

BAB III

ANALISA DAN PERANCANGAN SISTEM

Pada bab ini penulis akan menjelaskan tahapan-tahapan analisa dan perncangan sistem yang meliputi deskripsi sistem. Selain itu pada bab ini non fungsional yakni kebutuhan perangkat keras dan perancangan interface. menjelaskan tentang kebutuhan fungsional yakni kebutuhan perangkat lunak, kebutuhan.

3.1. Jenis Pendekatan Penelitian

Pada penelitian kali ini, peneliti menggunakan pendekatan kualitatif. Pada pendekatan kualitatif, peneliti lebih menekankan pada pemahaman mengenai permasalahan dalam kehidupan sehari-hari berdasarkan kondisi realitas. Hal ini menyebabkan bentuk desain yang bervariasi, fleksibel, dan dapat diubah sesuai dengan rencana yang sudah dibuat sebelumnya (Lestari, 2017).

3.2. Kebutuhan Non Fungsional

analisa kebutuhan non fungsional dilakukan untuk mengetahui perangkat yang digunakan untuk membuat web analisis sentimen pengguna twitter mengenai pemilihan presiden tahun 2024 menggunakan metode Support Vector Machine , kebutuhan non fungsional dibedakan menjadi 2, yaitu:

1. Perangkat keras (*hardware*)

Perangkat keras atau hardware merupakan alat yang berfungsi untuk memberikan masukan, mengolah dan menampilkan keluaran. Hardware juga berperan sebagai eksekutor dari perintah yang telah di programkan melalui perangkat lunak yang ditanamkan didalamnya.

Spesifikasi hardware yang digunakan dalam membangun pembuatangame ini adalah sebagai berikut:

- a. Processor 2,5 GHz Intel core i5
- b. RAM 6 GB
- c. HDD 500 GB
- d. VGA Intel HD Graphics 4000

2. Perangkat lunak (*software*)

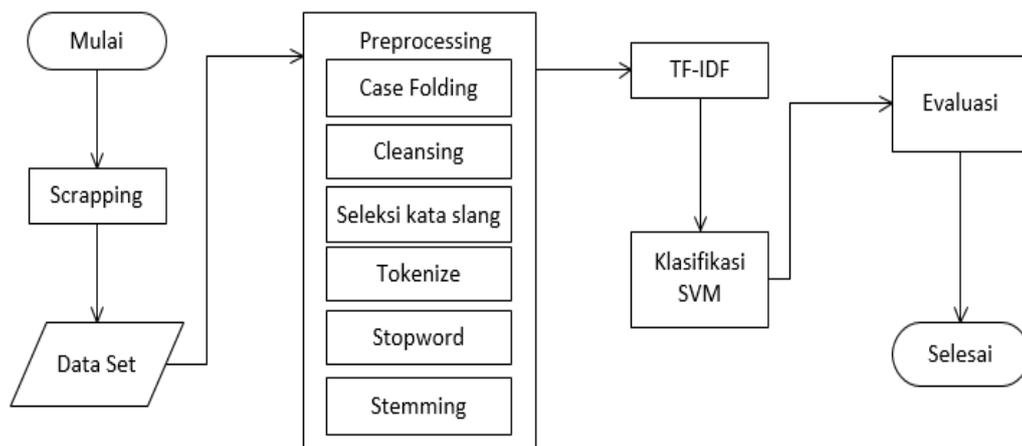
Perangkat lunak atau software merupakan kumpulan perintah atau instruksi yang dijalankan oleh pengguna computer untuk diteruskan atau di proses oleh perangkat keras. Software berfungsi sebagai sarana interaksi antara pengguna dengan perangkat keras.

- a. Windows 7/10
- b. Microsoft Word
- c. Microsoft Excel
- d. Jupyter Notebook

3.3. Perancangan sistem

3.3.1. Alur Sistem

Alur sistem adalah sebuah tahapan dari awal hingga akhir yang dibuat berdasarkan pemikiran peneliti dalam melakukan sebuah penelitian dan dibuat kedalam bentuk flowchart . Berikut alur sistem penulis adalah sebagai berikut.



