# BAB III ANALISA DAN PERANCANGAN

## Analisa

### Identifikasi Masalah

Indonesia sedang mengalami pandemi dan sedang berjuang menurunkan angka positif COVID-19 dengan melakukan vaksinasi. Penggunaan vaksin COVID-19 pastinya memunculkan respon dan opini dari masyarakat. Persepsi masyarakat terhadap vaksin sangat beragam dan saat ini cenderung negatif, karena masyarakat Indonesia saat ini masih banyak terpengahuh oleh informasi bohong tidak sesuai fakta yang beredar di media sosial. Penolakan pada vaksin COVID-19 sempat menjadi pembicaraan topik di Twitter. Menurut data vaksinasi COVID-19 Nasional kementrian Kesehatan per tanggal 08 Oktober 2021, menunjukkan bahwa jumlah vaksinasi dosis 1 masih mencapai 47% dari 208.265.720 penduduk sasaran vaksinasi. Menurut survei Kementerian Kesehatan RI, persentase responden yang ingin divaksinasi adalah 64,8%. 27,6% responden menyatakan tidak tahu, dan 7,6% responden menyatakan sangat menentang vaksinasi terhadap COVID-19. Survei dilakukan pada November 2020 dengan lebih dari 112.000 responden (Rosa, 2021). Berdasarkan penelitian yang dilakukan oleh (Lestari & Saepudin, 2021) dengan judul “Analisis Sentimen Vaksin Sinovac pada Twitter Menggunakan Algoritma Naïve Bayes”. Penelitian ini bertujuan untuk memberikan solusi penentuan presentase pendapat positif atau negatif *user* akun media sosial Twitter mengenai vaksin Sinovac. Perspektif masyarakat yang cenderung negatif terhadap vaksin COVID-19 khususnya vaksin Sinovac, vaksin Astrazeneca, vaksin Pfizer dan vaksin Moderna, jika tidak segera diatasi dan dibiarkan berlarut-larut maka akan menyebabkan masyarakat semakin enggan melakukan vaksinasi.

### Pemecahan Masalah

Berdasarkan uraian permasalahan yang dijabarkan diatas, maka dapat dilakukan penelitian tentang analisis persepsi masyarakat terhadap vaksin COVID-19 di Indonesia.

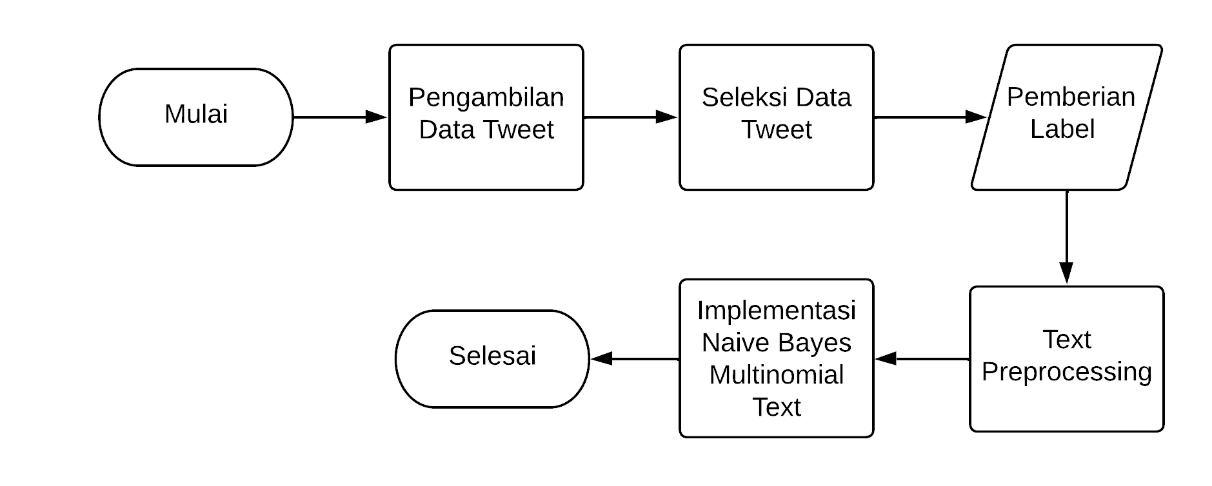
Penelitian yang akan dilakukan ini mempunyai kelebihan sebagai berikut:

1. Hasil analisis sentimen diharapkan dapat mengedukasi masyarakat terkait vaksin COVID-19 dan mempercepat pemerintah dalam melakukan vaksinasi kepada masyarakat.
2. Menjadi evaluasi dan masukan bagi pemerintah tentang pengadaan vaksin COVID-19.
3. Aplikasi *web crawling* data *tweet* yang akan dibuat, nantinya akan memiliki fitur *text preprocessing* secara otomatis dan bisa langsung men-*download* hasil *text preprocessing* secara otomatis dengan berformatkan .arff.
4. Bisa melakukan pengujian kalimat *tweet* apakah memiliki sentimen positif atau negatif dengan hasil yang akurat jika dibandingkan dengan software WEKA.

## Perancangan

### Perancangan Sistem

#### Flowchart

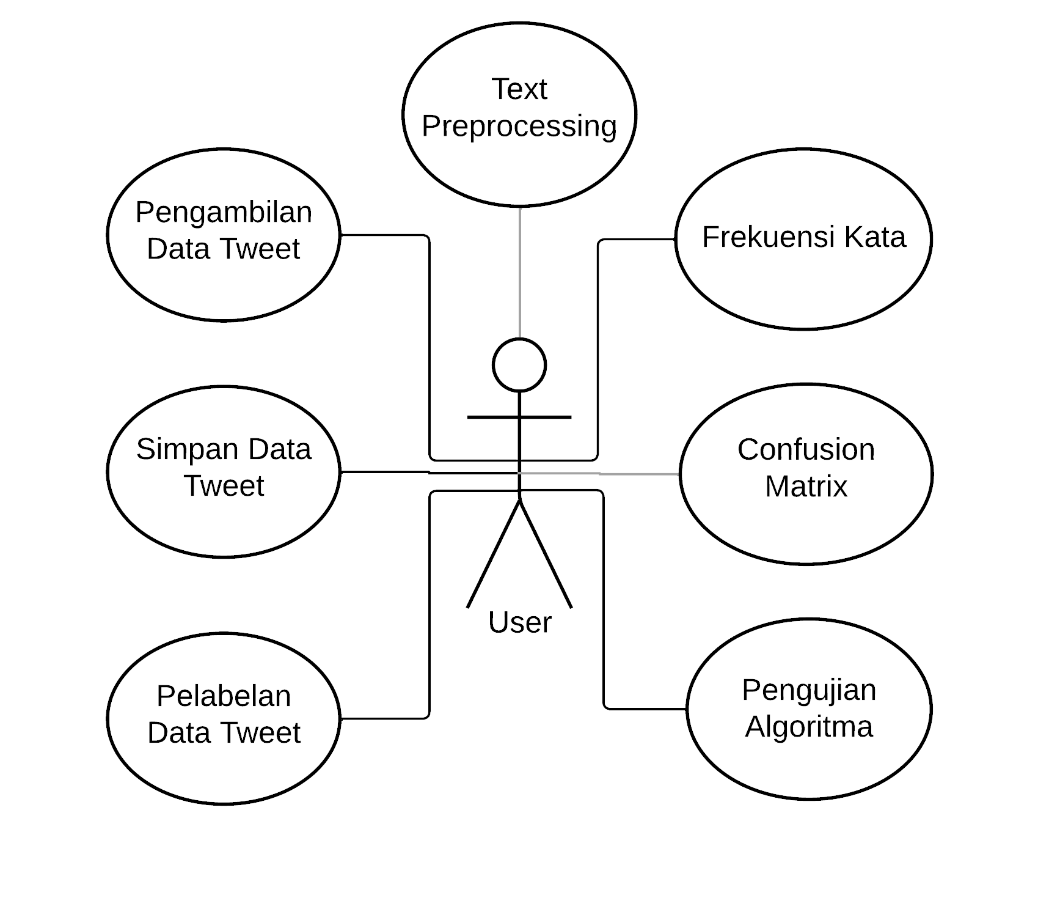


Gambar 3. 1 *Flowchart* Sistem

*Flowchart* diatas menjelaskan bahwa user harus melakukan proses pengambilan data *tweet* berdasarkan fitur / *keyword* yang sudah ditentukan. Selanjutnya setelah melakukan pengambilan data *tweet*, dilanjutkan dengan prosesmenyimpan data *tweet* yang dipilih ke dalam *database*. Setelah data *tweet* disimpan kedalam *database*, data tersebut dilakukan proses *text preprocessing*. Setelah proses *text preprocessing* selesai, proses selanjutnya yaitu implementasi naïve bayes multinomial text mulai dari proses menghitung frekuensi kata, confusion matrix dan pengujian algoritma naïve bayes multinomial ltext.

#### Use Case Diagram

*Use case* diagram adalah teknik untuk merekam persyaratan fungsional pada sebuah sistem, menggambarkan fungsi dari sebuah sistem.



