

BAB 3

ANALISIS DAN PERANCANGAN

3.1 Analisis Masalah

Berdasarkan latar belakang, banyaknya pengguna media sosial facebook menyebabkan banyak juga yang membuat komentar berdasarkan suatu postingan. Setiap pengguna memiliki komentar yang beragam mulai dari komentar yang membangun hingga komentar *bully*. Untuk mendeteksi komentar yang mengandung unsur *bullying* dapat menggunakan *Machine Learning* yaitu dengan Analisis Sentimen.

Cara untuk mendeteksi komentar yang mengandung unsur *bullying* dapat menggunakan pendekatan *machine learning* yaitu dengan analisis sentimen. Analisis sentimen sangat diperlukan untuk menyaring komentar-komentar di media sosial. Proses yang dilakukan dalam analisis sentimen adalah dengan mengklasifikasikan informasi ke dalam kelas sentimen positif dan kelas sentimen negatif. Informasi akan diklasifikasikan ke dalam kelas positif apabila informasi yang disampaikan bernilai baik atau setuju terhadap sesuatu. Sebaliknya, informasi diklasifikasikan ke dalam kelas negatif apabila informasi yang disampaikan bernilai tidak baik atau tidak setuju[1].

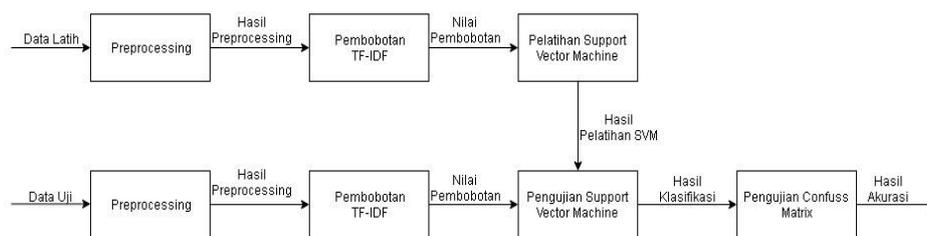
Untuk mengklasifikasikan komentar maka dibutuhkan pendekatan *machine learning* yang dapat memisahkan antara komentar yang mengandung *cyberbullying* dan tidak mengandung *cyberbullying*. Oleh karena itu, dipilihlah algoritma *support vector machine* untuk penelitian ini.

Pada penelitian sebelumnya yang berguna dalam mendukung pelaksanaan penelitian terkait adalah penelitian mengenai analisis sentimen yang menggunakan metode *Support Vector Machine* oleh Petrik[2]. Penelitian lain membahas mengenai kategorisasi teks Bahasa Indonesia oleh Wulandini dan Nugroho, didapat algoritma SVM memiliki akurasi yang paling tinggi yaitu 92,5 % dibanding algoritma yang lain seperti *K-Nearest Neighbors*, *Naïve Bayes Classification*, *Information Fuzzy Networks*[3].

Didapatkan kesimpulan bahwa penggunaan metode *Support Vector Machine* (SVM) memberikan hasil akurasi paling baik dibandingkan metode lainnya. Berdasarkan hal itu dalam penelitian ini akan digunakan metode metode *Support Vector Machine* untuk mendeteksi adanya sentimen serta mengetahui nilai akurasi pada metode yang digunakan.

3.2 Analisis Proses

Analisis proses merupakan tahapan untuk menganalisis suatu cara atau menganalisis metode-metode yang digunakan. Tahapan yang digunakan untuk mengetahui adanya kalimat yg mengandung bullying dalam teks pada penelitian ini dibagi menjadi dua tahap, yaitu tahap pelatihan dan pengujian. Berikut gambaran tahapan proses yang akan dilakukan pada gambar 3.1.



Gambar 3.1 Tahapan Proses yang Dilakukan

1. Tahap pertama yang dilakukan adalah pengumpulan dataset. Dalam proses pelatihan data latih yang berupa data komentar dari *Facebook* akan diberi label lalu melalui beberapa proses utama yaitu proses *preprocessing*, TF-IDF, *Lexicon Based Features* lalu klasifikasi *Support Vector Machine* tahap pelatihan. Proses *preprocessing* itu sendiri terdiri dari enam proses yaitu *Case Folding*, *Cleansing*, *Normalisasi Bahasa*, *Convert Negation*, *Stopword Removal*, *Tokenisasi*. Setelah *preprocessing*, dilakukan proses pembobotan kata dengan menggunakan pembobotan TF-IDF dengan menggunakan data hasil *preprocessing*. Pada tahap klasifikasi *Support Vector Machine* tahap pelatihan dilakukan untuk mendapatkan *hyperplane* yang didapat dari data latih yang telah dimasukkan kemudian data *hyperplane* tersebut dimasukkan ke dalam database.

2. Tahap selanjutnya yang dilakukan adalah tahap pengujian. Data uji dimasukkan melalui proses yang sama seperti data latih yaitu proses *preprocessing*, Pembobotan TF-IDF, lalu hasil dari pengujian ini adalah klasifikasi sentimen berupa komentar facebook dari data uji yang dimasukkan. Lalu untuk pengujian sistem dilakukan metode *confuss matrix*. Hasil dari pengujian tersebut akan menghasilkan akurasi dari metode *Support Vector Machine*.

3.3 Analisis Data Masukan

Data masukan yang digunakan dalam penelitian ini adalah komentar di facebook dari postingan facebook yang telah ditentukan. Data masukan diperoleh dengan cara:

1. Data komentar diperoleh dengan menggunakan *tools* Microsoft Excel; yaitu Power Query untuk dilakukannya *scrapping* komentar pada kiriman *fanspage* facebook “*Sepakbola Indonesia Fansbook*”.
2. Data komentar yang diperoleh disimpan dengan format .csv
3. Data yang diproses merupakan data yang berbahasa Indonesia.
4. Terdapat 150 data komentar yang akan digunakan, Dimana terdiri dari, 100 data komentar sebagai data latih, dan 50 data komentar menjadi data uji.

Data yang digunakan dalam penelitian ini terdiri dari dua jenis data; yaitu data latih dan data uji. Berikut adalah sampel data dari latih dan uji yang akan dilakukan dalam penelitian ini pada Tabel 3.1 dan Tabel 3.2.

