

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Dengan perkembangan teknologi informasi saat ini, telah menciptakan jenis-jenis dan peluang-peluang bisnis yang baru di mana transaksi-transaksi bisnis makin banyak dilakukan secara elektronika. Sehubungan dengan perkembangan teknologi informasi tersebut memungkinkan setiap orang dengan mudah melakukan perbuatan hukum seperti misalnya melakukan jual-beli. Perkembangan internet memang cepat dan memberi pengaruh signifikan dalam segala aspek kehidupan kita. Internet membantu kita sehingga dapat berinteraksi, berkomunikasi, bahkan melakukan perdagangan dengan orang dari segala penjuru dunia dengan murah, cepat dan mudah. beberapa tahun terakhir ini dengan begitu merembaknya media internet menyebabkan banyaknya perusahaan yang mulai mencoba menawarkan berbagai macam produknya dengan menggunakan media ini. Dan salah satu manfaat dari keberadaan internet adalah sebagai media promosi suatu produk. Suatu produk yang dionlinekan melalui internet dapat membawa keuntungan besar bagi pengusaha karena produknya di kenal di seluruh dunia.

Penggunaan internet tidak hanya terbatas pada pemanfaatan informasi yang dapat diakses melalui media ini, melainkan juga dapat digunakan sebagai sarana untuk melakukan transaksi perdagangan yang sekarang di Indonesia telah mulai diperkenalkan melalui beberapa seminar dan telah mulai penggunaannya oleh beberapa perusahaan yaitu electronic commerce atau yang lebih dikenal dengan E-Commerce, yang merupakan bentuk perdagangan secara elektronik melalui media

internet Dalam Tugas Akhir kali ini akan membuat pengembangan dari sistem berbasis website dengan penerapan webscraping menggunakan ekstrasi data yang cepat .

Pembelian produk barang yang dilakukan secara offline biasanya tidak menggunakan sistem sehingga pemilihan produk barang hanya didasari dari bagian yang berwenang sesuai dengan yang dilihat (subjektif). Pemilihan yang subjektif sangat memungkinkan terjadi kesalahan atau tidak sesuai antara nilai harga saat ini dengan nilai harga yang telah diperbaharui sehingga dapat mengurangi kinerja dari pemilihan barang secara offline tersebut.

Web Scraping (WS) adalah sistem komputer yang mengolah data website menjadi informasi untuk mengambil data-data yang terpenting di dalam website dari masalah semi terstruktur yang spesifik. Sistem informasi sangat penting untuk mendukung proses pengambilan data scraping [3]. Maka dari itu diusulkan penelitian untuk pembuatan web scraping dalam pengambilan produk menggunakan salah satu metode Web Scraping yaitu *Vision Based Page Segmentation* (VIPS) yang dikombinasikan dengan algoritma *Cosine Similarity* dengan metode *TF-IDF* yang diharapkan dapat menjadi acuan data untuk rekomendasi pengguna melakukan penilaian secara baik dan objektif untuk mengambil produk yang akan digunakan untuk perbandingan harga yang tepat dalam pembelian produk yang sesuai.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang sudah diuraikan di atas maka dapat diambil rumusan masalah pada penelitian ini diantaranya:

1. Bagaimana prinsip kerja algoritma *Cosine Similarity* dalam melakukan *scraping* data pada string proses ?
2. Bagaimana prinsip kerja algoritma *Vision Based Page Segmentataion* dalam melakukan *scraping* data atau pengambilan data tag *html* ?
3. Bagaimana implementasi ekstrasi pada data dom *html* dan data *json* yang telah di dapat menjadi *txt/html* ?
4. Bagaimana implementasi proses *afiliasi* dalam *scraping* produk ?

1.3. Tujuan dan Manfaat

1.3.1. Tujuan Penelitian

Tujuan dari penelitian yang dilakukan oleh penulis ini adalah sebagai berikut:

1. Mengetahui prinsi kerja algoritma *Cosine Similarity* dalam melakukan *scraping* data padaa string proses.
2. Mengetahui prinsip kerja algoritma *Vision Based Page Segmentation* dalam melakukan *scraping* data atau pengambilan data tag *html*.
3. Mengetahui implementasi ekstrasi pada data dom *html* dan data *json* menjadi *txt/html* data .
4. Mengetahui implementasi proses *afiliasi* dalam *scraping* produk.