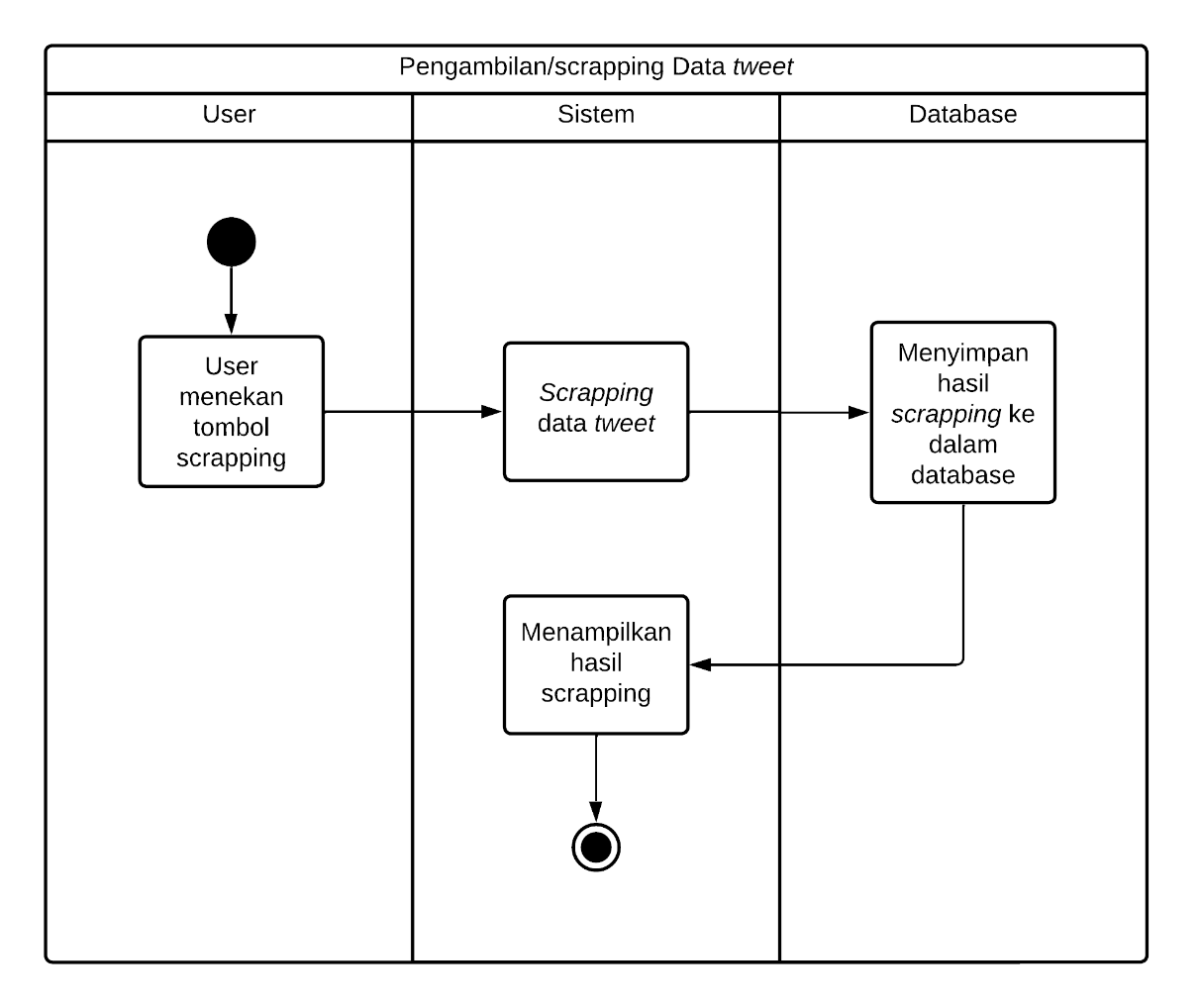
Gambar 3. 2 Use Case Diagram

Dari diagram *use case* diatas, terdapat aktor yang merupakan pengguna sistem. Pengguna dapat melakukan:

1. Pengguna dapat melakukan proses pengambilan data *tweet*
2. Pengguna dapat memilih *tweet* dan menyimpan data *tweet* ke dalam database
3. Pengguna dapat melakukan pelabelan positif atau negated terhadap suatu *tweet*.
4. Pengguna dapat melakukan perhitungan frekuensi setiap kata.
5. Pengguna dapat melakukan perhitungan *confusion matrix*.
6. Pengguna dapat melakukan pengujian suatu *tweet* apakah *tweet* tersebut memiliki sentimen positif atau negatif.

#### Activity Diagram

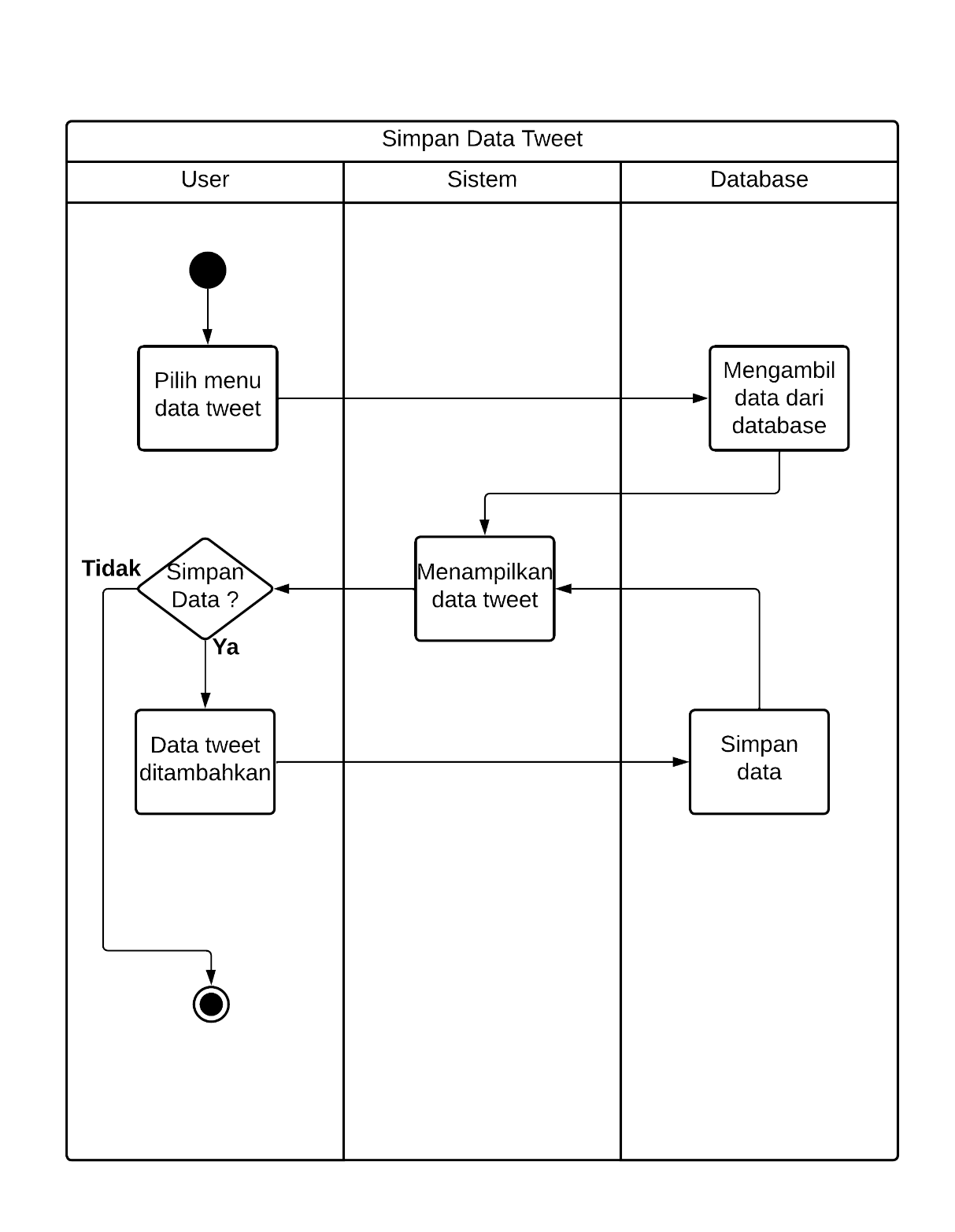
*Activity* diagram ini menggambarkan aliran yang berbeda dari aktivitas dalam sistem di bawah desain, bagaimana setiap aliran dimulai, keputusan yang mungkin terjadi, dan bagaimana mereka berakhir.Activity Diagram Pengambilan/Scrapping Data



Gambar 3. 3 Activity Diagram Pengambilan Data Tweet

Pengguna dapat melakukan *scrapping* data *tweet* dengan keyword Sinovac, Astrazeneca, Pfizer dan Moderna. Kemudian data *tweet* tersebut akan disimpan ke dalam *database*.

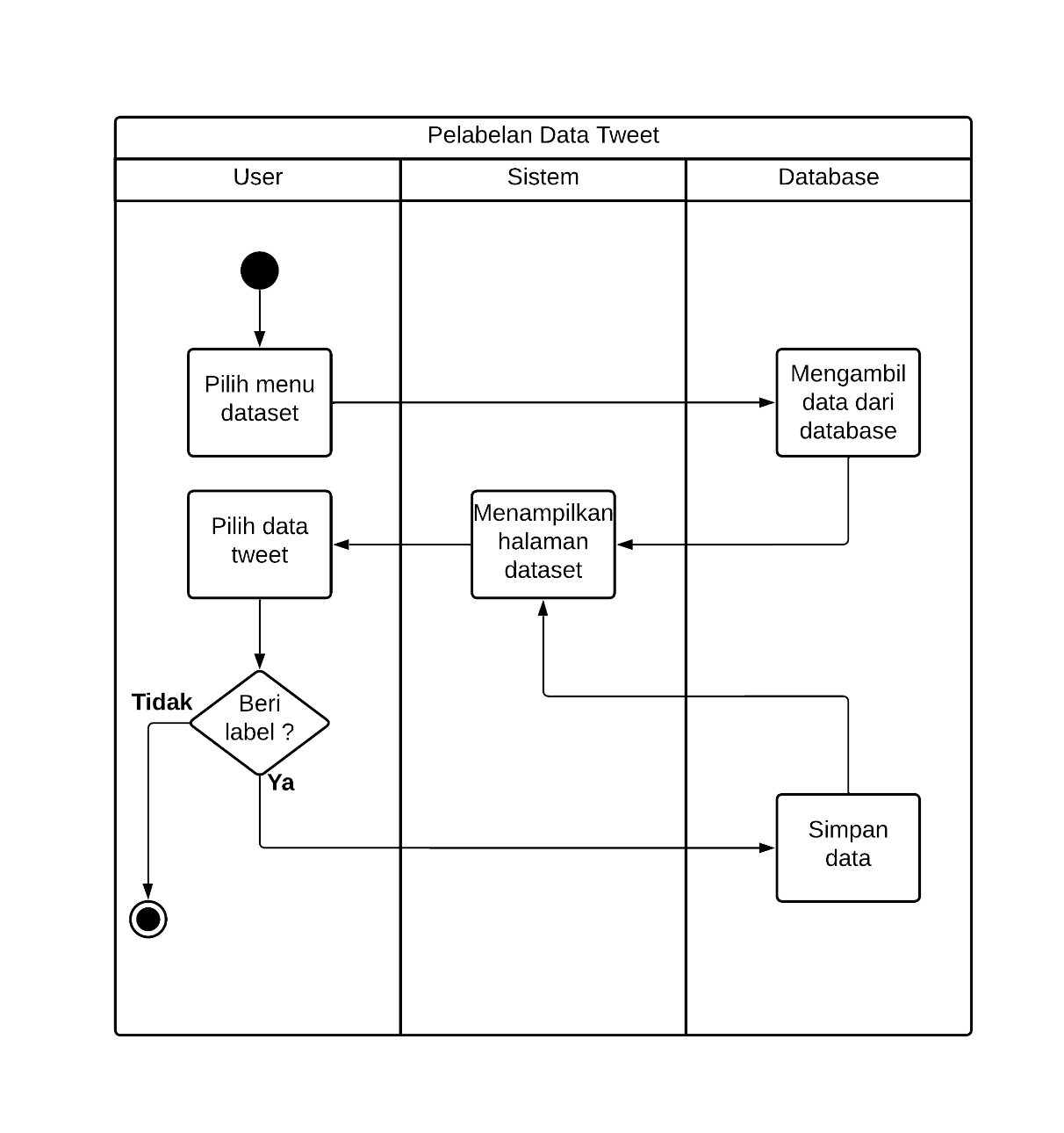
##### Activity Diagram Simpan Data Tweet



Gambar 3. 4 Activity Diagram Simpan Data Tweet

Pengguna dapat menyimpan data *tweet* dengan menyeleksi terlebih dahulu kalimat *tweet* mana yang akan disimpan. Apabila sudah menyeleksi data *tweet* tersebut, setelah menekan tombol simpan maka data *tweet* akan disimpan ke dalam *database*.