Tabel 3.1 Contoh Data Latih

Data Latih	Kategori	Komentar
P1	Negatif	Wasit indonesia sulit fokus kalo udah dapet tekanan heu ! semoga mental wasit juga bisa diperbaiki padahal fase grup mantap
P2	Positif	Kejadian yg diprotes keras oleh ps tira mnurut opiniku wasit udah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan Terimakasih wasit
Data Latih	Kategori	Komentar
P3	Negatif	Wkwkwk orang goblog yg ga ngerti bola pada ngamuk semua Sebenarnya saya bukan fans tim manapun
P4	Negatif	Dear The jak & aremania wasit salah apa coba ? Penalti diving ? Jelas pemain lawan sudah bermain sangat keras Jangan hujat WASIT
P5	Positif	Beberapa menit terakhir cukup menyakitkan bagi tibo Selamat Persebaya dan wasit harus lebih baik lagi ya kedepannya
P6	Negatif	Kenapa ga ditonjok aja itu wasitnya ? Kalo perlu sekalian aja kroyok ! Biar tau rasaaa
P7	Negatif	Ada dorongan terhadap Osvaldo Hay ya walaupun minor Harusnya wasit lebih jeli lagi biar ga kena protes pemain pemain yang ga ngeliat jelas
P8	Positif	Wasit sudah benar untuk masalah penalty ituAbduh memang melakukan pelanggaran sangat keras

Tabel 3.2 Contoh Data Uji

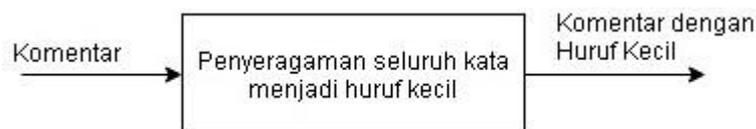
Data Uji	Kategori	Komentar
P9	Negatif	Pdhl jelas penalti dan tidak offside Jdi apanya yang harus diprotes ? Tolong !! dipake lagi ya matanya
P10	Positif	bodo amat apa kata netizen nyinyir Yang penting persebaya menang hahaha

3.4 Analisis Tahap *Preprocessing*

Analisis *preprocessing* merupakan tahap awal dan salah satu langkah yang penting dalam sebuah pengklasifikasian sebuah teks. Adapun tahapan *preprocessing* yang akan dilakukan pada penelitian ini yaitu *Case folding*, *Cleansing*, *Normalisasi Bahasa*, *Convert Negation*, *Stopword Removal*, *Tokenisasi*. *Preprocessing* ini dilakukan pada data latih yang menjadi data masukan, berikut penjelasan dari tahapan *preprocessing* yang akan dilakukan:

3.4.1 Case Folding

Pada tahap ini dilakukan perubahan pada huruf kapital dalam dokumen menjadi huruf kecil dengan maksud untuk menyeragamkan karakter dalam dokumen. Seperti yang terdapat pada gambar 3.2.

**Gambar 3.2 Block Diagram Case Folding**

Adapun contoh tahap *Case Folding* diterapkan pada data latih dan data uji dapat dilihat pada Tabel 3.3. Berikut ini adalah penjelasan dari tabel 3.3 :

- a. Cek huruf kapital dari setiap data latih mulai dari data latih P1 sampai data latih P6.
- b. Bila ada huruf kapital yang terdapat pada data latih akan diubah menjadi huruf kecil.

Contoh : pada data latih P4 kata “WASIT” akan diubah yang semula tertulis semua dengan huruf kapital menjadi huruf kecil yaitu “semangat”.

Dan hasil case folding terdapat pada Tabel 3.3

Tabel 3.3 Hasil Case Folding

Sebelum Case Folding	Setelah Case Folding
Wasit indonesia sulit fokus kalo udah dapet tekanan heu ! semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	wasit indonesia sulit fokus kalo udah dapet tekanan heu ! semoga mental wasit juga bisa diperbaiki padahal fase grup mantap
Kejadian yg diprotes keras oleh ps tira mnurut opiniku wasit udah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan Terimakasih wasit	kejadian yg diprotes keras oleh ps tira mnurut opiniku wasit udah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terimakasih wasit
Wkwkwk orang goblog yg ga ngerti bola pada ngamuk semua Sebenarnya saya bukan fans tim manapun	wkwkwk orang goblog yg ga ngerti bola pada ngamuk semua sebenarnya saya bukan fans tim manapun
Dear The jak & aremania wasit salah apa coba ? Penalti diving ? Jelas pemain lawan sudah bermain sangat keras Jangan hujat WASIT	dear the jak & aremania wasit salah apa coba ? penalti diving ? jelas pemain lawan sudah bermain sangat keras jangan hujat wasit
Beberapa menit terakhir cukup menyakitkan bagi tibo Selamat Persebaya dan wasit harus lebih baik lagi ya kedepannya	beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya
Kenapa ga ditonjok aja itu wasitnya ? Kalo perlu sekalian aja kroyok ! Biar tau rasaaa	kenapa ga ditonjok aja itu wasitnya ? kalo perlu sekalian aja kroyok ! biar tau rasaaa

3.4.2 Cleansing

Pada tahap ini, akan dilakukan penghapusan karakter simbol atau tanda baca selain dari karakter *alphabet a-z*, dengan tujuan untuk mengurangi noise dan menyeragamkan jenis karakter dalam dokumen tersebut, untuk menanggulangi kelebihan spasi setelah melewati tahapan cleansing, maka dibutuhkan penghapusan spasi yang berlebih (*remove whitespace*). Seperti yang terdapat pada gambar 3.3 berikut.



Gambar 3.3 Block Diagram Cleansing

Adapun hasil tahap *Cleansing* diterapkan pada data latih dan data uji dapat dilihat pada Tabel 3.4.

