

BAB III

LANDASAN TEORI

3.1. Data Scraping

Data merupakan sekumpulan teks, angka, simbol yang tidak mempunyai arti. Sekumpulan data yang diolah, dikelompokkan dapat menghasilkan informasi yang mempunyai arti dan dapat dijelaskan. Informasi yang didapatkan dapat menjadi fakta ataupun dapat digunakan untuk memecahkan masalah. Pada tahap tersebut informasi sudah berubah menjadi pengetahuan [10]. Pengetahuan dibutuhkan untuk dapat memecahkan masalah atau menentukan keputusan yang akan dibuat.

Data *scraping* yang bisa disebut juga *web scraping* adalah sebuah teknik yang digunakan untuk mengumpulkan data tidak teratur berasal dari *website* [6]. Data yang tidak teratur ini akan dijadikan data yang lebih teratur. *Web scraping* ini dikenal juga dengan *web harvesting*, *web data extraction*, dan *web data mining*. Data *scraping* dijadikan pilihan untuk mendapatkan data dari internet dengan beberapa alasan, yaitu :

1. *Website* yang dituju tidak mempunyai *Application Programming Interface* (API),
2. API yang disediakan harus membayar,
3. API disediakan secara terbatas,
4. API tidak menyediakan semua data yang dibutuhkan sedangkan *website* menyediakan data tersebut.

Terdapat beberapa metode yang dapat dilakukan untuk melakukan data *scraping*, yaitu *copy-paste*, *hypertext transfer protocol* (HTTP) *programming*, *hyper text markup language* (HTML) *parsing*, *document object model* (DOM) *parsing*, *web scraping software*, dan *computer vision web-page analysers*. Metode yang paling mudah dilakukan yaitu yang pertama, *copy-paste*, tetapi metode ini juga paling rawan terjadi *error*.

3.2. Robotic Process Automation

Berdasarkan halaman pada *uipath.com*, dijelaskan *robotic process automation* (RPA) merupakan suatu teknologi perangkat yang mudah untuk dibuat, dikembangkan, dan dirawat untuk mengatur perangkat lunak robot yang bertujuan untuk menirukan interaksi manusia dengan perangkat lunak [11]. Sedangkan berdasarkan *Automation Anywhere* menjelaskan bahwa RPA memungkinkan manusia untuk membuat robot perangkat lunaknya untuk melakukan suatu proses bisnis. Selain itu, juga bisa menganggap RPA ini sebagai *digital worker*, dengan memberitahukan apa yang perlu dilakukan oleh bot, bot akan melakukannya sesuai dengan yang diberitahukan. Pengertian robot pada RPA bukan mengarah pada bentuk fisik tetapi lebih kepada perangkat lunak berbasis robot (bot).

Interaksi manusia atau proses bisnis yang dapat dilakukan oleh RPA adalah *copy-paste* data dari aplikasi ke aplikasi, membuka *website* lalu *login*, membuka *email*, memasukan data, memindahkan *file*, membuat folder, dan melakukan ekstraksi data dari *web*. Beberapa hal yang perlu diperhatikan untuk menentukan interaksi manusia atau proses bisnis yang dapat dikembangkan RPA nya, yaitu :

1. Proses bisnis yang dijelaskan dengan baik, hal ini karena robot pada RPA membutuhkan langkah yang akurat supaya dapat berjalan dengan sukses,
2. Mempunyai volume tinggi dalam melakukannya secara berulang-ulang,
3. Interaksi atau proses bisnis yang matang lebih baik dipilih, dimana mempunyai keluaran yang mudah diprediksi [7] .

RPA ini mempunyai 3 jenis robot yaitu *attended* RPA atau dikenal dengan *robotic desktop automation* (RDA) dan *unattended* RPA. *Attended* RPA tetap harus membutuhkan campur tangan manusia ketika menjalankan robot, terdapat beberapa proses bisnis yang dilakukan oleh manusia. Sedangkan *unattended* RPA tidak membutuhkan campur tangan manusia tetapi akan terjadi jika suatu hal terjadi. Contohnya ketika masuknya *e-mail*, maka akan otomatis membalasnya. Selain jenisnya, juga terdapat beberapa *tools* yang tersedia untuk mengembangkan RPA, yaitu UiPath, Automation Anywhere, dll.