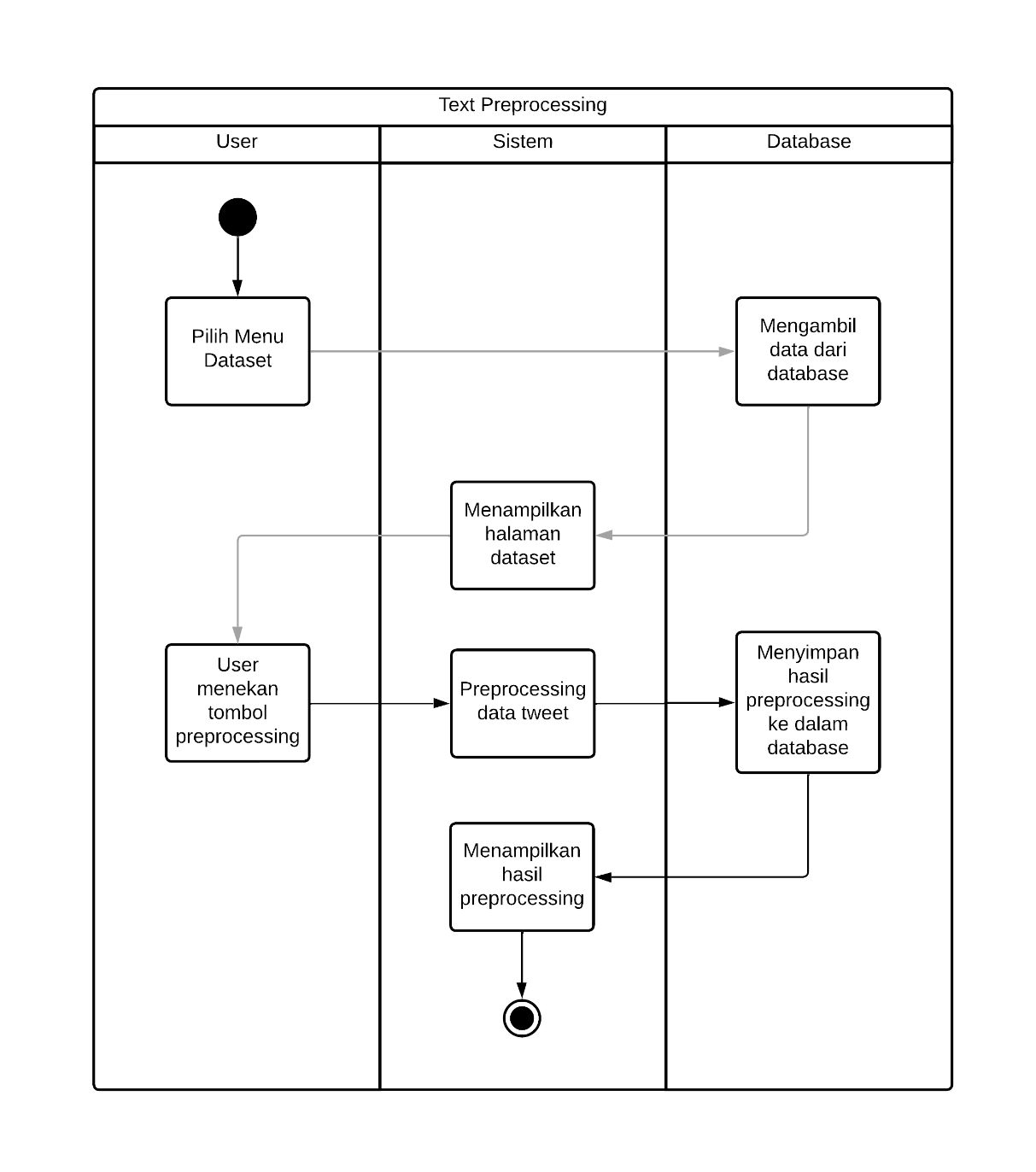
##### Activity Diagram Pelabelan Data Tweet



Gambar 3. 5 Activity Diagram Pelabelan Data Tweet

Pengguna dapat melakukan pelabelan data *tweet* dengan menyeleksi terlebih dahulu kalimat *tweet* mana yang akan diubah labelnya. Apabila sudah memberi label pada data *tweet* tersebut, setelah menekan tombol simpan maka label pada data akan berubah.

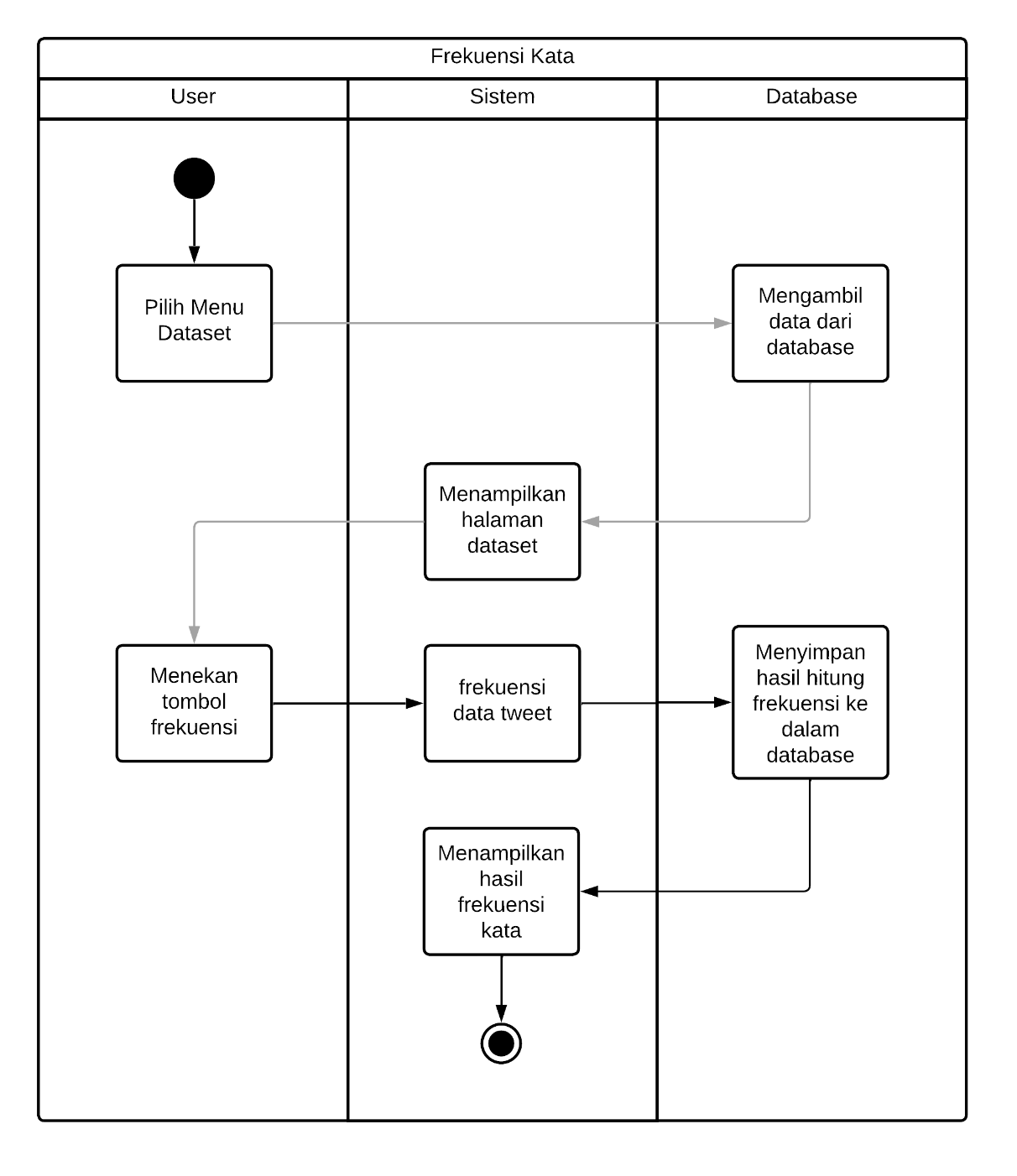
##### Activity Diagram Text Preprocessing



Gambar 3. 6 Activity Diagram Text Preprocessing

Pengguna dapat melakukan *text preprocessing* data *tweet* dengan menekan tombol *preprocessing*. Untuk preprocessing menggunakan 5 proses yaitu *Transform Cases, Tokenize, Filter Token (By Length), Stopword Removal, Stemming*.