Tabel 3.4 Hasil Cleansing

Sebelum Cleansing	Setelah Cleansing
wasit indonesia sulit fokus kalo udah dapet tekanan heu ! semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	wasit indonesia sulit fokus kalo udah dapet tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap
kejadian yg diprotes keras oleh ps tira mnurut opiniku wasit udah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terimakasih wasit	kejadian yg diprotes keras oleh ps tira mnurut opiniku wasit udah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terimakasih wasit
wkwkwk orang goblog yg ga ngerti bola pada ngamuk semua sebenarnya saya bukan fans tim manapun	wkwkwk orang goblog yg ga ngerti bola pada ngamuk semua sebenarnya saya bukan fans tim manapun
dear the jak & aremania wasit salah apa coba ? penalti diving ? jelas pemain lawan sudah bermain sangat keras jangan hujat wasit	dear the jak aremania wasit salah apa coba penalti diving jelas pemain lawan sudah bermain sangat keras jangan hujat wasit
beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya	beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya
kenapa ga ditonjok aja itu wasitnya ? kalo perlu sekalian aja kroyok ! biar tau rasaaa	kenapa ga ditonjok aja itu wasitnya kalo perlu sekalian aja kroyok biar tau rasaaa

Berikut penjelasan tahap-tahap dan proses Cleansing:

- a. Cek karakter yang mengandung simbol pada data masukan. Misalnya pada pernyataan “kenapa ga ditonjok aja itu wasitnya ? kalo perlu sekalian aja kroyok ! biar tau rasaaa”. Maka tanda tanya(?) akan terdeteksi dan akan terhapus setelah proses cleansing.

3.4.3 Normalisasi Bahasa

Pada tahap *preprocessing* dilakukan normalisasi bahasa terhadap kata yang tidak baku. Tahapan ini bertujuan untuk mengembalikan bentuk penulisan dari masing-masing kata yang sesuai dengan kamus data yang dibuat. Proses ini dilakukan dengan mencocokkan setiap kata pada dokumen data latih dan data uji dengan kata yang ada pada kamus tidak baku.

Adapun contoh tahap Normalisasi Bahasa diterapkan pada data latih dan data uji dapat dilihat pada Tabel 3.5.

Tabel 3.5 Hasil Normalisasi Bahasa

Sebelum Normalisasi Bahasa	Kata Tidak Baku	Setelah Normalisasi Bahasa
wasit indonesia sulit fokus kalo udah dapet tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	kalo,kalau udah,sudah dapet,dapat	wasit indonesia sulit fokus kalau sudah dapat tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap
kejadian yg diprotes keras oleh ps tira mnurut opiniku wasit udah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terimakasih wasit	yg,yang udah,sudah	kejadian yang diprotes keras oleh ps tira mnurut opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terima kasih wasit
wkwkwk orang goblog yg ga ngerti bola pada ngamuk semua sebenarnya saya bukan fans tim manapun	wkwkwk,ketawa yg,yang ga,tidak ngamuk,marah	ketawa orang goblog yang tidak ngerti bola pada marah semua sebenarnya saya bukan fans tim manapun

Sebelum Normalisasi Bahasa	Kata Tidak Baku	Setelah Normalisasi Bahasa
dear the jak aremania wasit salah apa coba penalti diving jelas pemain lawan sudah bermain sangat keras jangan hujat wasit		dear the jak aremania wasit salah apa coba penalti diving jelas pemain lawan sudah bermain sangat keras jangan hujat wasit
beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya		beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya
kenapa ga ditonjok aja itu wasitnya kalo perlu sekalian aja kroyok biar tau rasaaa	ga,tidak aja,saja kalo,kalau kroyok,keroyok tau,tahu	kenapa tidak ditonjok saja itu wasitnya kalau perlu sekalian saja keroyok biar tahu rasaaa

3.4.4 Convert Negation

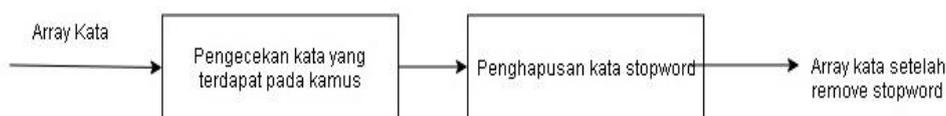
Convert Negation merupakan proses konversi kata negasi yang terdapat pada suatu komentar, karena kata negasi mempunyai pengaruh dalam merubah nilai sentimen pada suatu komentar. Jika terdapat kata negasi, maka kata tersebut akan disatukan dengan kata setelahnya. Kata negasi “tidak” akan disambung dengan kata selanjutnya. Contohnya kata “tidak ngerti” yang disambung menjadi kata “tidakngerti” maka nilainya menjadi negatif Adapun contoh tahap *Convert Negation* diterapkan pada data latih dan data uji dapat dilihat pada Tabel 3.6.

Tabel 3.6 Hasil Convert Negation

Sebelum <i>Convert Negation</i>	Kata Negasi	Setelah <i>Convert Negation</i>
wasit indonesia sulit fokus kalo sudah dapat tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap		wasit indonesia sulit fokus kalau sudah dapat tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap
kejadian yang diprotes keras oleh ps tira mnurut opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terima kasih wasit		kejadian yang diprotes keras oleh ps tira mnurut opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terima kasih wasit
ketawa orang goblog yang tidak ngerti bola pada marah semua sebenarnya saya bukan fans tim manapun	tidakngerti	ketawa orang goblog yang tidakngerti bola pada marah semua sebenarnya saya bukan fans tim manapun
dear the jak aremania wasit salah apa coba penalti diving jelas pemain lawan sudah bermain sangat keras jangan hujat wasit		dear the jak aremania wasit salah apa coba penalti diving jelas pemain lawan sudah bermain sangat keras jangan hujat wasit
beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya		beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya
kenapa tidak ditonjok saja itu wasitnya kalau perlu sekalian saja keroyok biar tahu rasaaa	tidakditonjok	kenapa tidak ditonjok saja itu wasitnya kalau perlu sekalian saja keroyok biar tahu rasaaa

3.4.5 Stopword Removal

Pada tahap ini kata-kata yang tidak deskriptif (tidak penting) seperti kata ganti orang, kata penghubung, penunjuk, dan sebagainya dapat dibuang. Proses *stopword removal* menggunakan database dari kumpulan kata-kata yang tidak deskriptif (tidak penting), jika terdapat kata yang terdapat pada database tersebut maka kata tersebut dibuang. Seperti yang terdapat pada gambar 3.4.



Gambar 1.4 Block Diagram Stopword Removal

Berikut tahap-tahap dari proses *stopword removal*.