Gambar 3. 1 Alur Sistem

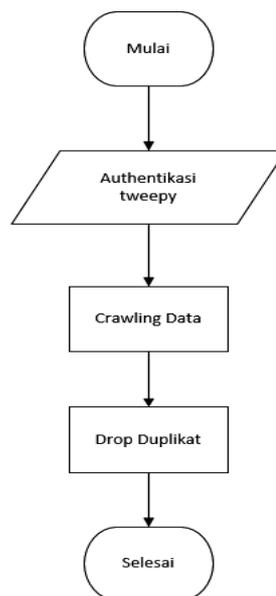
3.3.2. Data Set

Pada tahap pertama ini, penulis mengumpulkan data yang akan digunakan dalam penelitian dengan menyesuaikan berdasarkan topik yang dipilih. Dikarenakan penelitian menggunakan salah satu media sosial twitter, penulis akan

menggunakan teknik crawling untuk pengumpulan data. Alasan penulis menggunakan media sosial tersebut karena twitter menyediakan sebuah website untuk para developer yaitu <https://developer.twitter.com/en>.

Crawling dilakukan dengan menggunakan bahasa python dengan library sncrape yang digunakan untuk mengambil data tweet dari twitter menggunakan web scraping. sncrape menggunakan Teknik web scraping untuk mengambil data tweet dari halaman pencarian twitter. Itu secara otomatis mengekstrak konten tweet, tanggal, pengguna, dan atribut lainnya dari halaman HTML.

Hasil dari crawling berupa tweet berbahasa indonesia dengan kata kunci dari Ganjar pranowo, Prabowo subianto, Anis baswedan, Puan Maharani dan Muhaimin Iskandar Penulis berhasil mengumpulkan data dari kata kunci tersebut sebanyak 25.000 data dengan alur sebagai berikut.



Gambar 3. 2 Flowchart Crawling data

3.3.3. Preprocessing

Dalam penelitian *text mining* tahapan preprocessing perlu dilakukan. Penulis melakukan tahapan preprocessing ini menggunakan bahasa python, dengan tahapan *case folding*, *cleansing*, *Seleksi kata slang*, *tokenize*, *stopword*, dan *stemming*. Hasil dari tahapan ini nanti akan digunakan sebagai data pembelajaran

mesin oleh SVM.

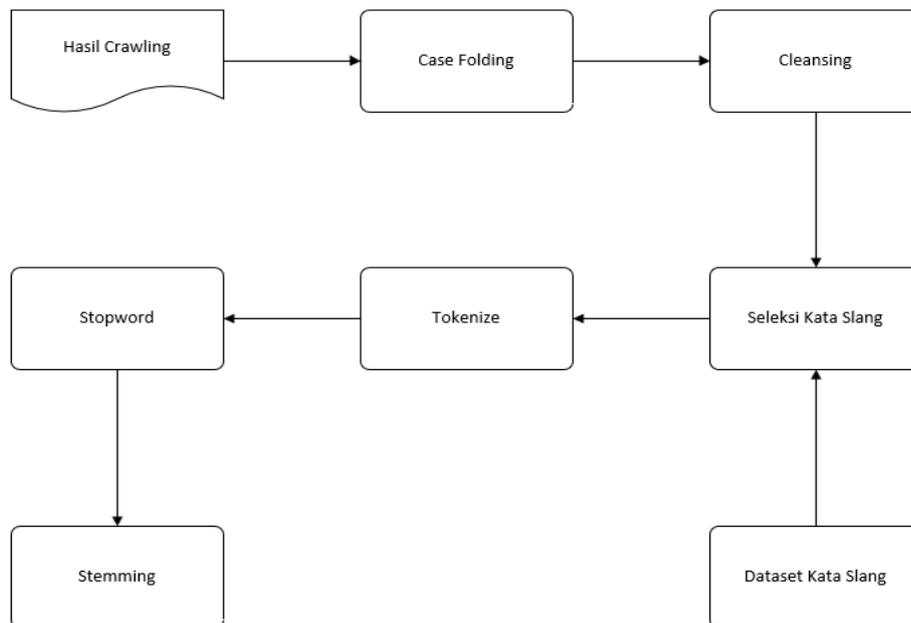
Tujuan dari preprocessing ini adalah sebagai berikut.

1. Menghilangkan noise
2. Menyetarakan bentuk kata
3. Mengurangi volume kata

Hal yang dilakukan penulis dalam tahapan ini adalah meng-*import library* yang akan digunakan dalam tahapan ini dan membuat alur yang digunakan dalam tahapan ini. Berikut library yang dijabarkan dalam bentuk tabel dan alur pada preprocessing.