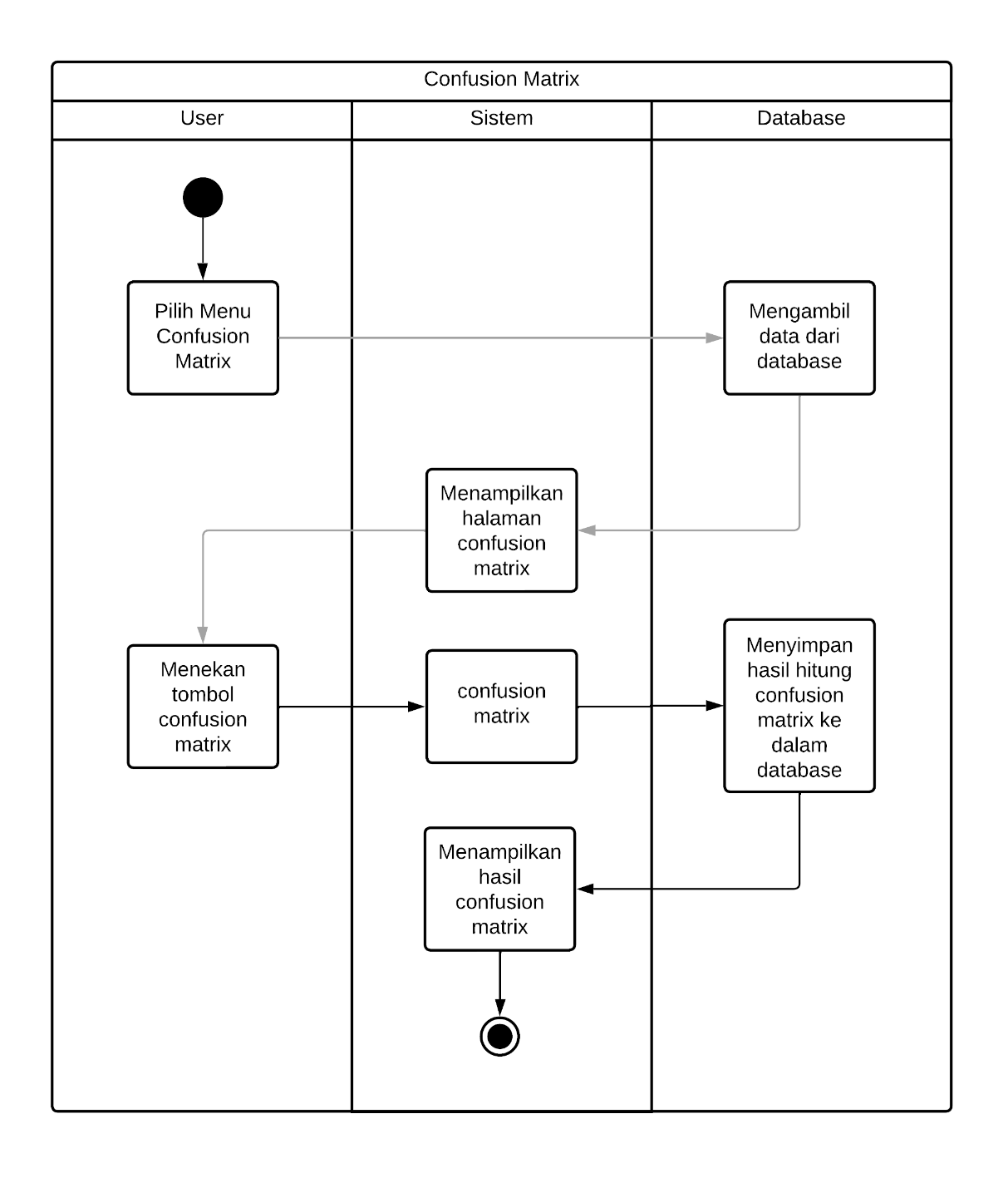
##### Activity Diagram Frekuensi Kata



Gambar 3. 7 Activity Diagram Frekuensi Kata

Pengguna dapat menghitung frekuensi setiap kata dengan menekan tombol frekuensi. Dalam proses ini nantinya akan menghitung probabilitas negatif dan probabilitas positif untuk setiap kata.

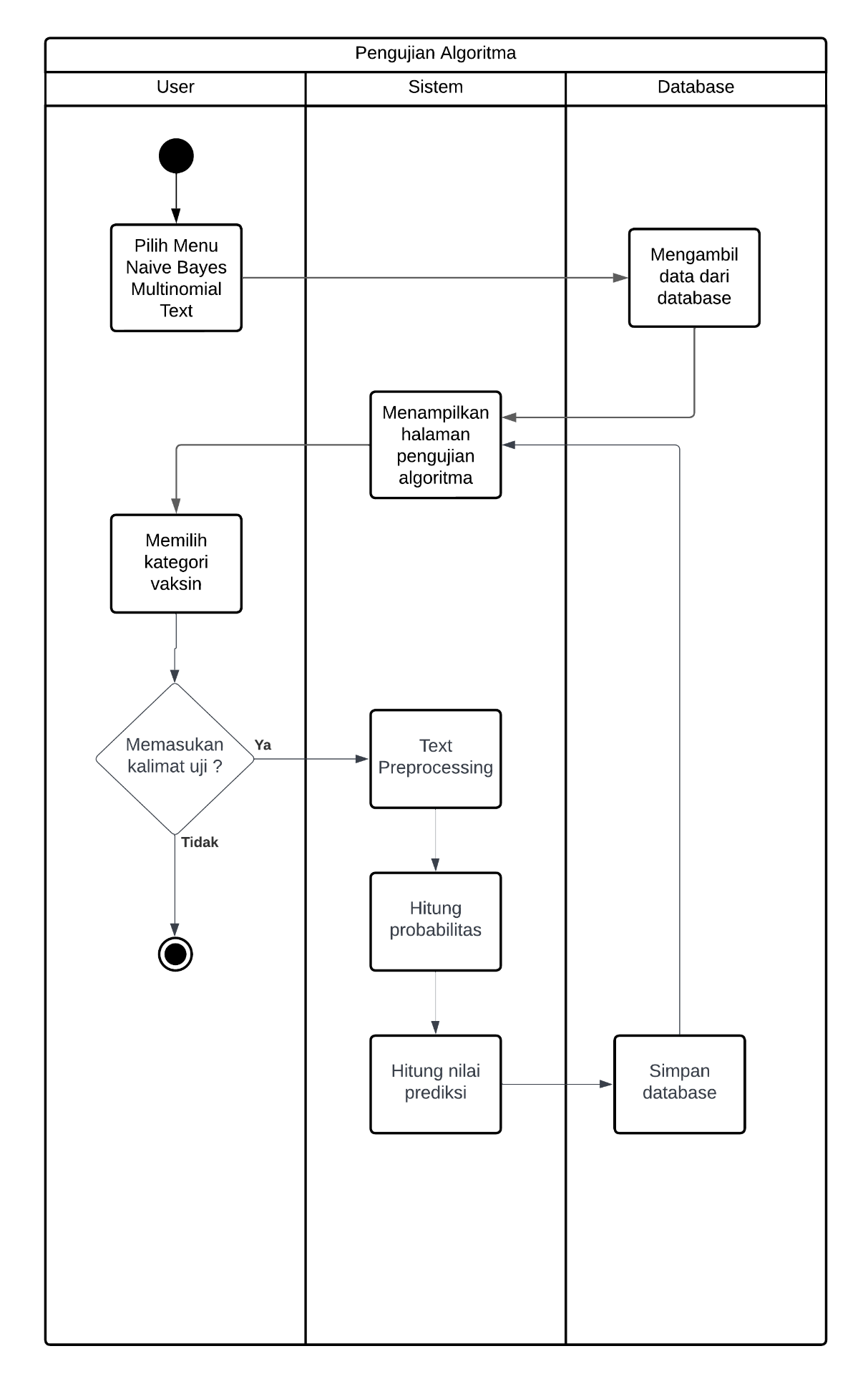
##### Activity Diagram Confusion Matrix



Gambar 3. 8 Activity Diagram Confusion Matrix

Pengguna dapat menghitung *confusion matrix* dengan menekan tombol Confusion Matrix. Dalam proses ini nantinya akan *Accuracy*, *Precision*, *Recall* dan *F-Measure*.