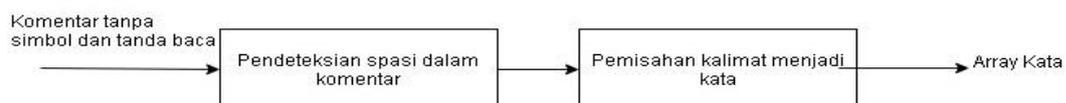
- a. Misalkan pada data latih P1 memiliki kata “kalau” yang sebelumnya telah dimasukkan ke dalam database stoplist. Sebelum melalui tahapan *stopword removal*, data latih P1 berisi kalimat “wasit indonesia sulit fokus kalau sudah dapat tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap”. Setelah dilakukan *stopword removal* maka kata “kalau” dalam data latih P1 akan terhapus sehingga akan menjadi “wasit indonesia sulit fokus sudah dapat tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap”. Adapun contoh tahap *Stopword Removal* diterapkan pada data latih dan data uji dapat dilihat pada Tabel 3.7.

Tabel 3.7 Hasil Stopword Removal

Sebelum Stopword Removal	Setelah Stopword Removal
wasit indonesia sulit fokus kalau sudah dapat tekanan heu semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	wasit indonesia sulit fokus sudah dapat tekanan semoga mental wasit juga bisa diperbaiki padahal fase grup mantap
kejadian yang diprotes keras oleh ps tira mnurut opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras dari belakang oleh pemain lawan terima kasih wasit	kejadian diprotes keras opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras belakang pemain lawan terima kasih wasit
ketawa orang goblog yang tidakngerti bola pada marah semua sebenarnya saya bukan fans tim manapun	ketawa orang goblog tidakngerti bola marah semua sebenarnya saya bukan fans tim manapun
dear the jak aremania wasit salah apa coba penalti diving jelas pemain lawan sudah bermain sangat keras jangan hujat wasit	wasit salah coba penalti jelas pemain lawan sudah bermain sangat keras jangan hujat wasit
beberapa menit terakhir cukup menyakitkan bagi tibo selamat persebaya dan wasit harus lebih baik lagi ya kedepannya	beberapa menit terakhir menyakitkan selamat persebaya wasit harus lebih baik ya kedepannya
kenapa tidak ditonjok saja itu wasitnya kalau perlu sekalian saja keroyok biar tahu rasaaa	kenapa tidak ditonjok itu wasitnya perlu sekalian saja keroyok tahu

3.4.6 Tokenizing

Pada tahap ini, teks yang telah diproses sebelumnya akan dipecah menjadi bagian-bagian kata yang disebut token, *Tokenizing* merupakan tahap pemotongan kalimat berdasarkan tiap kata yang menyusunnya. Pada tahap ini dilakukan penguraian deskripsi yang semula berupa kalimat-kalimat menjadi kata tunggal(term). Setelah melalui proses tokenization kita bisa mendapatkan jumlah kemunculan setiap tokennya. Seperti yang terdapat pada gambar 3.5 berikut.

**Gambar 3.5 Block Diagram Tokenizing**

Adapun contoh tahap *Tokenizing* diterapkan pada data latih dapat dilihat pada Tabel 3.8.

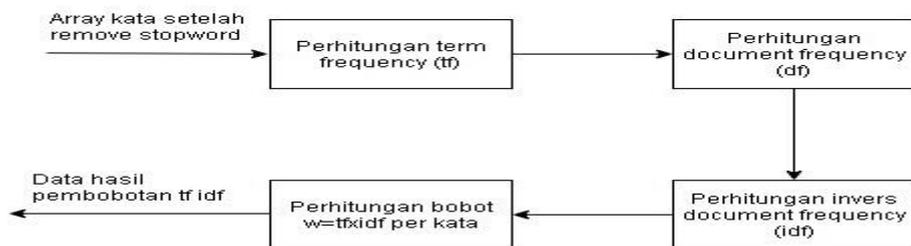
Tabel 3.8 Hasil Tokenizing

Sebelum Tokenisasi	Setelah Tokenisasi	
wasit indonesia sulit fokus sudah dapat tekanan semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	wasit indonesia sulit fokus sudah dapat tekanan semoga mental	wasit juga bisa diperbaiki padahal fase grup mantap
kejadian diprotes keras opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras belakang pemain lawan terima kasih wasit	kejadian diprotes keras opiniku wasit sudah bertindak benar proses pinalti memang	karena ada tekel keras belakang pemain lawan terima kasih wasit
ketawa orang goblog tidakngerti bola marah semua sebenarnya saya bukan fans tim manapun	ketawa orang goblog tidakngerti bola marah semua	sebenarnya saya bukan fans tim manapun
wasit salah coba penalti jelas pemain lawan sudah bermain sangat keras jangan hujat wasit	wasit salah coba penalti jelas pemain lawan	bermain sangat keras jangan hujat wasit

Sebelum Tokenisasi	Setelah Tokenisasi	
	sudah	
beberapa menit terakhir menyakitkan selamat persebaya wasit harus lebih baik ya kedepannya	beberapa menit terakhir menyakitkan selamat persebaya	wasit harus lebih baik ya kedepannya
kenapa tidak ditonjok itu wasitnya perlu sekalian saja keroyok tahu	kenapa tidak ditonjok itu wasitnya	perlu sekalian saja keroyok tahu

3.5 Analisis Pembobotan TF-IDF

Dalam penelitian ini dilakukan perhitungan bobot yang diperoleh dari kemunculan term(kata tunggal) dalam satu dokumen (tf) dan jumlah kemunculan dokumen (idf). Untuk mendapatkan nilai IDF maka dilakukan perhitungan dengan **Error! Reference source not found.** Setelah mendapatkan nilai tf dan idf, maka selanjutnya menghitung pembobotan pada term, untuk mendapatkan bobot (W) masing-masing dokumen pada setiap term(kata tunggal) dengan persamaan 2.3. Adapun *block diagram* dari proses pembobotan dapat dilihat pada gambar 6.



Gambar 3.6 Block Diagram Pembobotan

Berikut adalah contoh data yang sudah melewati tahapan *preprocessing* dapat dilihat pada Tabel 3.9.

