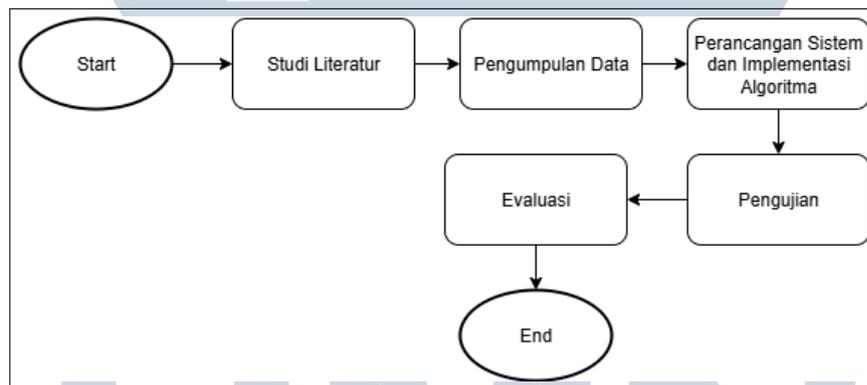


## BAB 3 METODOLOGI PENELITIAN

### 3.1 Alur Penelitian

Untuk mencapai tujuan penelitian ini, sebuah alur kerja yang sistematis telah dirancang dan divisualisasikan dalam *flowchart* pada Gambar 3.1. Struktur Alur Penelitian. *Flowchart* ini dimulai dengan Studi Literatur untuk memahami referensi yang relevan. Kemudian, dilanjutkan dengan Pengumpulan Data sebagai bahan penelitian. Setelah itu, dilakukan Perancangan Sistem dan Implementasi Algoritma untuk membangun model yang sesuai. Tahap berikutnya adalah Pengujian untuk mengevaluasi performa sistem. Selanjutnya, Evaluasi dilakukan untuk menganalisis hasil pengujian sebelum mencapai kesimpulan penelitian.



Gambar 3.1. Struktur alur penelitian

#### 3.1.1 Studi Literatur

Pada tahap ini, perlu dilakukan penelitian lebih lanjut mengenai topik yang akan dibahas. Beberapa literatur yang dipelajari adalah *Text Mining*, *Sentiment Analysis*, *Preprocessing*, *Term Frequency-Inverse Document Frequency (TF-IDF)*, *Support Vector Machine*, dan *Confusion Matrix*.

#### 3.1.2 Pengumpulan Data

Pada tahap ini, dilakukan pengumpulan data yang terkait dengan fokus penelitian. Pengumpulan data dilakukan dengan *scraping* data media sosial X (Twitter). *Scraping* data menggunakan *command-line tool* yaitu *Tweet-Harvest*.

Data yang diambil didasarkan pada kata kunci "indodax hack" dan mengacu pada komentar di salah satu *posting* Indodax yang memberikan klarifikasi terkait insiden peretasan. Rentang waktu data ini dibatasi dari 11 September 2024 sampai 13 September 2024. Setelah data dikumpulkan, langkah selanjutnya yaitu tahap *preprocessing*.

### 3.1.3 Perancangan Sistem

Pada bagian ini dibahas mengenai alur sistem yang dibuat. Alur sistem yang dibuat dalam penelitian ini dijelaskan melalui *flowchart*. Terdapat beberapa alur sistem yang dibuat dengan *flowchart*, antara lain gambaran umum sistem, proses *scraping* data, *preprocessing* data, proses *labeling*, implementasi TF-IDF, implementasi *Support Vector Machine*.

### 3.1.4 Pengujian

Pada tahap ini, dilakukan pengujian terhadap model *Support Vector Machine* (SVM) untuk menilai kinerjanya dalam klasifikasi sentimen. Pengujian dilakukan dengan menggunakan data uji yang telah dipersiapkan setelah tahap *preprocessing* dan pelabelan sentimen. Model diuji untuk melihat kemampuannya dalam mengklasifikasikan sentimen dengan benar berdasarkan pola yang telah dipelajari dari data latih.