##### Activity Diagram Pengujian Algoritma

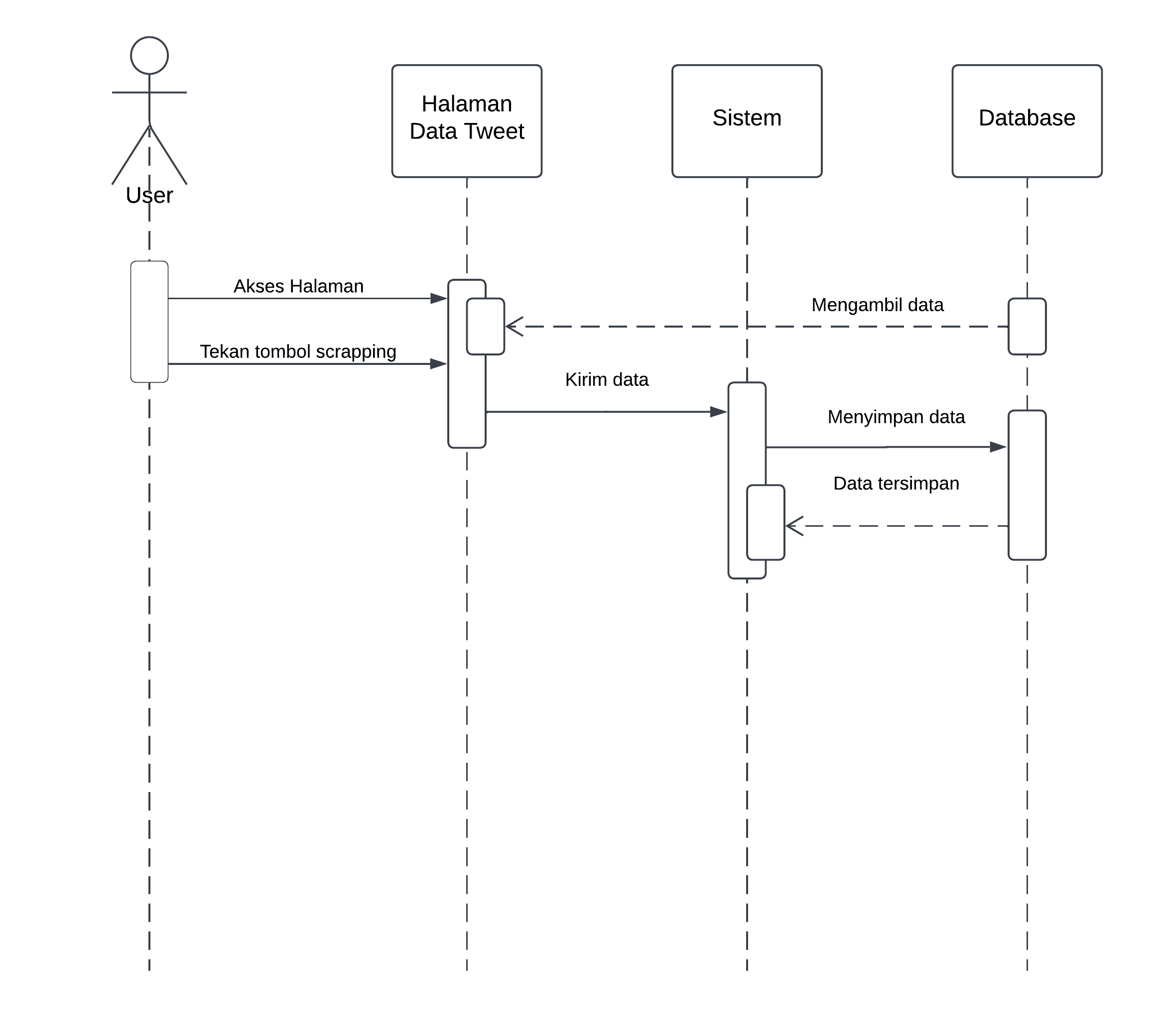


Gambar 3. 9 Activity Diagram Pengujian Algoritma

Pengguna dapat menguji suatu kalimat apakah kalimat *tweet* tersebut mempunyai sentimen positif atau negatif dengan mengisi kalimat *tweet* di form isian yang ada.

#### Sequence Diagram

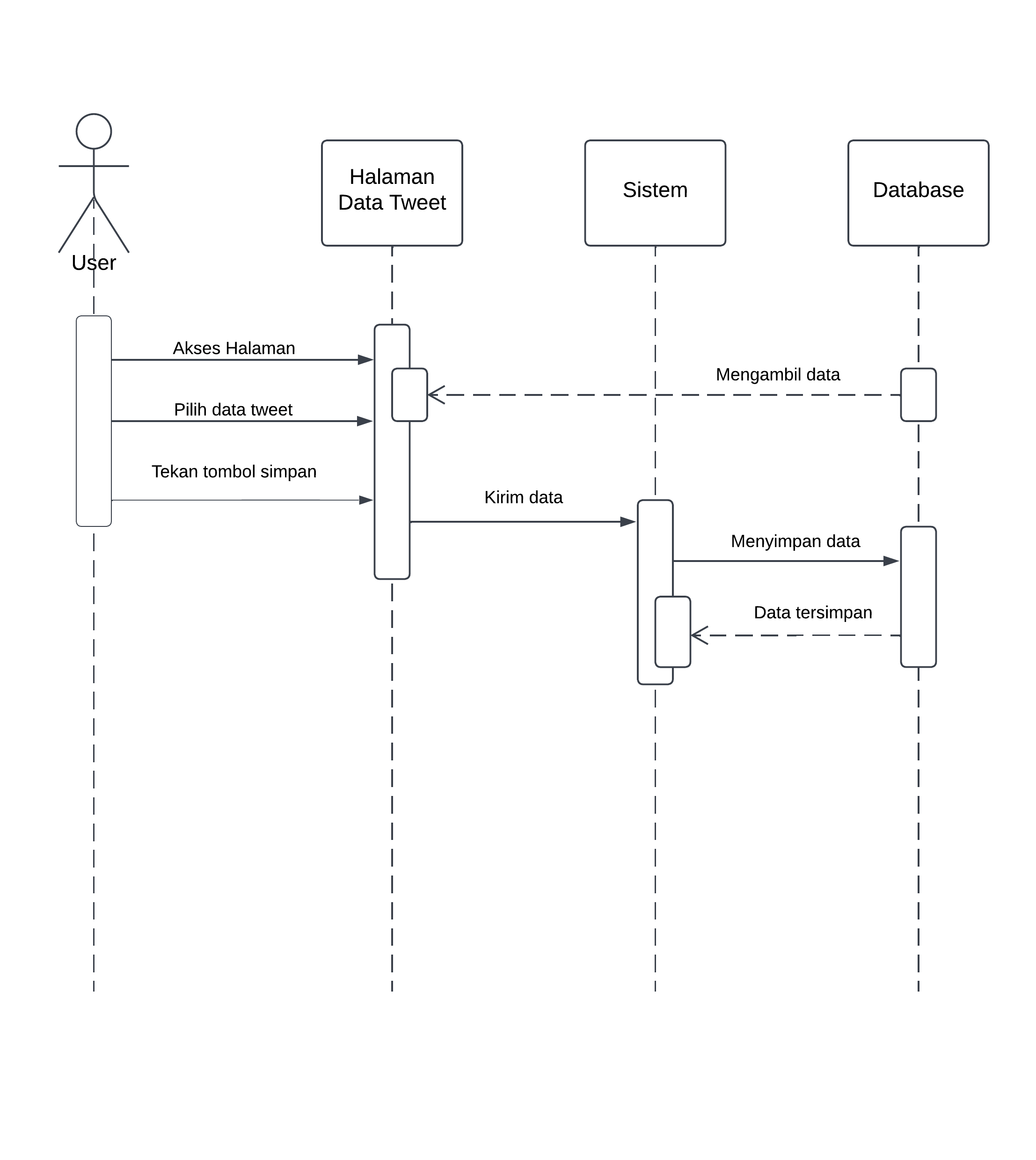
##### Sequence Diagram Pengambilan / Scrapping Data



Gambar 3. 10 Sequence Diagram Pengambilan Data Tweet

Sequence diagram pengambilan data atau *scrapping* data menjelaskan tentang urutan yang dilakukan untuk pengambilan data menggunakan API twitter.

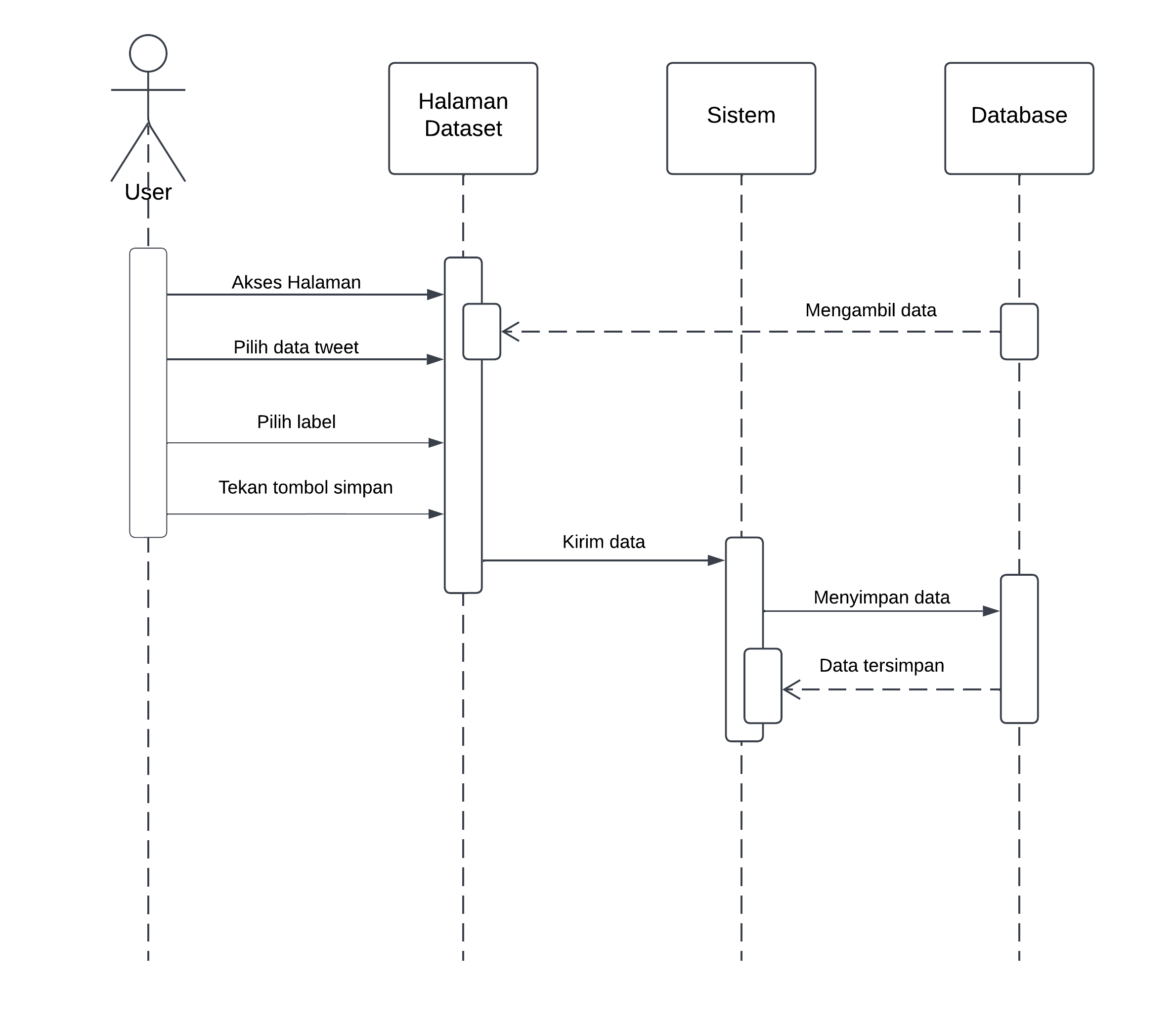
##### Sequence Diagram Simpan Data Tweet



Gambar 3. 11 Sequence Diagram Simpan Data Tweet

Sequence diagram simpan data *tweet* menjelaskan tentang urutan yang dilakukan untuk menyimpan data *tweet* yang akan digunakan sebagai data uji.