Tabel 3.9 Hasil *Preprocessing*

Pernyataan	Kata	Pernyataan	Kata
P1	wasit indonesia sulit fokus sudah dapat tekanan semoga mental wasit juga bias diperbaiki padahal fase grup mantap	P2	kejadian diprotes keras opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras belakang pemain lawan terima kasih wasit
P3	ketawa orang goblog tidakngerti bola marah semua sebenarnya saya bukan fans tim manapun	P4	wasit salah coba penalti jelas pemain lawan sudah bermain sangat keras jangan hujat wasit
P5	beberapa menit terakhir menyakitkan selamat persebaya wasit harus lebih baik ya kedepannya	P6	kenapa tidak ditonjok itu wasitnya perlu sekalian saja keroyok tahu

Proses awal dilakukan perhitungan term (kata tunggal) pada setiap dokumen, sehingga akan mendapatkan frekuensi term. Selanjutnya adalah menghitung df, karena df adalah banyaknya dokumen dimana munculnya suatu term. Setelah memperoleh nilai df, maka dilakukan perhitungan idf dengan $idf_t = \log\left(\frac{N}{df}\right)$ (1.1)

$$idf_t = \log\left(\frac{N}{df}\right) \quad (1.1)$$

Diambil contoh pada kata “wasit”. Maka didapatkan banyak dokumen (N) = 7, dan df = 3. Maka perhitungannya seperti berikut.

$$idf_t = \log\left(\frac{7}{3}\right)$$

Selanjutnya untuk mendapatkan bobot term, maka dilakukan perhitungan tf dan idf dengan

$$W_t = tf_{dt} * idf_t \quad (1.2)$$

$$W_t = tf_{dt} * idf_t \quad (1.2)$$

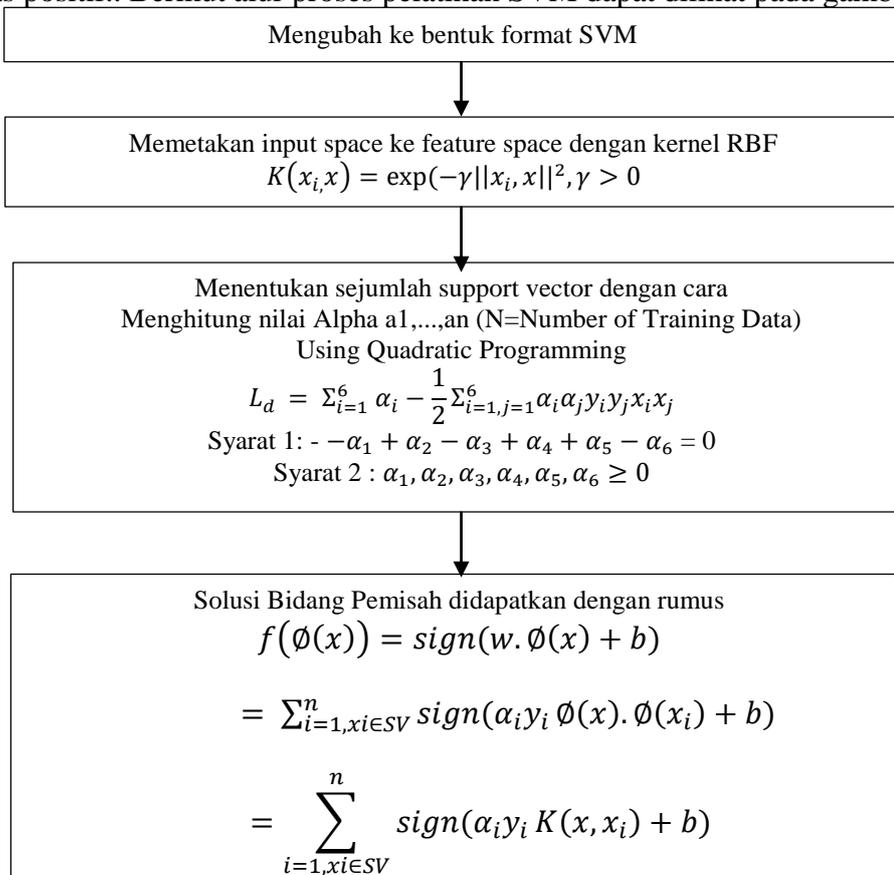
Diperoleh tf = 3, dan idf = 0,477. Maka perhitungannya menjadi seperti berikut.

$$Wt = 3 * 0,368 = 1,104$$

Sehingga kata “wasit” memiliki bobot sebesar 1,104. Hasil dari perhitungan pembobotan TF-IDF dapat dilihat pada Tabel 3.10 pada lampiran tabel.

3.6 Analisis Klasifikasi *Support Vector Machine*

Pelatihan SVM bertujuan untuk menemukan vektor α , nilai W dan konstanta b untuk mendapatkan hyperplane terbaik. Pada penelitian kali ini, data yang digunakan sebagai pelatihan adalah data komentar. Dalam pelatihan SVM, setiap model klasifikasi dilatih pada data dari dua kelas ke- i dan kelas- j . Data masukan yang akan digunakan untuk proses pelatihan yaitu data komentar P1 sampai P6 telah diberi kelas dan telah melalui tahapan preprocessing. Sesuai dengan data masukan, data komentar diberi kelas positif dan kelas negatif, lalu diberikan label kelas 1 atau -1 yang dimana kelas -1 merupakan kelas negatif sedangkan kelas 1 merupakan kelas positif. Berikut alur proses pelatihan SVM dapat dilihat pada gambar 3.7.



Gambar 3.7 Block Diagram Proses Pelatihan SVM

3.6.1 Mengubah ke bentuk format SVM

Pada proses ini, data masukan yang telah melalui tahapan preprocessing diubah kedalam data vektor. Data yang akan digunakan dalam proses pelatihan adalah data komentar P1 sampai dengan P6 yang diberikan label 1 dan 0. Berikut data komentar P1 sampai dengan P6 yang akan diubah menjadi data vektor yang ditampilkan pada Tabel 3.10.

