetahui diskon terbesar dari beberapa toko online. Dalam proses pencarian promo barang, konsumen dapat memilih ingin mencari promo barang pada beberapa toko online sesuai minat konsumen dan untuk pengembangan aplikasi tersebut diperlukan teknik Web Scraping menggunakan metode *Xpath Selector*.

## **1. 2. Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan, yang menjadi pokok permasalahan dalam hal ini adalah:

1. Bagaimana membuat crawler untuk pencarian barang.
2. Bagaimana membuat bot yang dapat melakukan Web Scraping pada toko online.
3. Bagaimana mendapatkan informasi secara jelas tentang produk yang sedang promo.
4. Bagaimana membuat interface untuk menampilkan informasi hasil Scraping.

## **1. 3. Batasan Masalah**

Setelah dilakukan penelitian lebih lanjut, maka dapat ditarik beberapa batasan masalah pada tugas akhir ini adalah sebagai berikut:

1. Proses Scraping dilakukan pada 2 toko online yaitu klikklik.com dan erafone.com
2. Untuk penyimpanan dan pengelolaan data menggunakan mysql dan django web framework
3. Proses scraping berjalan secara backend melalui command.
4. UI hanya berfungsi untuk menampilkan data hasil scraping

## **1. 4. Tujuan**

Penelitian ini bertujuan untuk membantu konsumen dalam melakukan pencarian promo barang yang dibutuhkan dengan menggunakan teknik *web scraping*. Informasi yang telah didapat kemudian dikumpulkan untuk selanjutnya diolah dan ditampilkan ke konsumen berupa data promo barang yang berurutan.

## **1. 5. Manfaat**

### **1. 5. 1. Manfaat bagi Penulis**

Manfaat yang akan diperoleh oleh penulis ialah penulis dapat menerapkan teknik *web scraping* untuk mencari promo barang sesuai kebutuhan konsumen dengan menggunakan metode XPath Selector.

### **1. 5. 2. Manfaat bagi Konsumen**

Manfaat yang akan diperoleh oleh konsumen apabila aplikasi web scraping telah dibuat berupa proses pencarian promo barang sesuai kebutuhan konsumen akan semakin mudah, konsumen dapat membandingkan promo barang yang sama pada toko online yang berbeda sehingga konsumen dapat lebih selektif dalam memilih barang kebutuhannya.