Berdasarkan halaman yang ada pada website UiPath, terdapat beberapa manfaat yang didapatkan dari RPA, yaitu:

1. Merupakan komponen utama dalam transformasi digital,
2. Menghemat pembiayaan,
3. Bersifat adaptif,
4. Akurasi yang tinggi,
5. Meningkatkan kepatuhan,
6. Meningkatkan produktivitas pekerja,
7. Mempunyai nilai lebih strategis, dan
8. Pekerja yang lebih senang.

3.3. UiPath

UiPath merupakan salah satu *tools* yang dapat digunakan untuk mengembangkan RPA. UiPath didirikan pada 2005 sebagai *DeskOver* di Romania yang dipimpin oleh Daniel Dines. Lalu pada tahun 2012 menfokuskan kepada RPA, yang mana sebelumnya membuat *automation libraries* yang digunakan oleh perusahaan – perusahaan besar. Pertama kali meluncurkan UiPath Desktop Automation pada tahun 2013. UiPath selalu berkembang setiap tahunnya dengan menambahnya pembukaan kantor yang lebih banyak. Selain itu juga bekerja sama dengan BPO / *consulting firms*, sampai pada 2018 menjadi peringkat pertama RPA menurut *Forrester WaveTM* dan peringkat 14 untuk *cloud* menurut *Forbes*. Tidak berhenti sampai situ, UiPath terus berkembang dan pada 2020 mencapai pendanaan Series E.

UiPath menggunakan VB.Net sebagai kode bahasa yang digunakan untuk RPA. Selain itu UiPath hanya dapat berjalan pada platform Windows. Salah satu karakteristik UiPath yaitu lingkungan untuk mengembangkan RPA menggunakan *drag and drop activities*. *Activities* ini merupakan fitur yang disediakan oleh UiPath, seperti membuka *web*, membaca dan menuliskan pada *excel*, mengklik tombol ataupun *link*, dan data *scraping* juga telah disediakan.

UiPath sebagai *tools* untuk mengembangkan RPA mempunyai tiga komponen utama, yaitu UiPath Robot, UiPath Studio, dan UiPath Orchestrator. UiPath Studio digunakan untuk membuat *flow* dari proses bisnis atau interaksi manusia yang akan di-RPA-kan. UiPath Orchestrator adalah komponen yang

digunakan untuk memantau kerja bot yang telah dikembangkan. UiPath Robot merupakan *flow* yang telah dibuat sebelumnya untuk dijalankan pada mesin tersebut. Untuk dapat menjalankan UiPath Robot harus terkoneksi dengan UiPath Orchestrator.

3.4. Metode Haversine

Haversine merupakan metode yang menerapkan bentuk trigonometri, yang diterapkan pada bentuk bulat. Pada 1805, seorang ilmuwan membuat tabel Haversine untuk menentukan jarak antara dua titik [12]. Perhitungan yang dilakukan dengan metode ini mengabaikan pegunungan serta jurang yang ada di permukaan bumi. Meskipun begitu, metode ini dapat mengukur jarak antara dua titik dengan tepat dan akurat.

Metode ini menggunakan titik *latitude* dan *longitude* untuk dapat menentukan jarak antara dua titik. *Latitude* merupakan garis yang digunakan mengukur jarak antara selatan dan utara bumi terhadap garis ekuator. Sedangkan *longitude*, untuk mengukur jarak antara barat dan timur bumi dari garis meridian. Berikut rumus dari Metode Haversine,

$$d = 2r \sin^{-1} \left(\sqrt{\sin^2 \left(\frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left(\frac{\psi_2 - \psi_1}{2} \right)} \right)$$