Tabel 3. 1 Library preprocessing yang digunakan

Nama <i>Library</i>	Fungsi <i>Library</i>
<i>Regular expression</i> (RE)	Digunakan untuk melakukan cleansing dengan cara memeriksa apakah suatu string tersebut sesuai dengan pola yang dibuat atau Tidak
NLTK	Sebuah <i>library</i> yang biasa digunakan untuk analisis teks. <i>Library</i> ini memiliki kamus stopword bahasa indonesia beserta tanda baca. Selain stopword library ini dapat digunakan untuk tahapan <i>tokenize</i>
<i>StemmerFactory</i>	Library untuk melakukan <i>stemming</i> . Library yang digunakan yaitu <i>stemming</i> sastrawi yang dapat diakses di <i>StemmerFactory</i>



Gambar 3. 3 Alur Preprocessing

Berikut contoh dari setiap proses yang dilakukan untuk melakukan tahapan preprocessing.

1. Case Folding

Tabel 3. 2 Contoh case folding

Tweet	Hasil Case Folding
Sosok menteri BUMN @erickthohir menjadi figure cawapres yg terus menguat untuk pilpres 2024	sosok menteri bumn @erickthohir menjadi figure cawapres yg terus menguat untuk pilpres 2024

Tahap pertama dalam preprocessing ini merubah seluruh tweet menjadi huruf kecil.

2. Cleansing

Tabel 3. 3 Contoh cleansing

Tweet	Hasil Cleansing
sosok menteri bumn @erickthohir menjadi figure cawapres yg terus menguat untuk pilpres 2024	sosok menteri bumn menjadi figure cawapres yg terus menguat untuk pilpres

Pada tahapan cleansing penulis membuat pola untuk menghilangkan mention, hastag, url, tanda baca, angka, dll.

3. Seleksi kata slang

Tabel 3. 4 Contoh seleksi kata slang

Tweet	Hasil Seleksi Kata Slang
sosok menteri bumn menjadi figure cawapres yg terus menguat untuk pilpres	sosok menteri bumn menjadi figure cawapres yang terus menguat untuk pilpres

Pada tahapan ini penulis menggunakan dataset bahasa alay milik Salsabila. Pada kamus tersebut terdapat 2 kolom yaitu kata alay dan kata berdasarkan kamus bahasa indonesia. Dengan kamus ini merubah kata slang seperti pada tabel 4.4 yang terdapat kata “yg” diubah menjadi “yang”.

4. Tokenize

Tabel 3. 5 Contoh Tokenize

Tweet	Hasil Tokenize
sosok menteri bumn menjadi figure cawapres yang terus menguat untuk pilpres	[sosok, mentri, bumn, menjadi, figure, cawapres, yang, terus, menguat, untuk, pilpres]

Pada tokenize ini memisahkan kalimat tweet menjadi kata per kata.

5. Stopword

Tabel 3. 6 Contoh Stopword

Tweet	Hasil Stopword
[sosok, menteri, bumn, menjadi, figure, cawapres, yang, terus, menguat, untuk, pilpres]	[sosok, menteri, bumn, figure, cawapres, menguat, pilpres]

Pada tahapan stopwords ini menghapus kata – kata yang dianggap tidak penting seperti yang, juga, dari, dia, kami, kamu, aku.

6. Stemming

Tabel 3. 7 Contoh Stemming

Tweet	Hasil Stemming
[sosok, menteri, bumn, figure, cawapres, menguat, pilpres]	[sosok, menteri, bumn, figure, wapres, menguat, pilpres]

Tahapan terakhir dalam preprocessing ini berfungsi untuk merubah kata yang sudah sesuai dengan kamus bahasa Indonesia menjadi kata dasar. Seperti pada tabel 4.7 merubah kata “cawapres” menjadi “capres”.

3.3.4. Model training

Pada tahapan ini model dirancang untuk mengklasifikasi sentimen yang telah melalui tahapan – tahapan sebelumnya. Berikut alur pada tahapan ini.

1. Pembobotan kata TF-IDF

Tahapan ini merupakan tahapan pembobotan kata dimana menghitung banyaknya term atau kata yang muncul dari hasil tweet. Pada tahapan ini menggunakan library sklearn maka dapat digunakan rumus pada bab2 pembahasan tf-idf. Berikut potongan kode program dalam melakukan pembobotan kata.