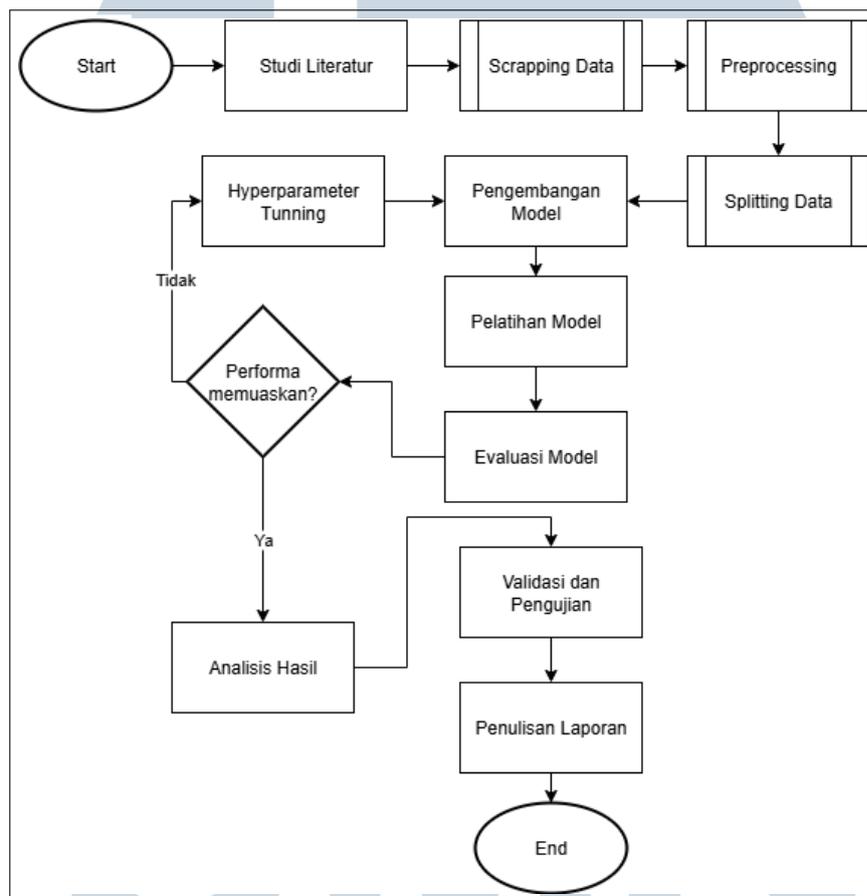
### 3.1.5 Evaluasi

Evaluasi model dilakukan dengan menggunakan metrik standar dalam machine learning, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik-metrik ini digunakan untuk mengukur sejauh mana model SVM mampu melakukan klasifikasi sentimen secara akurat dan konsisten. Hasil evaluasi ini akan memberikan gambaran mengenai keefektifan model SVM dalam menangani tugas klasifikasi sentimen.

## 3.2 Gambaran Umum Sistem

Pada Gambar 3.2, proses dimulai dengan studi literatur untuk memahami dasar teoritis, diikuti oleh pengumpulan data dan *preprocessing* untuk mempersiapkan data. Data kemudian dibagi menjadi *dataset* pelatihan dan

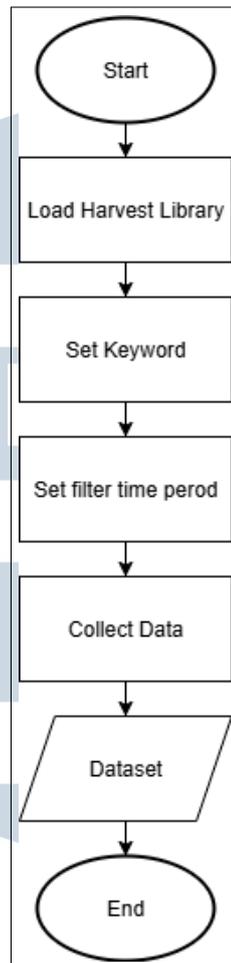
pengujian. Tahap selanjutnya adalah pengembangan model dan penyetelan *hyperparameter* untuk meningkatkan performa. Jika performa model memuaskan, proses dilanjutkan dengan analisis hasil, pelatihan ulang model, evaluasi, serta validasi dan pengujian. Jika performa tidak memuaskan, proses kembali ke tahap pengembangan model atau penyetelan *hyperparameter*. Setelah model dianggap siap, hasilnya dianalisis dan dilaporkan, menandai akhir dari proses.



