

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi saat ini banyak kemajuan dan memberikan inovasi, salah satunya memperoleh informasi dari situs *web*. Teknologi informasi melalui situs *web* memberikan manfaat yang memudahkan perorangan atau suatu group dalam waktu yang relative singkat mendapatkan berbagai informasi yang diinginkan (Yunita, 2020). Salah satu contoh perkembangan teknologi *Website portal* berita *Scraper* yang menyediakan kebutuhan data berupa informasi, yang dimana proses pengambilan data informasi dari situs-situs nantinya diambil untuk diamati perilaku dan persepsi untuk mendapatkan segmentasi pasar yang tepat.

Pengambilan data untuk saat ini masih banyak dilakukan secara manual, dapat ditemukan memiliki beberapa keterbatasan system yaitu lamanya mengumpulkan data sehingga memperlambat kinerja pengguna. Adapun risiko yang bisa terjadi yaitu tidak mendapatkannya segementasi pasar yang tepat. Di dalam mengatasi masalah tersebut maka diperlukan *web scarping* situs berita dengan metode menggunakan Bahasa pemograman python.

Sistem *Website portal berita Scraper* mampu menemukan hasil pencarian di beberapa situs secara bersamaan, (Deviacita et al., 2019) dengan menginputkan *keyword* mengenai topic artikel yang dicari oleh pengguna kemudian system akan memberikan *request* kepada *server* untuk mencaritautan yang berhubungan dengan *keyword* tersebut. Hasil *keyword* dari proses *scraping* yang telah di temukan *system* secara otomatis (Flores et al., 2020) mengunduh sebuah file berekstensi CSV untuk memudahkan pemindahannya dari satu program keprogram lain. Dalam proses pembuatan program tersebut akan menjelaskan melalui laporan ini

mengenai hasil kegiatan Tugas Akhir ini. Sehingga, penulis memberi judul laporan, yaitu “WEB SCRAPING SITUS BERITA MENGGUNAKAN BAHASA PEMOGRAMAN PYTHON”.

1.2 Rumusan Masalah

Bedasarkan latar belakang diatas dapat dirumuskan masalah sebagai berikut:

1. Bagaimana membangun *web scraping* situs berita untuk mengumpulkan data berita dari beberapa situs yang dibutuhkan pengguna?

1.3 Batasan Masalah

Dalam perancangan dan pembuatan *web scraping* situs berita, antara lain:

1. Pengambilan data berita hanya dari 4 situs portal berita.
2. Pengambilan data berita hanya menghasilkan judul, waktu post, dan rangkuman.

1.4 Tujuan Penelitian

Adapun tujuan dari pembuatan *web scraping* situs berita menggunakan bahasa pemograman python, antara lain:

1. *Web Scraping* situs berita, untuk dapat mengambil data dari beberapa *website* portal berita.

1.5 Manfaat Penelitian

Manfaat dari pembuatan *web scraping* situs berita menggunakan bahasa pemrograman *python*, antara lain:

1. *Web Scraping* situs berita ini dapat mengurangi keterbatasan yang menghambat pelaksanaan disaat pencarian data informasi dari penggunaannya.
2. Membantu untuk mengetahui informasi yang diberitakan dibeberapa *website* berita terkait kata kunci yang diberikan.
3. Membantu menyediakan data yang bisa digunakan untuk diamati mengetahui informasi yang sedang ramai diberitakan.

1.6 Sistematika Penulisan

Dalam penulisan laporan ini, penulis membagi dalam 5 (lima) bab. Masing–masing bab terdiri atas sub bab dengan maksud agar laporan ini dapat lebih terperinci dan mempermudah pemahaman dalam masing - masing bab. Detail sistematika laporan ini adalah antara lain sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisikan tentang uraian latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini memuat uraian tentang landasan teori yang berhubungan dengan topic tugas akhir.

BAB III METODE PENELITIAN

Bab ini memuat uraian tentang metode penelitian yang dipakai untuk menunjang pelaksanaan tuga sakhir.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memuat hasil pelaksanaan di lapangan dan memuat uraian tentang hasil permasalahan dan pemecahan tersebut dicapai dari setiap metode yang dilakukan selama tugas akhir.

BAB V PENUTUP

Bab ini berisikan uraian tentang kesimpulan dan saran yang dapat diambil dari tugas akhir yang dilaksanakan.