Tabel 3. 8 Contoh hasil pembobotan TF-IDF

TREM	TF				DF	TF Normalisasi				IDF	TF.IDF			
	d1	d2	d3	d4		d1	d2	d3	d4		d1	d2	d3	d4
kejut	1	0	0	0	1	0,167	0	0	0	0,602	0,101	0	0	0
pdip	1	0	1	0	2	0,167	0	0,125	0	0,301	0,05	0	0,038	0
ajak	1	0	0	0	1	0,167	0	0	0	0,602	0,101	0	0	0
demokrat	1	0	0	0	1	0,167	0	0	0	0,602	0,101	0	0	0
kerjasama	1	0	0	0	1	0,167	0	0	0	0,602	0,101	0	0	0
pilpres	1	1	1	1	4	0,167	0,143	0,125	0,25	0	0	0	0	0
sidoarjo	0	1	0	0	1	0	0,143	0	0	0,602	0	0,086	0	0
nih	0	1	0	0	1	0	0,143	0	0	0,602	0	0,086	0	0
deklarasi	0	1	0	0	1	0	0,143	0	0	0,602	0	0,086	0	0
dukung	0	1	0	0	1	0	0,143	0	0	0,602	0	0,086	0	0
ganjar	0	1	1	0	2	0	0,143	0,125	0	0,301	0	0,043	0,038	0
pranowo	0	1	1	0	2	0	0,143	0,125	0	0,301	0	0,043	0,038	0
kata	0	0	1	0	1	0	0	0,125	0	0,602	0	0	0,075	0
politis	0	0	1	0	1	0	0	0,125	0	0,602	0	0	0,075	0
mungkin	0	0	1	0	1	0	0	0,125	0	0,602	0	0	0,075	0
kalah	0	0	1	0	1	0	0	0,125	0	0,602	0	0	0,075	0
gus	0	0	0	1	1	0	0	0	0,25	0,602	0	0	0	0,151
muhaimin	0	0	0	1	1	0	0	0	0,25	0,602	0	0	0	0,151
tentu	0	0	0	1	1	0	0	0	0,25	0,602	0	0	0	0,151
panjang d	6	7	8	4										

Penggunaan TF-IDF dalam analisis sentimen dengan SVM membantu model untuk:

1. Mengurangi bobot kata-kata umum yang tidak relevan.
2. Mengidentifikasi kata-kata penting yang unik untuk setiap dokumen.
3. Menyeimbangkan frekuensi kata-kata agar kata-kata penting yang muncul lebih jarang tetap diberi bobot tinggi.
4. Meningkatkan representasi teks menjadi bentuk numerik yang lebih informatif.

2. Train-Split-data

Pada tahapan ini dilakukan pembagian data menjadi 2 yaitu data uji dan data latih. Data latih akan digunakan untuk melatih model untuk klasifikasi sedangkan data uji akan digunakan untuk menguji performa yang telah dibuat oleh data latih dengan 80% data uji dan 20% data latih.

3. Klasifikasi

Tahapan ini merupakan tahapan terakhir dari model training. Skenario yang dibuat akan digunakan untuk klasifikasi menggunakan Support vector machine.

Kernel yang digunakan pada tahapan ini adalah linier SVM.

3.4. Rancangan Interface

Rancangan interface ini menggambarkan desain website yang akan ditampilkan pada halaman web sesuai dengan apa yang diinginkan. Yakni desain interface yang berisikan beberapa menu seperti upload data ,preprocessing, tf-Idf, klasifikasi svm, Grafik dan tes model.

3.4.1. Interface Halaman Utama

Interface halaman ini menampilkan desain website yang akan ditampilkan pada halaman utama.