Gambar 3.2. Gambaran umum sistem

### 3.2.1 Scraping Data

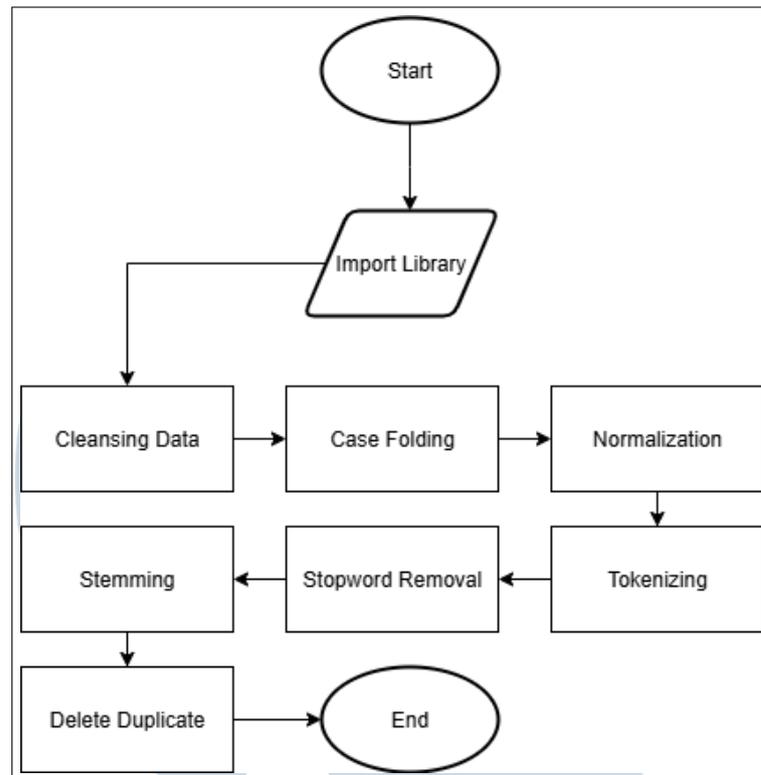
Alur proses *scraping data* ini dimulai dengan langkah inialisasi, diikuti dengan memuat *library* yang diperlukan. Kemudian, seperti yang ditunjukkan dalam Gambar 3.3, kata kunci (keyword) dan periode waktu filter ditetapkan untuk memastikan data yang dikumpulkan relevan dan sesuai dengan kebutuhan. Setelah itu, proses pengumpulan data dilakukan, menghasilkan dataset yang siap digunakan.



Gambar 3.3. Flowchart Scraping Data

### 3.2.2 Data Preprocessing

Setelah data terkumpul, proses *preprocessing* dilakukan untuk menyiapkan data. Gambar 3.4 mengilustrasikan bahwa proses ini dimulai dengan mengimpor *library* yang diperlukan. Selanjutnya, dilakukan penghapusan *stopword* (kata-kata umum yang tidak signifikan), mengubah semua teks menjadi huruf kecil (*case folding*) pembersihan data (*cleansing*) untuk menghilangkan karakter atau simbol yang tidak diinginkan, dan penghapusan data duplikat. Proses *stemming* dilakukan untuk mengubah kata ke bentuk dasarnya, dan *normalization* digunakan untuk menyamakan variasi kata ke bentuk standar.

