

## **BAB 3**

### **METODE PENELITIAN**

#### **3.1 ALAT DAN BAHAN**

Penelitian dilakukan dalam beberapa tahap yaitu tahap perancangan sistem, pembuatan dan yang terakhir yaitu tahap pengujian sistem. Alat dan bahan yang digunakan dalam penelitian ini meliputi bahasa pemrograman *python*, *visual studio code*, *dataset* yang dikumpulkan dari *website* pemberitaan dan juga *website turnbackhoax.id*, *google colab* yang digunakan untuk menyimpan *code* program. *Website turnbackhoax.id* merupakan situs yang dikelola oleh Mafindo, Masyarakat Anti *Hoax* Indonesia.

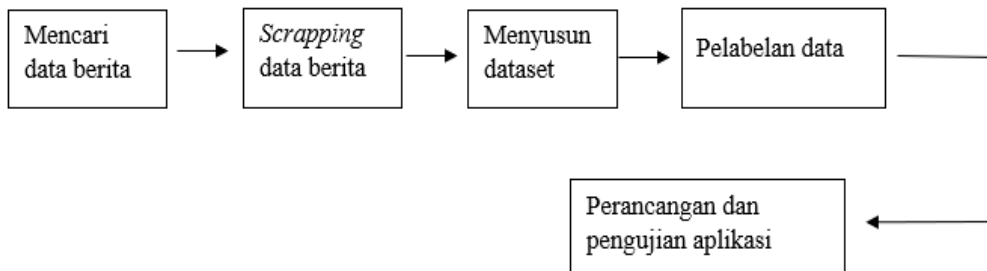
Sumber kontennya adalah Forum FAFHH, Forum Anti Fitnah Hasut dan *Hoax*. Mafindo adalah organisasi perkumpulan resmi yang didirikan pada tanggal 19 November 2016, Akta Notaris Nomor 1 Tanggal 19 November 2016 yang dibuat oleh ISMA JANUARTI, SH., M.KN, SK (Surat Keputusan) Pendirian Perkumpulan NOMOR AHU-0078919.AH.01.07. TAHUN 2016 Menteri Hukum dan Hak Asasi Manusia.

Kepala Biro Penerangan Masyarakat Divisi Humas, menuturkan orang yang menebarkan informasi palsu atau hoax di dunia maya akan dikenakan hukum positif. Hukum positif yang dimaksud adalah hukum yang berlaku. Maka, penebar hoax akan dikenakan KUHP, Undang-Undang No.11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (ITE), Undang-Undang No.40 Tahun 2008 tentang Penghapusan Diskriminasi Ras dan Etnis, serta tindakan ketika ujaran kebencian telah menyebabkan terjadinya konflik sosial. Pada pasal 45A ayat (1) UU ITE disebutkan, setiap orang yang sengaja menyebarkan berita bohong dan menyesatkan yang mengakibatkan kerugian konsumen dalam transaksi elektronik bisa dikenakan pidana penjara paling lama enam tahun dan/atau denda maksimal Rp 1 miliar.

### 3.2 ALUR PENELITIAN

Penelitian ini dimulai dengan mencari berita terlebih dahulu di *website* pemberitaan untuk berita benar, dan *website turnbackhoax.id* untuk berita *hoax*. Melakukan *scraping* untuk memperoleh data secara cepat, *Web scraping* merupakan teknik yang digunakan untuk mengumpulkan data dan informasi dari suatu *website* tertentu kemudian menyimpannya dalam format yang diinginkan. Kemudian menyusun data tersebut menjadi dataset (kumpulan data) meliputi data benar dan *hoax* dengan pemberian label 0 untuk data *hoax* dan 1 untuk data benar.

Dataset disusun berdasarkan jenis labelnya dimana untuk berita benar yang bersumber dari *kompas.com*, *detik.com*, *liputan6.com* dan *TribunNews.com* di beri label 1, untuk berita *hoax* yang bersumber dari *website turnbackhoax.id* diberi label 0. Melakukan pemrograman dengan bahasa pemrograman *python* hingga mencapai nilai akurasi. Untuk tampilan *website* maka dilakukan desain *frontend and backend* yang selanjutnya semua prosesnya di gabungkan hingga membentuk suatu aplikasi pendeteksi berita *hoax*. Untuk lebih jelasnya alur penelitian akan digambarkan melalui blok diagram berikut:



Gambar 3. 1 Blok Diagram Alur Penelitian

### 3.1.1 Tahap Mencari data berita

Pada tahap ini dilakukan observasi dan penentuan terlebih dahulu mengenai *website* yang akan digunakan untuk diambil data beritanya. Dalam penelitian ini untuk data berita benar di ambil dari *website* kompas.com, detik.com, liputan6.com dan TribunNews.com, sedangkan untuk berita *hoax* diambil dari *website* *turnbackhoax.id*.