Gambar 3.1. Rumus Metode Haversine [13]

dengan keterangan sebagai berikut:

d = hasil jarak

r = jari – jari bumi

ϕ_1 = titik latitude pertama

ϕ_2 = titik latitude kedua

ψ_1 = titik longitude pertama

ψ_2 = titik longitude pertama

3.5. Google Maps

Google Maps merupakan salah satu produk yang dimiliki oleh Google. Google Maps secara resmi diluncurkan enam belas tahun lalu, tepatnya pada 8 Februari 2005. Google Maps yang diluncurkan pertama kali adalah versi desktop sebagai solusi untuk membantu orang pergi dari titik A ke titik B. Lalu dua tahun setelah Google Maps diluncurkan terdapat fitur untuk mengetahui *real-time traffic* yang tersedia di kurang lebih 30 kota di Amerika. Tidak berhenti disitu, pada 2008 tersedia Google Maps untuk *mobile*.

Seiring berjalannya waktu, terdapat fitur – fitur yang ditambahkan pada Google Maps, salah satunya *Google My Business* yang diluncurkan pada 2014. *Google My Business* membantu pemilik bisnis untuk menempatkan informasi mengenai bisnis yang dimiliki, foto, serta *review* pada Google Maps. Berselang beberapa bulan terdapat *Local Guides Program* yang memperbolehkan orang – orang untuk menambahkan *review*, foto, ataupun informasi mengenai suatu tempat. Kebutuhan akan Google Maps terlihat dari banyaknya pengguna aktif bulanan yang berjumlah 1 milyar.

Google Maps dapat diakses dengan bebas melalui *web browser* dengan mengetikkan “<https://www.google.com/maps>” pada kolom URL. Halaman tersebut akan menampilkan peta yang disediakan Google Maps. Untuk melakukan pencarian perlu mengetikkan kata kunci yang akan dicari pada kolom pencarian. Google Maps akan menampilkan hasil yang dirasa sesuai dengan kata kunci yang diketikkan. Ketika pencarian dilakukan terlihat perubahan pada kolom URL, dimana terdapat tambahan untuk kata kunci.

3.6. Plus Codes

Plus Codes merupakan alamat untuk tempat yang tidak mempunyai nama [4]. Berbeda dengan alamat pada umumnya yang mempunyai nomor dan nama, Plus Codes berdasarkan *latitude* dan *longitude*, merupakan kombinasi antara huruf dan angka. Plus Codes diluncurkan pada tahun 2015 dimana kombinasi dari Plus Codes lebih pendek dibanding *latitude* dan *longitude*. Plus Codes mempunyai beberapa karakteristik, yaitu:

1. Gratis dan *open source*, tidak memerlukan biaya untuk menggunakannya,
2. Dapat diakses secara offline,
3. Mudah untuk digunakan,
4. Pendek dan sederhana,
5. Mudah diidentifikasi karena terdapat simbol “+”.

Plus Codes dibuat berdasarkan teknologi *grid-based*. *Grid* tersebut diberi nama dengan alfanumerik karakter yang berbeda satu sama lain. Penamaan Plus Codes hanya menggunakan 20 karakter alfanumerik, yaitu 2, 3, 4, 5, 6, 7, 8, 9, C, F, G, H, J, M, P, Q, R, V, W, X. Plus Codes dapat ditemukan untuk masing – masing tempat yang terdaftar di Google Maps. Selain itu terdapat dua jenis Plus Codes mempunyai karakter sebelum simbol “+” berjumlah 4 (Plus Codes pendek) dan berjumlah 8 (Plus Codes panjang). Pada Google Maps yang ditampilkan adalah Plus Codes pendek. Untuk mendapatkan Plus Codes panjang perlu ke halaman Plus Codes, yaitu “<https://plus.codes/map>”. Lalu mencari dengan menggunakan Plus Codes pendek yang didapatkan dari Google Maps. Hasil yang ditampilkan dari pencarian tersebut adalah Plus Codes pendek, Plus Codes panjang, *latitude* dan *longitude*.