Tabel 3.10 Vektor data komentar

Komentar (Pi)	Format Vektor	(x_i)	Kelas
wasit indonesia sulit fokus sudah dapat tekanan semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	[1:0,845 2:1,104 3:0,845 4:0,845 5:0,845 6:0,845 7:0,544 8:0,845 9:0,845 10:0,845 11:0,845 12:0,845 13:0,368 14:0,845 15:0,845 16:1,104 17:0,845 18:0,845 19:0,845 20:1,104 21:0 22:0 23:0 24:0 25:0 26:0 27:0 28:0 29:0 30:0 31:0 32:0 33:0 34:0 35:0 36:0 37:0 38:0,368 40:0 41:0 42:0 43:0 44:0 45:0 46:0 47:0 48:0 49:0,544 50:0 51:1,104 52:0 53:1,104 54:0 55:0 56:0 57:0,368 58:0 59:0 60:0 61:1,104 62:0 63:0 64:0 65:0 66:0 67:0 68:0 69:0 70:1,104 71:0 72:0 73:0 74:0 75:0 76:0 77:1,104 78:0]	x_1	Negatif
kejadian diprotes keras opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras belakang pemain lawan terima kasih wasit	[1:0 2:0 3:0 4:0 5:0 6:0 7:0 8:0 9:0 10:0 11:0 12:0 13:0 14:0 15:0 16:0 17:0 18:0 19:0 20:0 21:0,544 22:0,845 23:0,845 24:0,845 25:0,845 26:0,845 27:0,544 28:0,845 29:0 30:0 31:0 32:0 33:0 34:0 35:0 36:0 37:0 38:0 39:0 40:0 41:0 42:0,544 43:0 44:0 45:0,544 46:0 47:0 48:0 49:0 50:0 51:0 52:0 53:0 54:0 55:0 56:0 57:0 58:0 59:0 60:0 61:0 62:0 63:0 64:0 65:0 66:0 67:0 68:0 69:0 70:0 71:0 72:0 73:0 74:0 75:0 76:0 77:0 78:0]	x_2	Positif
ketawa orang goblog tidakngerti bola marah semua sebenarnya saya bukan fans tim manapun	[1: 0 2:0 3:0 4:0 5:0 6:0 7:0 8:0 9:0 10:0 11:0 12:0 13:0 14:0 15:0 16:0 17:0 18:0 19:0 20:0 21:0 22:0 23:0 24:0 25:0 26:0 27:0 28:0 29:0,845 30:0,845 31:0,845 32:0,845 33:0,845 34:0,845 35:0,845 36:0 37:0 38:0 39:0 40:0 41:0 42:0 43:0 44:0 45:0 46:0 47:0 48:0 49:0 50:0 51:0 52:0 53:0 54:0 55:0 56:0 57:0 58:0 59:0 60:0 61:0 62:0 63:0 64:0 65:0 66:0 67:0 68:0 69:0 70:0 71:0 72:0 73:0 74:0 75:0 76:0 77:0 78:0]	x_3	Negatif

Komentar (Pi)	Format Vektor	(x_i)	Kelas
wasit salah coba penalti jelas pemain lawan sudah bermain sangat keras jangan hujat wasit	1:0 2:0 3:0 4:0 5:0 6:0 7:0 8:0 9:0 10:0 11:0 12:0 13:0,368 14:0 15:0 16:0 17:0 18:0 19:0 20:0 21:0 22:0 23:0 24:0 25:0 26:0 27:0 28:0 29:0 30:0 31:0 32:0 33:0 34:0 35:0 36:0,845 37:0,845 38:0,368 39:0,845 40:0,845 41:0,845 42:0 43:0 44:0 45:0 46:0 47:0 48:0 49:0 50:0 51:0 52:0 53:0 54:0 55:0 56:0,368 57:0 58:0 59:0 60:0 61:0 62:0 63:0 64:0 65:0 66:0 67:0 68:0 69:0 70:0 71:0 72:0 73:0 74:0 75:0 76:0 77:0 78:0	x_4	Positif
beberapa menit terakhir menyakitkan selamat persebaya wasit harus lebih baik ya kedepannya	1:0 2:0 3:0 4:0 5:0 6:0 7:0,544 8:0 9:0 10:0 11:0 12:0 13:0 14:0 15:0 16:0 17:0 18:0 19:0 20:0 21:0,544 22:0 23:0 24:0 25:0 26:0 27:0,544 28:0 29:0 30:0 31:0 32:0 33:0 34:0 35:0 36:0 37:0 38:0 39:0 40:0 41:0 42:0,544 43:0,845 44:0,845 45:0,544 46:0,845 47:0,845 48:0,544 49:0,845 50:0 51:0 52:0 53:0 54:0 55:0 56:0 57:0 58:0 59:0 60:0 61:0 62:0 63:0 64:0 65:0 66:0 67:0 68:0 69:0 70:0 71:0 72:0 73:0 74:0 75:0 76:0 77:0 78:0	x_5	Positif
kenapa tidak ditonjok itu wasitnya perlu sekalian saja keroyok tahu	1:0 2:0,736 3:0 4:0 5:0 6:0 7:0 8:0 9:0 10:0 11:0 12:0 13:0,368 14:0 15:0 16:0,736 17:0 18:0 19:0 20:0,736 21:0 22:0 23:0 24:0 25:0 26:0 27:0 28:0 29:0 30:0 31:0 32:0 33:0 34:0 35:0 36:0 37:0 38:0,368 39:0 40:0 41:0 42:0 43:0 44:0 45:0 46:0 47:0 48:0 49:0 50:0,736 51:0,845 52:0,736 53:0,845 54:0,845 55:0,845 56:0,368 57:0,845 58:0,544 59:0,845 60:0,736 61:0 62:0 63:0 64:0,544 65:0 66:0 67:0 68:0 69:0,736 70:0 71:0 72:0 73:0 74:0 75:0 76:0 77:0,736 78:0	x_6	Negatif

3.6.2 Pemetaan Input Space Feature

Sebelum melakukan proses perhitungan kernelisasi RBF, hitung data latih berdasarkan fungsi pada RBF $x_i - x_j$ terlebih dahulu. Untuk perhitungan $x_i - x_j$ dapat dilihat pada Tabel 3.11.