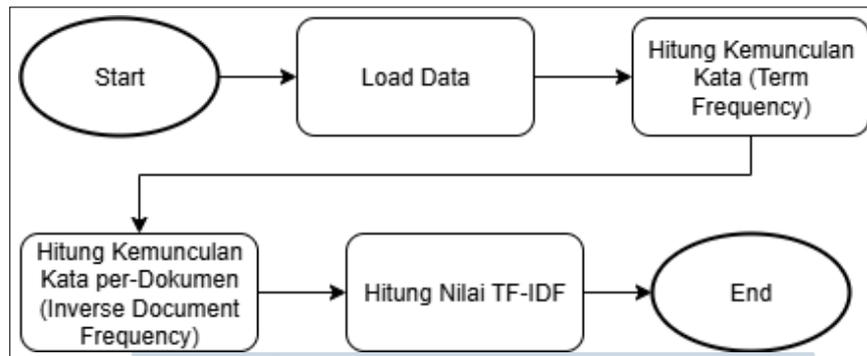


Gambar 3.4. Flowchart Data Preprocessing

### 3.2.3 Term Frequency Inverse Document Frequency (TF-IDF)

Pembobotan TF-IDF merupakan tahapan penting dalam analisis teks, yang alurnya diilustrasikan dalam Gambar 3.5. Proses ini dimulai dengan memuat data, lalu menghitung *term frequency (TF)*, yaitu jumlah kemunculan setiap kata dalam dokumen. Selanjutnya, *inverse document frequency (IDF)* dihitung untuk menilai seberapa penting kata tersebut dalam keseluruhan kumpulan dokumen. Kedua nilai ini kemudian dikombinasikan untuk memperoleh bobot TF-IDF, yang menunjukkan relevansi kata dalam dokumen tertentu. Setelah semua perhitungan selesai, proses pun berakhir.

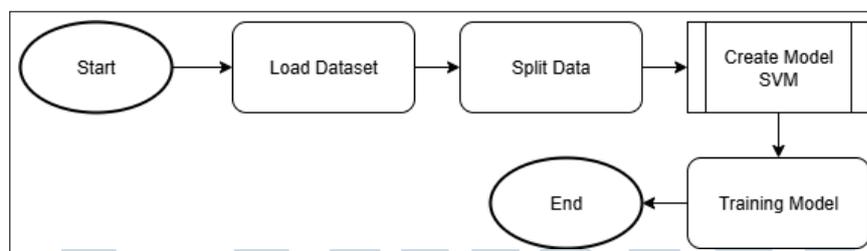
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA



Gambar 3.5. Flowchart TF-IDF

### 3.2.4 Splitting Data

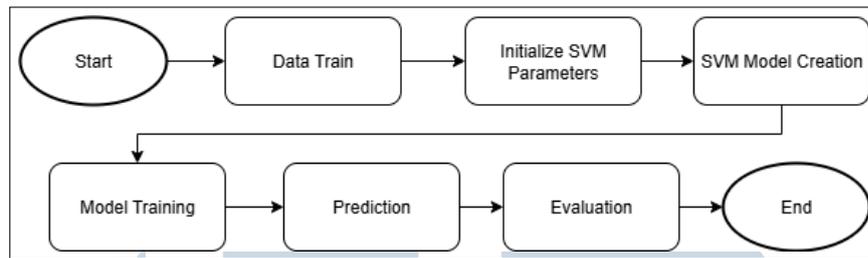
Untuk melatih model *Support Vector Machine (SVM)* secara efektif, tahapan *splitting data* sangat krusial dan alurnya disajikan dalam Gambar 3.6. Proses ini dimulai dengan memuat dataset, kemudian dilanjutkan dengan pembagian data menjadi set pelatihan dan pengujian. Setelah itu, dilakukan pembuatan model SVM, yang kemudian digunakan dalam tahap pelatihan model dengan data latih yang telah dibagi sebelumnya. Akhirnya, proses berakhir setelah model berhasil dilatih.



Gambar 3.6. Flowchart Splitting Data

### 3.2.5 Modeling Support Vector Machine

Seperti yang digambarkan dalam Gambar 3.7, proses pelatihan dan evaluasi model *Support Vector Machine (SVM)* dimulai dengan memuat data latih. Kemudian, parameter SVM seperti kernel,  $C$ , dan  $\gamma$  diinisialisasi. Setelah itu, model SVM dibentuk dan dilatih menggunakan data latih. Setelah pelatihan selesai, model digunakan untuk melakukan prediksi pada data uji. Hasil prediksi kemudian dievaluasi menggunakan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Setelah evaluasi selesai, proses pun berakhir.



Gambar 3.7. Flowchart Modeling SVM