### 3.1.2 Tahap Scrapping

*Web* scraping merupakan cara untuk mengumpulkan data dan informasi dari sebuah *website* tertentu kemudian disimpan dalam format yang diinginkan. Dalam penelitian ini menggunakan Teknik DOM (*Document Object Model*) yang berisi konten, *style* dan struktur *file* XML dalam suatu *website*. Berikut Langkah-langkah yang dilakukan dalam *scrapping* data:

- 1) Masuk ke aplikasi *chrome* dan ketikkan *chrome web store*



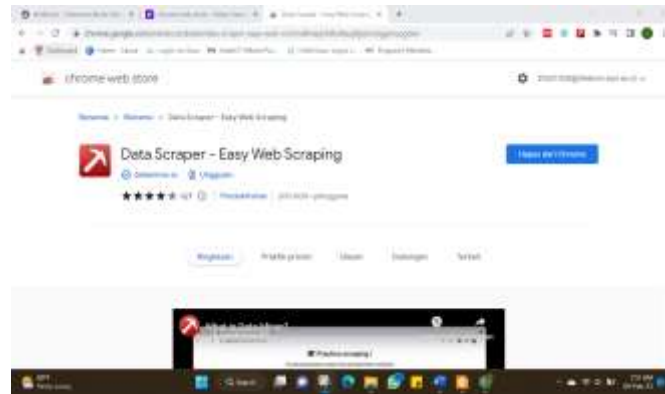
Gambar 3. 2 Tampilan *Chrome Web Store*

- 2) Kemudian pilih *Data Scraper - Easy Web Scraping*



Gambar 3. 3 Tampilan *Data Scraper - Easy Web Scraping*

3) Selanjutnya *install extension Data Scraper - Easy Web Scraping*



Gambar 3. 4 Tampilan *Install Data Scraper - Easy Web Scraping*

4. Setelah *extension Data Scraper - Easy Web Scraping* sudah berhasil di *install*, maka akan ada di samping kanan atas *browser* yang dipakai, seperti tertera pada gambar berikut.



Gambar 3. 5 Tampilan *Data Scraper - Easy Web Scraping* berhasil di *install*

5. Lalu lakukan *web scraping* dengan cara klik kanan kemudian pilih *inspect* dan klik *web scraper* di paling kanan.



Gambar 3. 6 Proses *Web Scraping*

- Selanjutnya mulai *scraping* dengan cara klik *create new sitemap* dan pilih *create sitemap*, isi *sitemap name* sesuai keinginan dan salin *link url website* tersebut dan simpan *sitemap*.



Gambar 3. 7 Proses *Web Scraping*

- Kemudian isikan *id* dengan judul berita yang ingin di ambil, *ceklis multiple*, dan kemudian pilih *select*. Selanjutnya pilih teks yang ingin dimasukkan ke data, jika sudah klik *done selecting*.



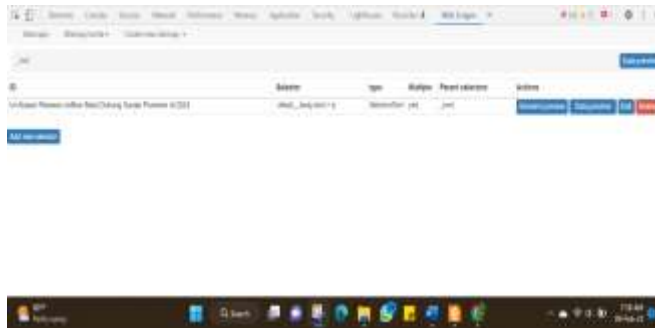
Gambar 3. 8 Proses *Web Scraping*

11) Simpan *sitemap*.



Gambar 3. 9 Proses *Web Scraping*

12) Berikut tampilan ketika sudah disimpan



Gambar 3. 10 Proses *Web Scraping*

13) Langkah terakhir klik *sitemap* berita kemudian pilih *Export* data. Disini ada dua jenis *file* yaitu *.xlsx* dan *.csv*. Jika sudah dipilih sesuai dengan kebutuhan otomatis data akan tersimpan di perangkat kita.



Gambar 3. 11 Proses *Web Scraping*

### 3.1.3 Tahap Menyusun Dataset

Pada tahap ini, data yang sudah di *scrapping* dan di simpan ke dalam computer kemudian disusun dengan urutan data berita benar dilanjutkan dengan data berita *hoax*. Hal ini dilakukan karena data di ambil dari *website* yang berbeda-beda dengan tujuan untuk menyatukan data menjadi satu dataset.