Tabel 3.11 Perhitungan kernelisasi $x_i - x_j$

Term	$x_1 - x_1$	$x_1 - x_2$	$x_1 - x_3$	$x_1 - x_4$	$x_1 - x_5$	$x_1 - x_6$
gaktonjok	0	0,845	0,845	0,845	0,845	0,845
wasit	0	1,104	1,104	1,104	1,104	0,368
keroyok	0	0,845	0,845	0,845	0,845	0,845
hantam	0	0,845	0,845	0,845	0,845	0,845
mulutnya	0	0,845	0,845	0,845	0,845	0,845
rifat	0	0,845	0,845	0,845	0,845	0,845
Manahati	0	0,544	0,544	0,544	0,544	0,544
Goodboy	0	0,845	0,845	0,845	0,845	0,845
Semangat	0	0,845	0,845	0,845	0,845	0,845
Tira	0	0,845	0,845	0,845	0,845	0,845
Kabo	0	0,845	0,845	0,845	0,845	0,845
Good	0	0,845	0,845	0,845	0,845	0,845
Persebaya	0	0,368	0,368	0	0,368	0
Kali	0	0,845	0,845	0,845	0,845	0,845
Hajar	0	0,845	0,845	0,845	0,845	0,845
Wasit	0	1,104	1,104	1,104	1,104	0,368
Efek	0	0,845	0,845	0,845	0,845	0,845
Tanggal	0	0,845	0,845	0,845	0,845	0,845
Tua	0	0,845	0,845	0,845	0,845	0,845
Wasit	0	1,104	1,104	1,104	1,104	0,368
Wasitnya	0	-0,544	0	0	-0,544	0
Payah	0	-0,845	0	0	0	0
Takut	0	-0,845	0	0	0	0
Tekanan	0	-0,845	0	0	0	0
Suporter	0	-0,845	0	0	0	0
Wajar	0	-0,845	0	0	0	0
Sepakbola	0	-0,544	0	0	-0,544	0
Gakberkembang	0	-0,845	0	0	0	0
komentar	0	0	-0,845	0	0	0
gakngerti	0	0	-0,845	0	0	0
offside	0	0	-0,845	0	0	0

Term	$x_1 - x_1$	$x_1 - x_2$	$x_1 - x_3$	$x_1 - x_4$	$x_1 - x_5$	$x_1 - x_6$
tolol	0	0	-0,845	0	0	0
gatau	0	0	-0,845	0	0	0
bola	0	0	-0,845	0	0	0
diam	0	0	-0,845	0	0	0
pertandingan	0	0	0	-0,845	0	0
seru	0	0	0	-0,845	0	0
persebaya	0	0,368	0,368	0	0,368	0
auto	0	0	0	-0,845	0	0
final	0	0	0	-0,845	0	0
juara	0	0	0	-0,845	0	0
sepakbola	0	-0,544	0	0	-0,544	0
indonesia	0	0	0	0	-0,845	0
gakmaju	0	0	0	0	-0,845	0
wasitnya	0	-0,544	0	0	-0,544	0
kaya	0	0	0	0	-0,845	0
gini	0	0	0	0	-0,845	0
manahati	0	0,544	0,544	0,544	0	0,544
lestusen	0	0	0	0	-0,845	0
wasit	0	1,104	1,104	1,104	1,104	0,368
gaksalah	0	0	0	0	0	-0,845
wasit	0	1,104	1,104	1,104	1,104	0,368
manusia	0	0	0	0	0	-0,845
keputusannya	0	0	0	0	0	-0,845
mainnya	0	0	0	0	0	-0,845
persebaya	0	0,368	0,368	0	0,368	0
bagus	0	0	0	0	0	-0,845
fans	0	0	0	0	0	-0,544
manapun	0	0	0	0	0	-0,845
wasit	0	1,104	1,104	1,104	1,104	0,368
gak	0	0	0	0	0	0
mata	0	0	0	0	0	0
bangsat	0	0	0	0	0	0
fans	0	0	0	0	0	-0,544
bali	0	0	0	0	0	0

Term	$x_1 - x_1$	$x_1 - x_2$	$x_1 - x_3$	$x_1 - x_4$	$x_1 - x_5$	$x_1 - x_6$
united	0	0	0	0	0	0
jengkel	0	0	0	0	0	0
kepemimpinan	0	0	0	0	0	0
wasit	0	1,104	1,104	0	1,104	0,368
sungguh	0	0	0	0	0	0
menyakitkan	0	0	0	0	0	0
dihukum	0	0	0	0	0	0
pinalti	0	0	0	0	0	0
semoga	0	0	0	0	0	0
pssi	0	0	0	0	0	0
mengusut	0	0	0	0	0	0
wasit	0	0	0	0	0	-0,736
dipake	0	0	90	0	0	0
$(x_i - x_j)^2$	0	5488,13	5018,41	3588,73	4773,81	4888,8

Didapatkan hasil dari $\sum_{i=1, j=1}^n (x_i - x_j)^2$ dari tabel 3.10 masing-masing komentar, kemudian lakukan perhitungan kernelisasi dengan dimasukkan ke dalam persamaan 3.3, yaitu

$$K(x_i x) = \exp(-\gamma \|x_i - x\|^2), \gamma > 0 \quad (1.3)$$

Nilai gamma akan diuji menggunakan nilai gamma (γ) = 0,5. Berikut perhitungan fungsi kernel sebagai berikut.

$$\begin{aligned} K(x_1 x_1) &= \exp(-\gamma \|x_1 - x_1\|^2) \\ &= \exp(-0,5(0)) \\ &= \exp(0) \\ &= 1 \end{aligned}$$

Lakukan perhitungan yang sama untuk iterasi selanjutnya. Untuk hasil nilai perhitungan kernel lainnya ditampilkan pada Tabel 3.12.