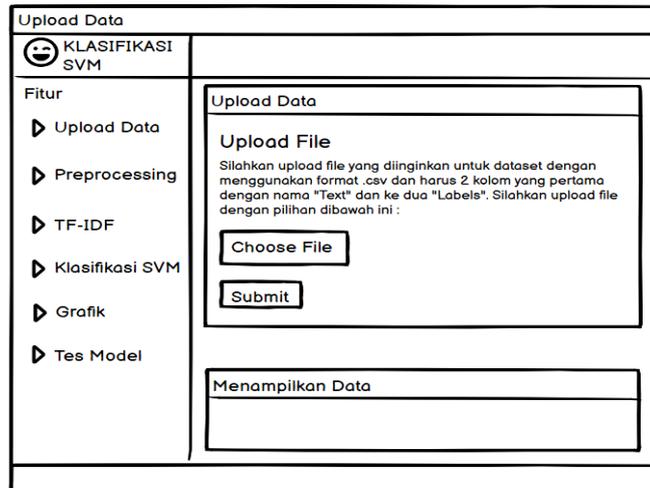
Klaifikasi SVM	
 KLASIFIKASI SVM	
Fitur	ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PEMILIHAN PRESIDEN TAHUN 2024 MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)
▶ Upload Data	
▶ Preprocessing	
▶ TF-IDF	
▶ Klasifikasi SVM	
▶ Grafik	
▶ Tes Model	

Gambar 3. 4 Interface Utama

Pada Gambar 3.4 ini menggambarkan beberapa menu yang bisa diakses. seperti menu upload data, preprocessing, tf-idf, klasifikasi SVM dan menu tes model.

3.4.2. Interface Halaman Upload Data

Interface halaman ini menampilkan desain website dari halaman upload data yang akan ditampilkan pada halaman upload data yang bisa di akses oleh *user*.

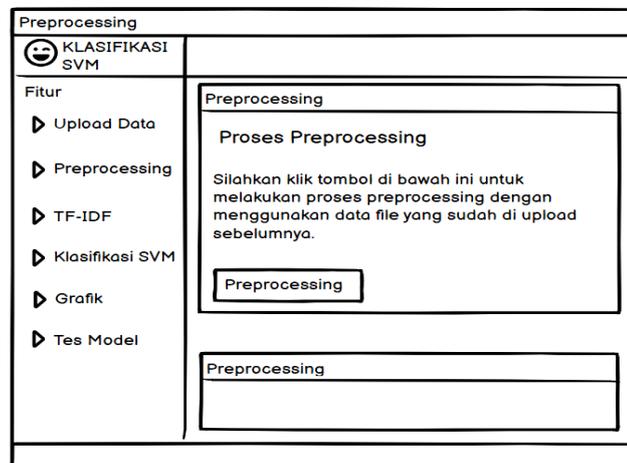


Gambar 3. 5 Interface upload data

Pada Gambar 3.5 ini *user* harus memasukkan data berupa file csv lalu data akan di tampilkan dibagian bawah.

3.4.3. Interface Halaman Preprocessing

Setelah mengimport data berupa csv, maka selanjutnya menuju ke halaman Preprocessing, berikut merupakan gambaran dari halaman preprocessing.

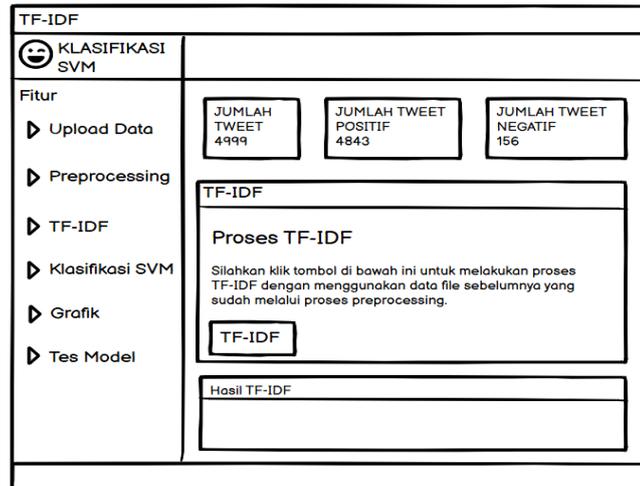


Gambar 3. 6 Interface preprocessing

Pada Gambar 3.6 ini akan dilakukan proses Preprocessing dan hasil Preprocessing akan di tampilkan di bagian bawah.

3.4.4. Interface Halaman TF-IDF

Setelah melakukan proses Preprocessing selanjutnya menuju halaman TF-IDF, berikut merupakan gambaran dari halaman TF-IDF.