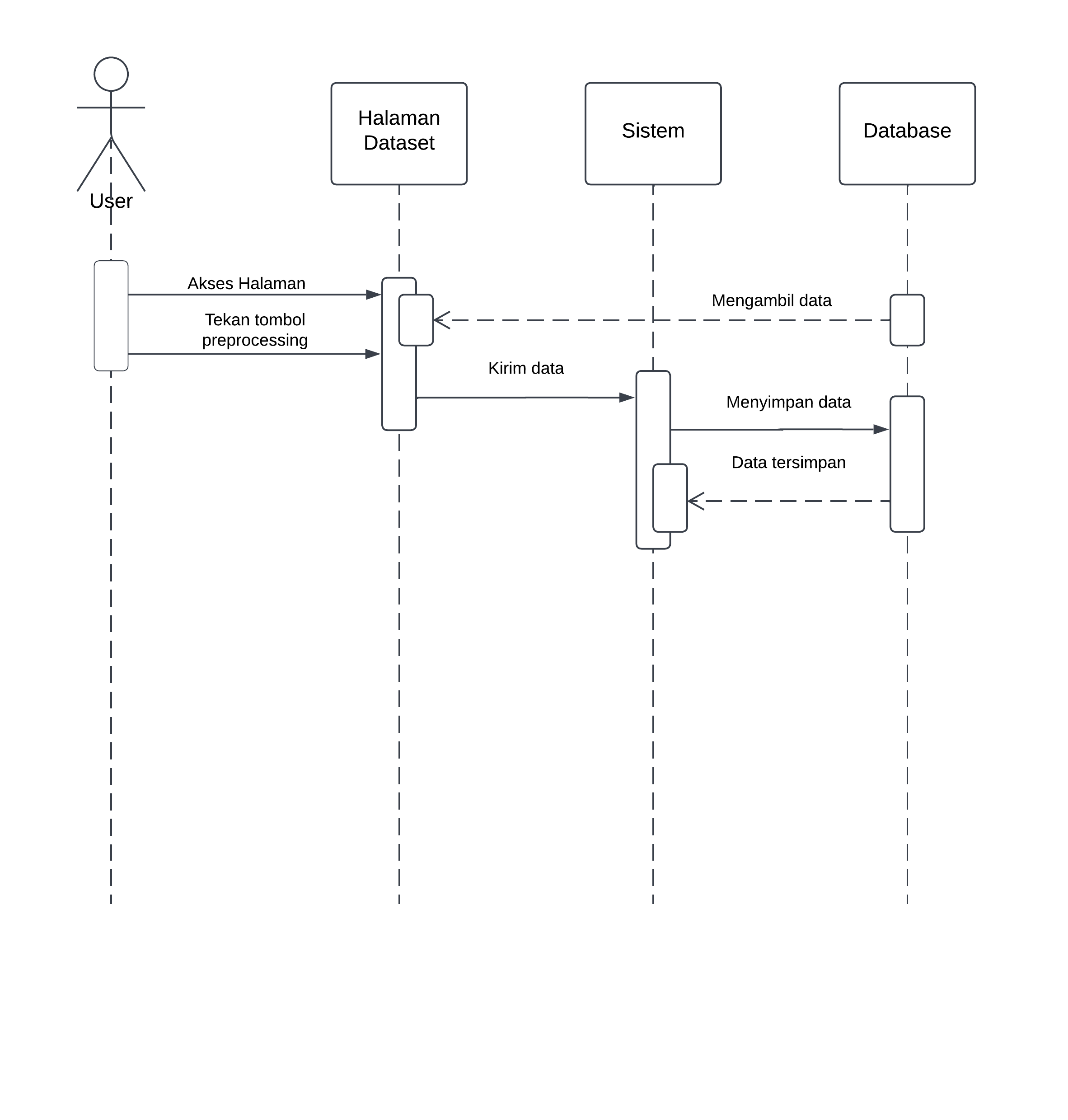
##### Sequence Diagram Pelabelan Data Tweet



Gambar 3. 12 Sequence Diagram Pelabelan Data

Sequence diagram pelabelan data *tweet* menjelaskan tentang urutan yang dilakukan untuk mengubah label positif atau negatif yang akan digunakan sebagai data uji.

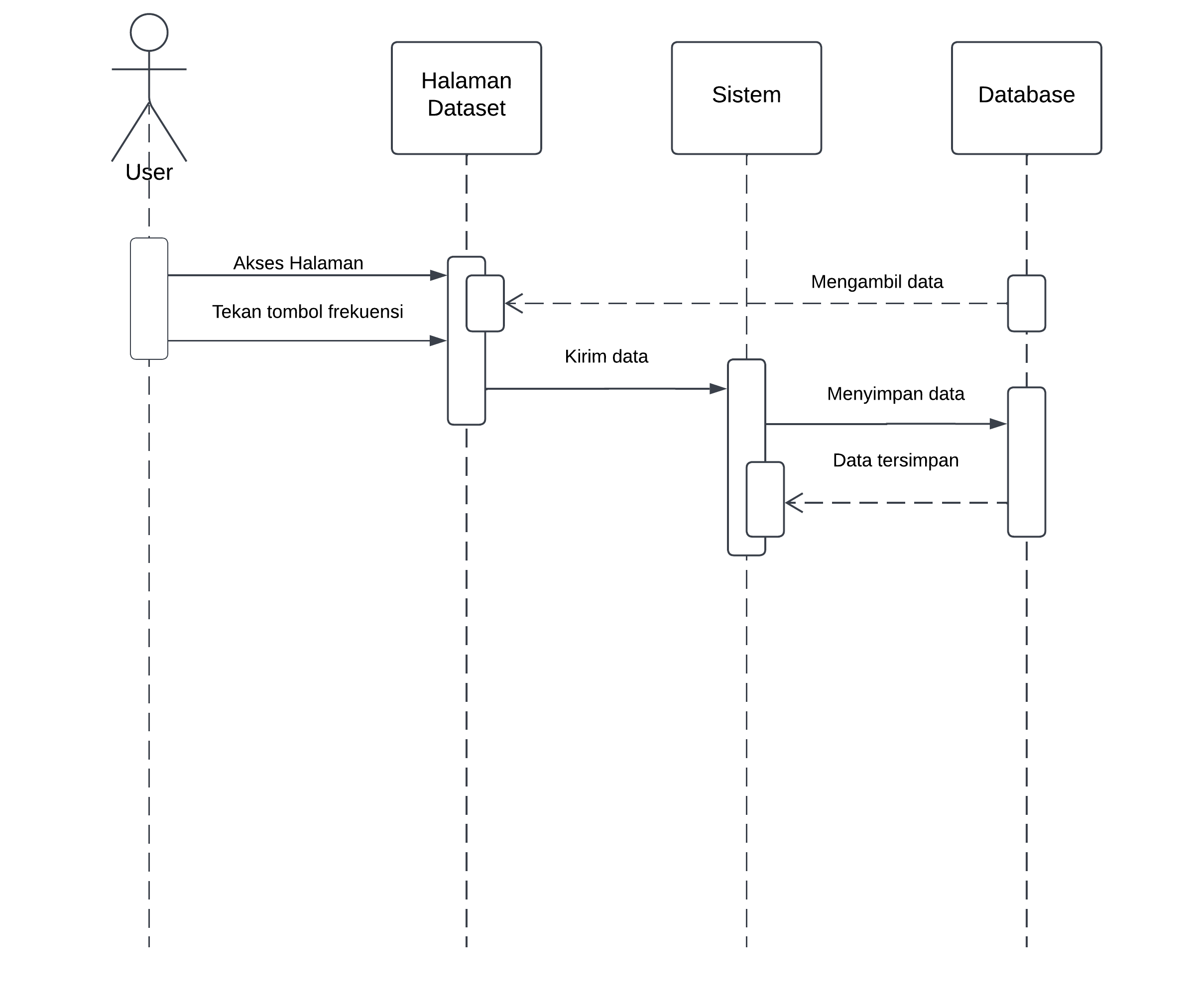
##### Sequence Diagram Text Preprocessing



Gambar 3. 13 Sequence Diagram Text Preprocessing

Sequence diagram text *preprocessing* menjelaskan tentang urutan yang dilakukan untuk melakukan proses *preprocessing* data.

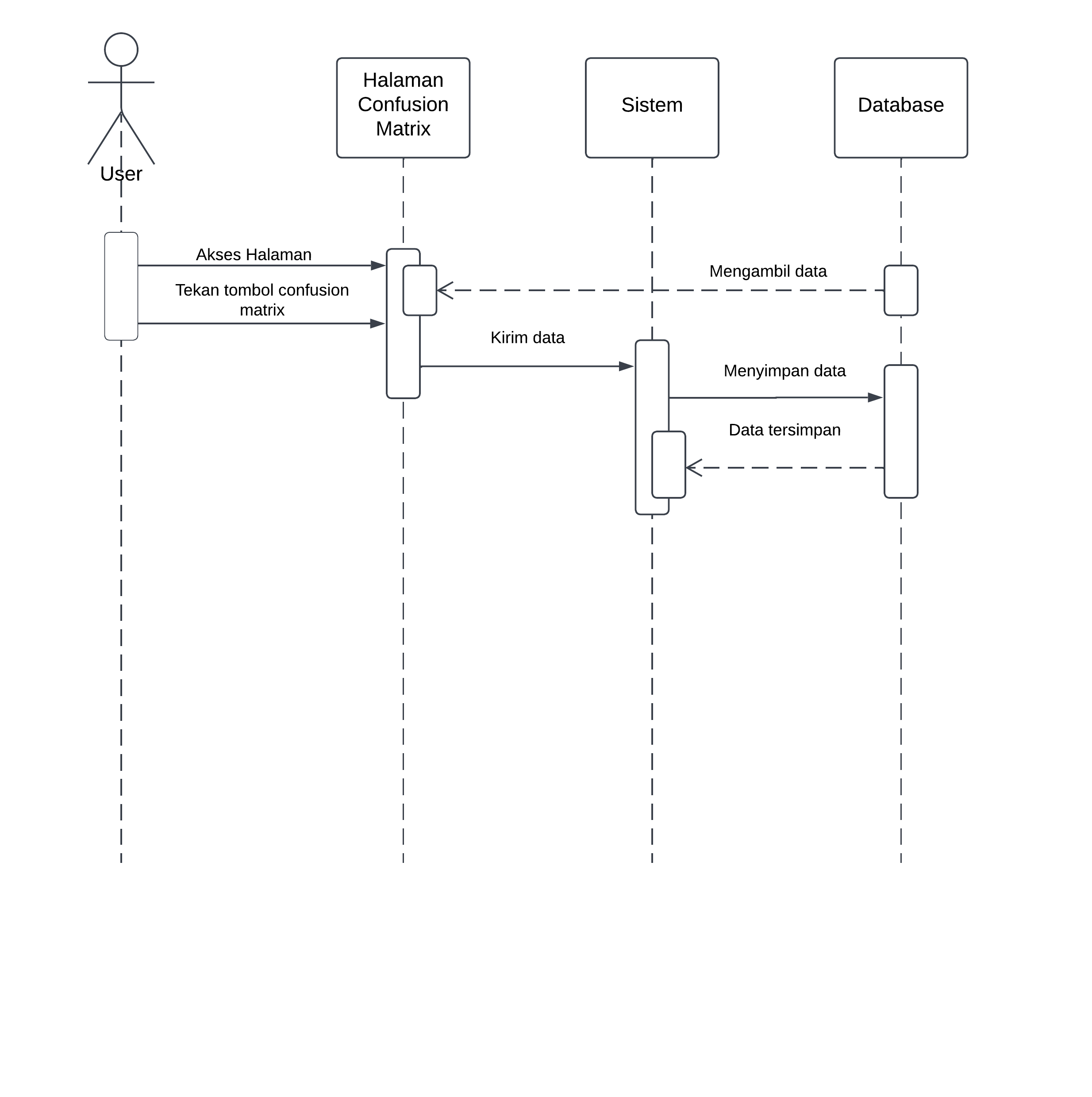
##### Sequence Diagram Frekuensi Kata



Gambar 3. 14 Sequence Diagram Frekuensi Kata

Sequence diagram frekuensi kata menjelaskan tentang urutan yang dilakukan untuk menghitung probabilitas positif atau negatif untuk setiap kata.