### 3.1.4 Tahap Pelabelan Data

Pada tahap ini dilakukan pelabelan pada data yang telah disusun menjadi satu dataset. Dataset disusun berdasarkan jenis labelnya dimana untuk berita benar yang bersumber dari *kompas.com*, *detik.com*, *liputan6.com* dan *TribunNews.com* di beri label 1, untuk berita *hoax* yang bersumber dari *website turnbackhoax.id* diberi label 0. Pelabelan data bertujuan untuk memudahkan mengenali mana berita yang benar dan yang *hoax*.

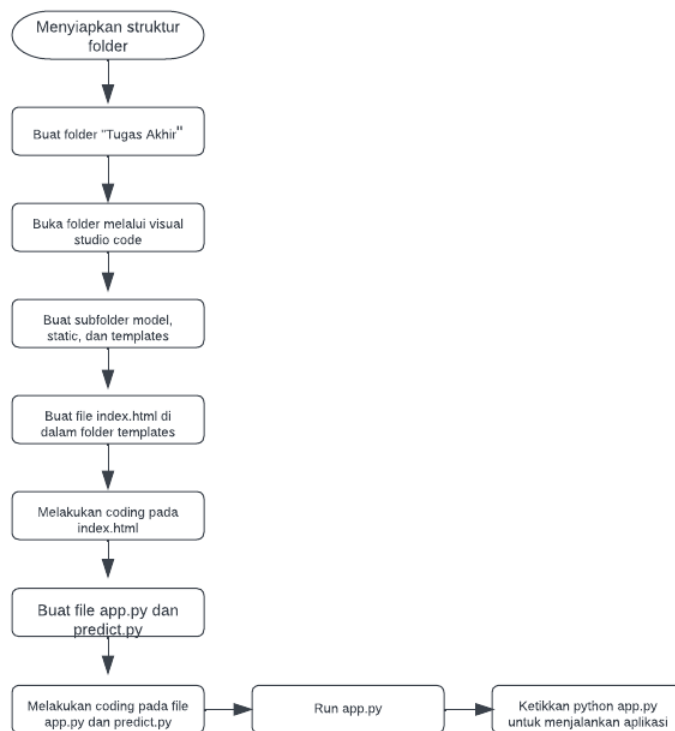
### 3.1.5 Tahap Perancangan dan Pengujian

Penelitian dimulai dengan pencarian berita di *website* hingga mengumpulkannya menjadi dataset sedangkan untuk tahap pemodelan AI dimulai dari *asking the right question* yang berarti menanyakan pertanyaan yang tepat. Tujuan dari tahap ini adalah untuk mengidentifikasi masalah dan mengidentifikasi aspek terpenting dari proyek, khususnya fase awal, yang akan mengungkap bagaimana data proyek akan disimpan. Masalah utama dengan proyek AI merupakan masalah yang serius dengan tujuan (objektif). Ini sangat penting karena tidak setiap saat mengharuskan penggunaan AI untuk mendapatkan hasil yang diinginkan. Kita harus menentukan apakah sebuah masalah dapat diselesaikan dengan solusi yang lebih komprehensif (seperti resolusi otomatis), atau apakah membutuhkan bahan pendukung yang lebih kompleks, seperti kecerdasan buatan.

Penelitian ini juga menggunakan pemodelan *front end* dan *back end*. Yang mana *front end* adalah bagian *website* yang terlihat. Seperti namanya, *front end* mengacu pada *front end website*. *Back end* mengacu pada situs *web* atau perangkat lunak manapun yang tidak bisa diakses oleh *user*. Sebagian besar konten di *internet* dibangun dengan cepat, ketika halaman dibaca, disimpan seluruhnya atau sebagian di *server web*.

Skrip ini didasarkan pada konten yang ditampilkan di *browser* pengguna, semua yang terjadi sebelum halaman ditampilkan di *browser* adalah bagian dari *back end*.

Pada penelitian ini tahap data *training* dengan menggunakan 1430 data yang dibagi kedalam 5 kelompok. Dengan jumlah minimal data yang digunakan yaitu untuk data *real* sebanyak 20 dan data *fake* 20 dengan jumlah data *class* sebanyak 4 dengan berita benar diambil dari Kompas.com, detik.com, liputan6.com, dan TribunNews.com untuk berita benar dan *website turnbackhoax.id* untuk berita *hoax*. Algoritma yang digunakan yaitu *naïve bayes* dimana algoritma ini digunakan untuk meningkatkan klasifikasi, ini didasarkan pada teori *bayesian* probabilistik. Hal ini digunakan untuk klasifikasi teks yang melibatkan set *information training* dimensi tinggi. Dataset yang digunakan dalam proses penelitian ini berjumlah 1004 data untuk berita benar dan 1001 data untuk berita *hoax*. Dalam dataset terdiri dari *web scrap*, judul, teks berita, dan *link* berita



Gambar 3. 12 *Flowchart* Perancangan Aplikasi