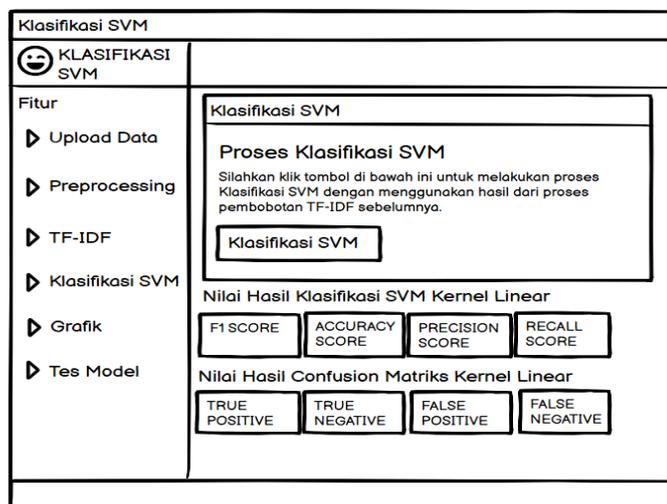


Gambar 3. 7 Interface TF-IDF

Pada Gambar 3.7 ini akan dilakukan proses perhitungan TF-IDF dan hasil dari perhitungan TF-IDF akan di tampilkan di bagian bawah.

3.4.5. Interface Halaman Klasifikasi SVM

Setelah melakukan perhitungan TF-IDF selanjutnya menuju halaman klasifikasi SVM, berikut merupakan Gambaran dari halaman klasifikasi SVM.



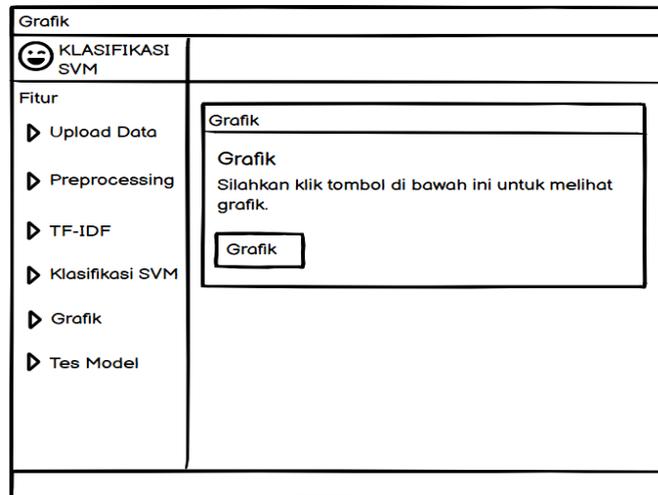
Gambar 3. 8 Interface klasifikasi SVM

Pada Gambar 3.8 ini akan dilakukan proses klasifikasi dengan metode SVM, hasil klasifikasi akan di tampilkan di bagian bawah.

3.4.5. Interface Halaman Grafik

Setelah melakukan klasifikasi SVM selanjutnya melihat grafik batang.

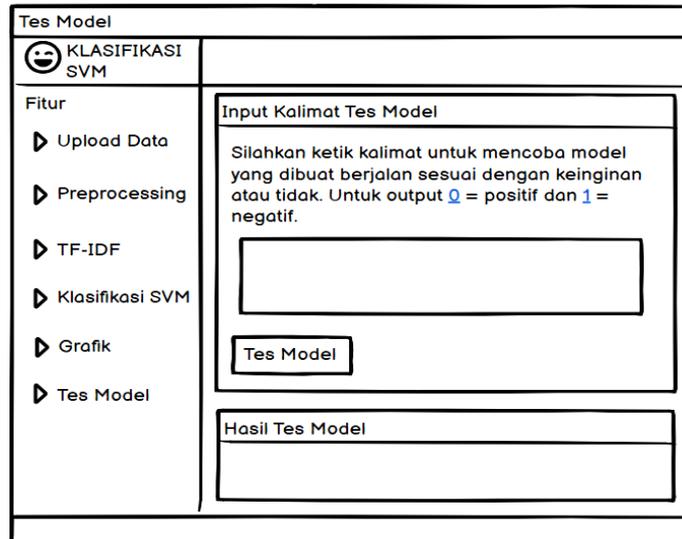
Berikut merupakan gambaran dari halaman Grafik.



Gambar 3. 9 Interface Grafik

3.4.6. Interface Halaman Tes Model

Setelah melihat Grafik selanjutnya menuju halaman tes model, berikut merupakan gambaran dari halaman tes model.



Gambar 3. 10 Interface Tes Model

Pada Gambar 3.10 ini akan dilakukan pengetesan model dengan memasukkan kalimat baru, hasil pengetesan model akan ditampilkan di bagian bawah.