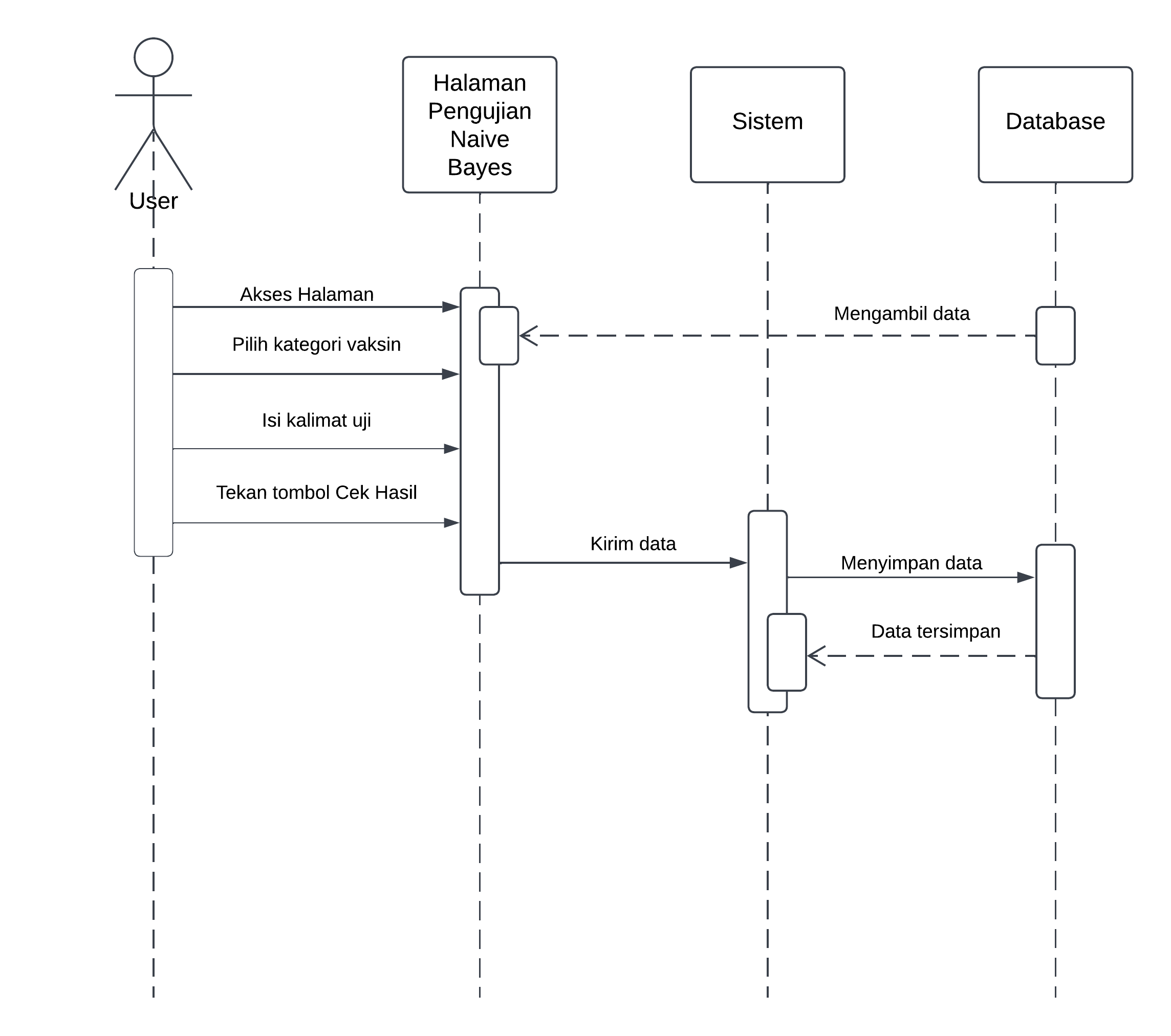
##### Sequence Diagram Confusion Matrix



Gambar 3. 15 Sequence Diagram Confusion Matrix

Sequence diagram *confusion matrix* menjelaskan tentang urutan yang dilakukan untuk menghitung *confusion matrix*, *Accuracy, Precision, Recall* dan *F-Measure*.

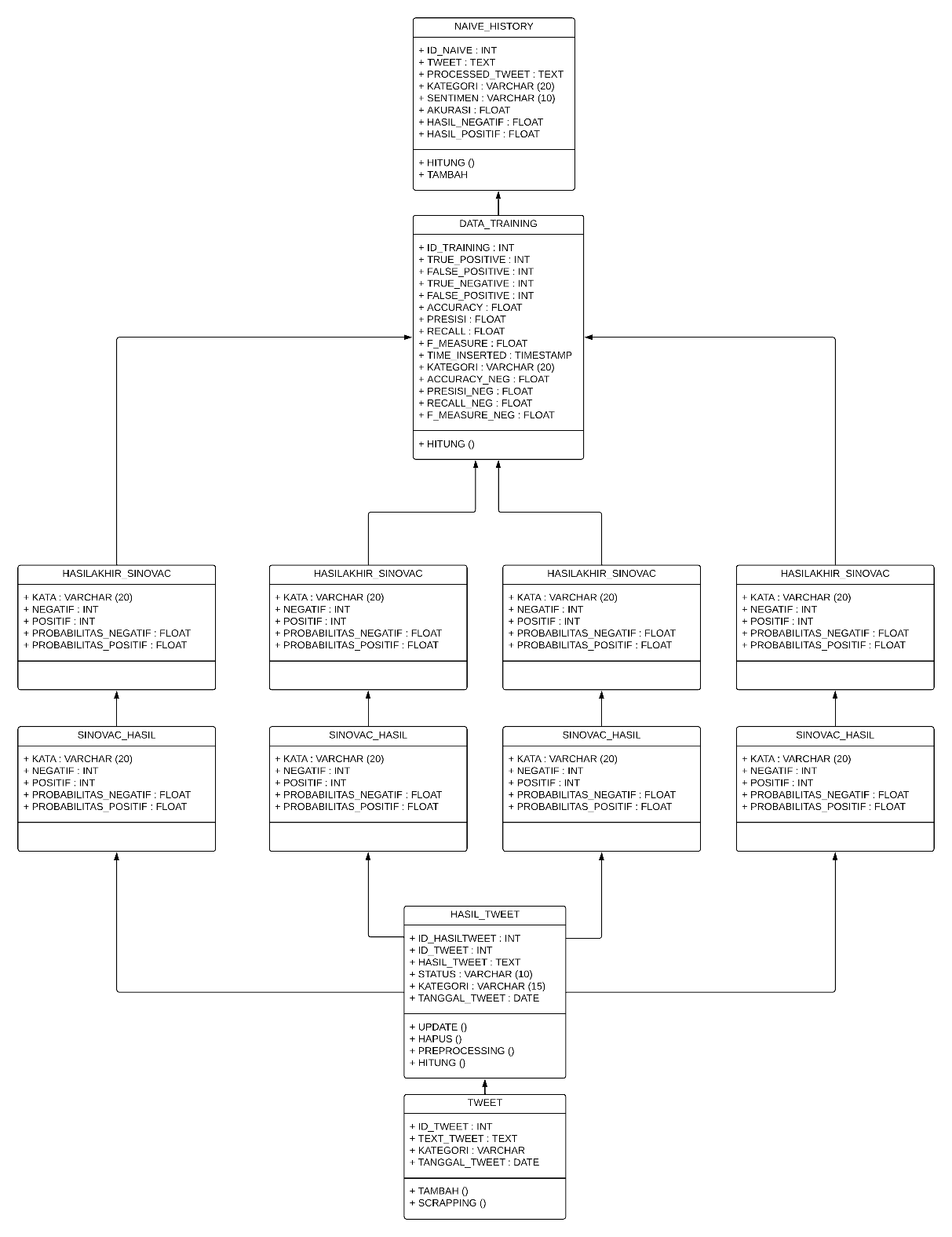
##### Sequence Diagram Pengujian Algoritma



Gambar 3. 16 Sequence Diagram Pengujian Algoritma

Sequence diagram pengujian algoritma menjelaskan tentang urutan yang dilakukan untuk menguji hasil sentimen sebuah kalimat *tweet*.