1.3.2. Manfaat Penelitian

Berdasarkan perumusan dan tujuan penelitian yang sudah diuraikan diatas, maka penulis menguraikan berapa hal yang menjadi manfaat dari penelitian ini adalah sebagai berikut:

1. Dapat memahami bagaimana pembobotan string query dalam pemilihan produk yang tepat pada website marketplace.

2. Dapat memahami bagaimana merancang sistem web scraping dengan menggunakan metode *Cosine Similarity* dan *Vision Based Page Segmentation* (VIPS).
3. Dapat memahami kinerja dari sistem web scraping menggunakan metode *Cosine Similarity* dan *Vision Based Page Segmentation*.
4. Menjadikan hasil sistem produk sebagai proses *afiliasi* data untuk mendapatkan komisi sebagaimana seharusnya.

1.4. Batasan Masalah

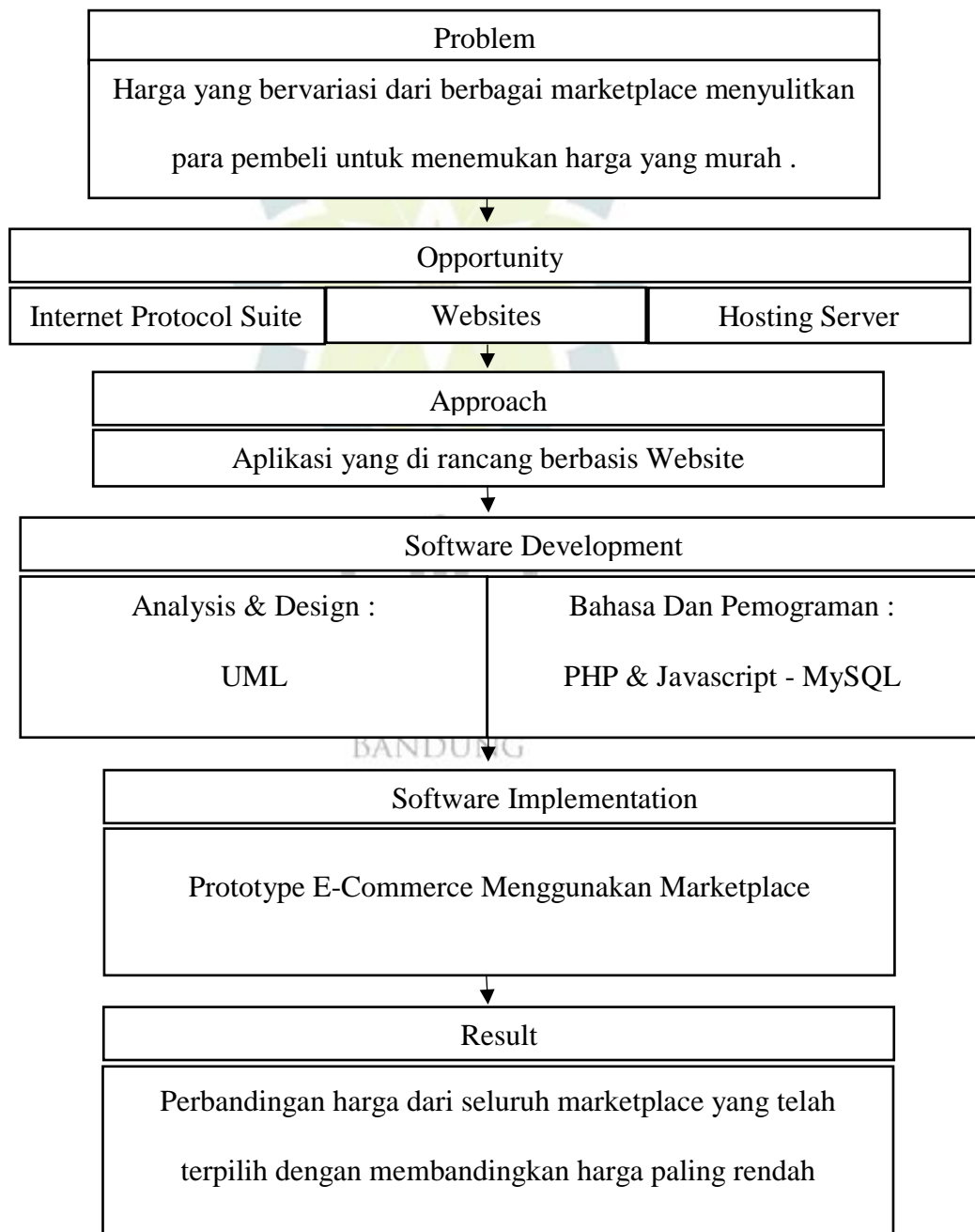
Pada penelitian ini, diperlukan beberapa batasan masalah agar sesuai dengan apa yang akan diteliti dan direncanakan sebelumnya. Adapun batasan masalah dalam penelitian ini sebagai berikut:

1. Sistem ini hanya dapat melakukan scraping pada website Marketplace Indonesia terpilih.
2. Data awal dari penelitian ini diambil dari data milik P.T Tokopedia Indonesia dari situs www.tokopedia.com sehingga data akhir hasil produk akan disesuaikan dengan data yang ada di www.tokopedia.com.
3. Metode yang digunakan adalah gabungan dari metode *Cosine Similarity* dan *Vision Based Page Segmentation* (VIPS).
4. Data yang dipakai berupa data *API* dan *close session* dalam selang waktu 30 hari.
5. *Output* yang di hasilkan yaitu berupa daftar produk dari setiap marketplace yang terpilih.

6. Pembuatan aplikasi web scraping ini menggunakan Bahasa PHP dengan teknik *OOP*.

1.5. Kerangka Pemikiran

Penelitian ini dilakukan untuk menyelesaikan permasalahan dalam proses penerapan web scraping marketplace Indonesia. Kerangka pemikiran penelitian ini dapat dilihat pada gambar 1.1.



Gambar 1.1 Kerangka Pemikiran Web Scraping Marketplace

Gambar 1.1 diatas menjelaskan keseluruhan proses penelitian. Sesuai analisis permasalahan diatas maka dibangun suatu sistem scraping menggunakan algoritma Vision Based Page Segmentation berbasis Website dalam hal perancangan data marketplace dan data produk marketplace Indonesia dengan menggunakan metode pengembangan prototype agar proses pembangunan sistem dapat dilakukan perbaikan secara langsung. Proses perhitungan aritmatik sistem dibantu dengan algoritma *Cosine Similarity* dan *Vision Based Page Segmentation* yang tepat untuk kasus dengan banyak alternatif karena algoritma VIPS ini akan menghasilkan rekomendasi berupa urutan pilihan DOM (Document Object Model) dengan menggunakan css selector sebagai alat atau tools untuk penyortiran data dari data yang akan di ambil dan di ekstraksi menjadi data yang dapat di tampilkan yaitu berupa data json , selanjutnya data json akan di tampilkan dalam website berbasis website dengan model data berupa text atau txt . Sistem ini dibangun dengan mempertimbangkan nilai dari kriteria yang disajikan dalam bentuk matriks yang akan dinormalisasi dan diberi bobot menggunakan metode *Cosine Similarity* dan untuk pencarian data dalam Website yang akan di bangun akan menampilkan data yang sesuai dengan pembobotan nilai dari query produk dan di cari persamaannya menggunakan TF-IDF , dengan kedua metode algoritma maka di nyatakan bahwa metode ini sangatlah penting untuk membantu jalannya aplikasi dari pencarian data produk maupun pencarian data dari marketplace yang telah di tentukan dalam aplikasi .