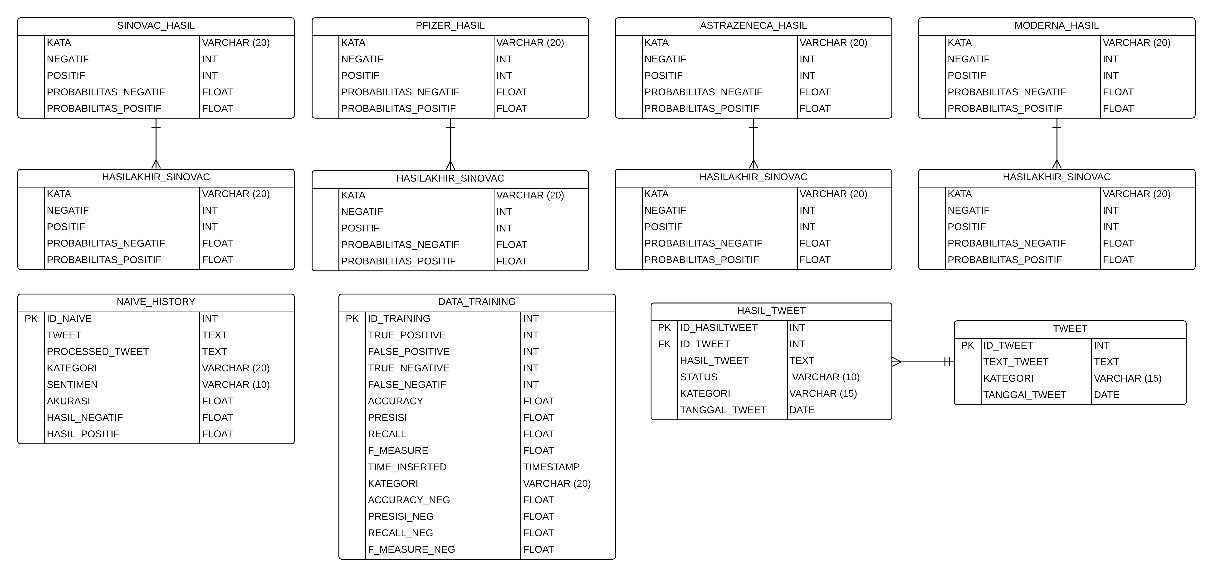
#### Class Diagram



Gambar 3. 17 Class Diagram

### Perancangan Data

#### Desain Database



Gambar 3. 18 Desain Database

#### Desain Tabel

Tabel 3. 1 Tabel tweet

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
| PRIMARY KEY | ID\_TWEET | INT |
|  | TEXT\_TWEET | TEXT |
|  | KATEGORI | VARCHAR (15) |
|  | TANGGAL\_TWEET | DATE |

Tabel 3. 2 Tabel hasil\_tweet

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
| PRIMARY KEY | ID\_HASILTWEET | INT |
| FOREIGN KEY | ID\_TWEET | INT |
|  | HASIL\_TWEET | TEXT |
|  | STATUS | VARCHAR (10) |
|  | KATEGORI | VARCHAR (15) |
|  | TANGGAL\_TWEET | DATE |

Tabel 3. 3 Tabel sinovac\_hasil

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 4 Tabel astrazeneca\_hasil

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 5 Tabel pfizer\_hasil

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 6 Tabel moderna\_hasil

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 7 Tabel hasilakhir\_sinovac

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 8 Tabel hasilakhir\_astrazeneca

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 9 Tabel hasilakhir\_pfizer

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 10 Tabel hasilakhir\_moderna

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
|  | KATA | VARCHAR (20) |
|  | NEGATIF | INT |
|  | POSITIF | INT |
|  | PROBABILITAS\_NEGATIF | FLOAT |
|  | PROBABILITAS\_POSITIF | FLOAT |

Tabel 3. 11 Tabel data\_training

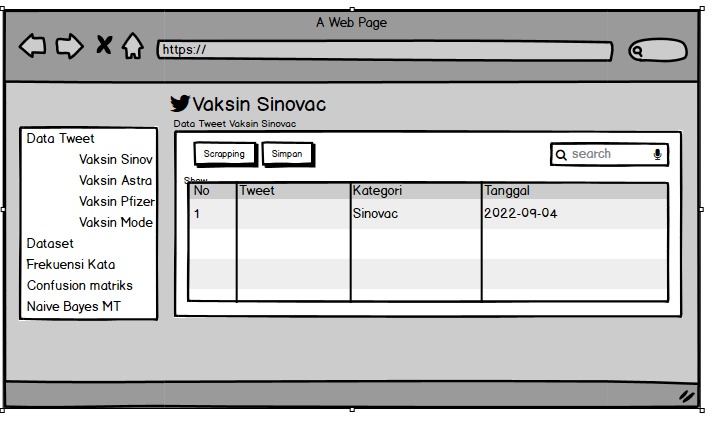
|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
| PRIMARY KEY | ID\_TRAINING | INT |
|  | TRUE\_POSITIVE | INT |
|  | FALSE\_POSITIVE | INT |
|  | TRUE\_NEGATIVE | INT |
|  | FALSE\_NEGATIVE | INT |
|  | ACCURACY | FLOAT |
|  | PRESISI | FLOAT |
|  | RECALL | FLOAT |
|  | F\_MEASURE | FLOAT |
|  | TIME\_INSERTED | TIMESTAMP |
|  | KATEGORI | VARCHAR (20) |
|  | ACCURACY\_NEG | FLOAT |
|  | PRESISI\_NEG | FLOAT |
|  | RECALL\_NEG | FLOAT |
|  | F\_MEASURE\_NEG | FLOAT |

Tabel 3. 12 Tabel naïve\_history

|  |  |  |
| --- | --- | --- |
| INDEX | NAMA | TIPE DATA |
| PRIMARY KEY | ID\_NAIVE | INT |
|  | TWEET | TEXT |
|  | PROCESSED\_TWEET | TEXT |
|  | KATEGORI | VARCHAR (20) |
|  | SENTIMEN | VARCHAR (10) |
|  | AKURASI | FLOAT |
|  | HASIL\_NEGATIF | FLOAT |
|  | HASIL\_POSITIF | FLOAT |

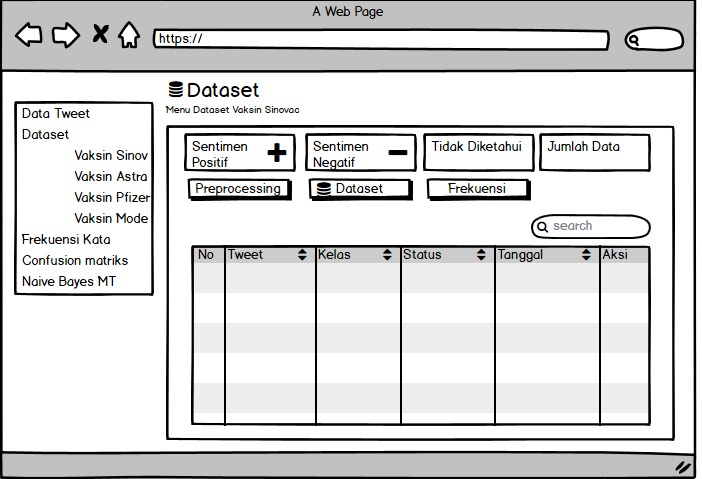
### Perancangan User Interface / *Mock-up* aplikasi

#### User Interface Halaman Data Tweet



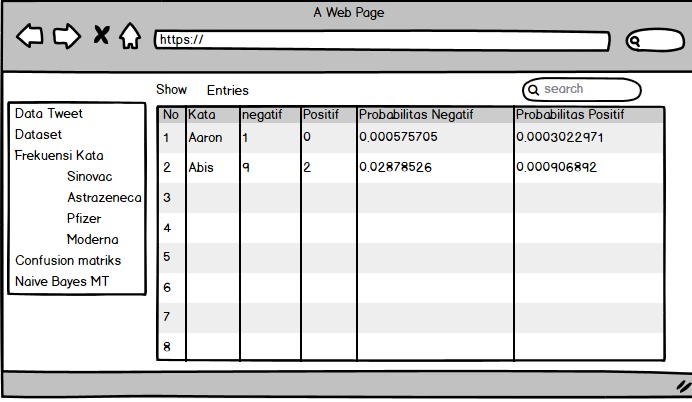
Gambar 3. 19 User Interface Halaman Data Tweet

#### User Interface Halaman Dataset



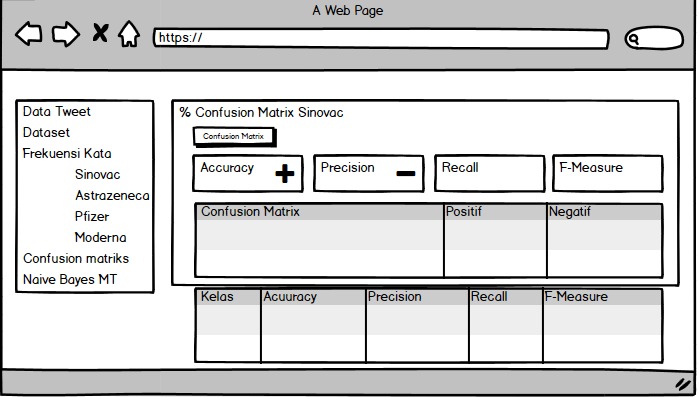
Gambar 3. 20 User Interface Halaman Dataset

#### User Interface Halaman Frekuensi Kata



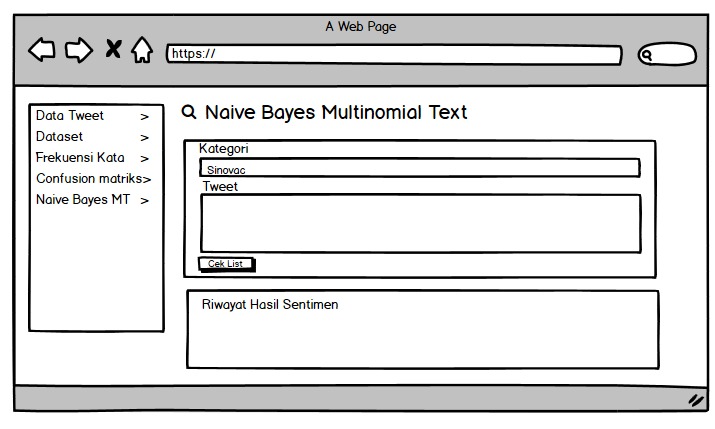
Gambar 3. 21 User Interface Halaman Frekuensi Kata

#### User Interface Halaman Confusion Matrix



Gambar 3. 22 User Interface Confusion Matrix

#### User Interface Halaman Pengujian Kalimat Tweet



Gambar 3. 23 User Inteface Halaman Pengujian

## Rancangan Pengujian

Tahap rancangan pengujian terdiri dari dua bagian, yaitu pengujian keberhasilan penerapan algoritma dan perbandingan pengujian akurasi sistem dengan *software* WEKA.