1.6. Metodologi Pengerjaan Tugas Akhir

Metode yang akan digunakan dalam pengembangan aplikasi ini adalah prototyping. Prototipe merupakan suatu metode dalam pengembangan sistem yang menggunakan pendekatan untuk membuat sesuatu program dengan cepat dan bertahap sehingga segera dapat dievaluasi oleh pemakai. Hal ini berbeda dengan SDLC tradisional (konvensional) yang lebih banyak menghabiskan waktu untuk menghasilkan spesifikasi yang sangat rinci sebelum pemakai dapat mengevaluasi sistem. Melihat dan mempercayai menjadi hal yang diharapkan untuk dicapai dalam prototipe. Dengan menggunakan pendekatan ini, konsumen dan tim pengembang dapat mengklarifikasi kebutuhan dan interpretasi mereka. Selain itu, prototipe membuat proses pengembangan sistem informasi menjadi lebih cepat dan lebih mudah, terutama pada keadaan kebutuhan pemakai sulit untuk diidentifikasi. [4]

1. Observasi

Melakukan pengamatan terhadap data yang diteliti, melakukan *interview* dengan pihak-pihak yang berkaitan dengan pembuatan sistem pendukung keputusan khususnya di bagian kepegawaian kampus UIN SGD Bandung.

2. Analisa data

Membuat analisa terhadap data yang sudah diperoleh dari hasil observasi yaitu menggabungkan dengan laporan survei dan kebijakan pemakai menjadi spesifikasi yang terstruktur dengan menggunakan pemodelan,

3. Perancangan aplikasi

Memahami rancangan aplikasi sesuai data yang ada dan mengimplementasikan model yang diinginkan oleh pemakai. Pemodelan

aplikasi ini berupa *Block Diagram Application*, serta perancangan *database* dengan didukung pembuatan *Use case*, *activity diagram*, dan *flowchart* (Diagram Alir), guna mempermudah dalam proses-proses selanjutnya,

4. Pembuatan program

Membuat program dan merepresentasikan hasil desain ke dalam pemrograman berdasarkan aplikasi yang sudah dirancang,

5. Evaluasi program

Menguji coba seluruh spesifikasi terstruktur dan aplikasi secara keseluruhan. Pada tahap ini, dilakukan uji coba aplikasi yang telah selesai disusun. Proses uji coba ini diperlukan untuk memastikan bahwa aplikasi yang telah dibuat sudah benar, sesuai dengan karakteristik yang ditetapkan dan tidak ada kesalahan-kesalahan yang terkandung di dalamnya.

1.7. Sistematika Penulisan

Sistematika penulisan skripsi ini dibagi menjadi beberapa bab yang masing-masing bab telah dirancang dengan suatu tujuan tertentu. Berikut penjelasan tentang masing-masing bab :

BAB I PENDAHULUAN

Menguraikan tentang latar belakang permasalahan, merumuskan inti permasalahan yang dihadapi, menentukan tujuan dan kegunaan penelitian, yang kemudian diikuti dengan pembatasan masalah, metodologi penelitian, serta sistematika penulisan.

BAB II STUDI PUSTAKA

Membahas studi literatur dari berbagai referensi baik dari penelitian jurnal yang sudah pernah dilakukan maupun jurnal yang berkaitan dengan bahasan pada penelitian yang dilakukan penulis. Dilengkapi dengan berbagai penjelasan dasar dan teori yang berhubungan dengan penelitian yang dilakukan atau dasar teori yang dapat membantu pada proses analisa masalah.

BAB III PERANCANGAN SISTEM

Membahas penganalisan masalah yang dihadapi dalam pembuatan laporan dan pembuatan perangkat lunak berdasarkan metode *Prototype*. Proses yang dilakukan meliputi perancangan arsitektur, perancangan aktivitas dan perancangan kelas.

BAB IV IMPLEMENTASI SISTEM

Membahas cara mengimplementasikan perangkat lunak yang dibuat dan cara mengoperasikannya serta kebutuhan *hardware* dan bahasa pemrograman yang digunakan serta tabel pengujian dengan menggunakan metode *Blackbox*.

BAB V PENUTUP

Bab ini berisi uraian tentang kesimpulan, usulan, solusi dan saran terhadap perangkat lunak yang akan dibangun dan dikembangkan lebih lanjut.