Tabel 3.12 Hasil perhitungan kernelisasi

$K(x_1x_1) =$ 1	$K(x_1x_2) =$ 0,05311	$K(x_1x_3) =$ 0,24446	$K(x_1x_4) =$ 0,02876	$K(x_1x_5) =$ 0,21687	$K(x_1x_6) =$ 0,24654
$K(x_2x_1) =$ 0,05311	$K(x_2x_2) =$ 1	$K(x_2x_3) =$ 0,02046	$K(x_2x_4) =$ 0,01612	$K(x_2x_5) =$ 0,12934	$K(x_2x_6) =$ 0,13421
$K(x_3x_1) =$ 0,24446	$K(x_3x_2) =$ 0, 02046	$K(x_3x_3) =$ 1	$K(x_3x_4) =$ 0,02146	$K(x_3x_5) =$ 0,11824	$K(x_3x_6) =$ 0,02353
$K(x_4x_1) =$ 0,02876	$K(x_4x_2) =$ 0, 0161	$K(x_4x_3) =$ 0,02146	$K(x_4x_4) =$ 1	$K(x_4x_5) =$ 0,02456	$K(x_4x_6) =$ 0,13456
$K(x_5x_1) =$ 0,21687	$K(x_5x_2) =$ 0, 12934	$K(x_5x_3) =$ 0,11824	$K(x_5x_4) =$ 0, 02456	$K(x_5x_5) =$ 1	$K(x_5x_6) =$ 0.05321
$K(x_6x_1) =$ 0,24654	$K(x_6x_2) =$ 0, 13421	$K(x_6x_3) =$ 0, 023534	$K(x_6x_4) =$ 0,13456	$K(x_6x_5) =$ 0, .05321	$K(x_6x_6) =$ 1

Kemudian tahap selanjutnya adalah melakukan perhitungan untuk nilai y adalah nilai dari label atau nilai dari kelas yang telah diberikan. Nilai tersebut dapat dilihat pada Tabel 13.

Tabel 3.13 Nilai Label pada y_i

y_1	y_2	y_3	y_4	y_5	y_6
-1	1	-1	1	1	-1

Selanjutnya dilakukan perhitungan $y_i y_j$ sebanyak 6 data dimana $i, j = 1, 2, \dots, n$. perhitungan $y_i y_j$ sebagai berikut.

$$y_1 y_1 = (-1)(-1) = 1$$

Lakukan perhitungan sebanyak 6 data $y_i y_j$, Sehingga terbentuk perhitungan $y_i y_j$ hasilnya akan menjadi seperti Tabel 3.14.

Tabel 3.14 Hasil Perhitungan label $y_i y_j$

$y_1 y_1$	$y_1 y_2$	$y_1 y_3$	$y_1 y_4$	$y_1 y_5$	$y_1 y_6$
1	-1	1	-1	-1	-1
$y_2 y_1$	$y_2 y_2$	$y_2 y_3$	$y_2 y_4$	$y_2 y_5$	$y_2 y_6$
-1	1	-1	1	1	-1
$y_3 y_1$	$y_3 y_2$	$y_3 y_3$	$y_3 y_4$	$y_3 y_5$	$y_3 y_6$
1	-1	1	-1	-1	1
$y_4 y_1$	$y_4 y_2$	$y_4 y_3$	$y_4 y_4$	$y_4 y_5$	$y_4 y_6$
-1	1	-1	1	1	-1
$y_5 y_1$	$y_5 y_2$	$y_5 y_3$	$y_5 y_4$	$y_5 y_5$	$y_5 y_6$
-1	1	-1	1	1	-1
$y_6 y_1$	$y_6 y_2$	$y_6 y_3$	$y_6 y_4$	$y_6 y_5$	$y_6 y_6$
-1	-1	1	-1	-1	1

3.6.3 Menghitung Lagrange

Matriks kernel diatas setiap elemennya merupakan hasil $K(x_i x) = \exp(-\gamma ||x_1 - x||^2)$ yang akan berkorelasi dengan α_i, α_j . Dengan menggunakan kernel-kernel sebagai pengganti dot-product $x_i x_j$ dalam persamaan dualitas

$$L_d = \sum_{i=1}^6 \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^6 \alpha_i \alpha_j y_i y_j x_i x_j$$

(1.4).

$$L_d = \sum_{i=1}^6 \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^6 \alpha_i \alpha_j y_i y_j x_i x_j \quad (1.4)$$

Syarat 1: $-\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 + \alpha_5 - \alpha_6 = 0$

Syarat 2 : $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \geq 0$

Sehingga didapatkan :

$$\begin{aligned}
 L_d &= \sum_{i=1}^6 \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^6 \alpha_i \alpha_j y_i y_j x_i x_j \\
 &= -\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 + \alpha_5 - \alpha_6 + \alpha_1 \alpha_1 y_1 y_1 x_1 x_1 + \alpha_1 \alpha_2 y_1 y_2 x_1 x_2 + \dots + \dots \\
 &\quad + \dots + \dots + \alpha_6 \alpha_4 y_6 y_4 x_6 x_4 + \alpha_6 \alpha_5 y_6 y_5 x_6 x_5 + \alpha_6 \alpha_6 y_6 y_6 x_6 x_6 \\
 &= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 - \frac{1}{2} (\alpha_1 \alpha_1 (-1)(1) + \alpha_1 \alpha_2 (1)(0,05311))
 \end{aligned}$$

$$\begin{aligned}
& +\alpha_1\alpha_3(-1)(0,24446) + \alpha_1\alpha_4(1)(0,02876) + \alpha_1\alpha_5(1)(0,21687) \\
& \quad + \alpha_1\alpha_6(-1)(0,24654) + \alpha_2\alpha_1(-1)(0,05311) + \dots + \dots + \dots \\
& \quad + \dots + \alpha_6\alpha_5(-1)(0,05321) + \alpha_6\alpha_6(1)(1) \\
= & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 + \alpha_1^2(-1) + \alpha_1\alpha_2(-0,05311) + \dots + \dots + \dots \\
& \quad + \alpha_6\alpha_4(-0,13456) + \alpha_6\alpha_5(-0,05321) + \alpha_6^2(1)
\end{aligned}$$

Dalam fungsi tujuan, suku kedua sudah dikalikan dengan $y_i y_j$. Persamaan tersebut memenuhi standar *Quadratic Programming* sehingga dapat dibantu penyelesaiannya dengan solver komersial untuk *Quadratic Programming* (QP). Dengan bantuan perangkat lunak, didapatkan hasil sebagai berikut:

$$\alpha_1 = 0,255 \quad \alpha_2 = 0,122 \quad \alpha_3 = 0,243 \quad \alpha_4 = 0,060 \quad \alpha_5 = 0,102 \quad \alpha_6 = 0,123 \quad b = 0,053$$

Sehingga didapat nilai alpha dan beta dari masing-masing kelas dapat dilihat pada Tabel 3.15.

Tabel 3.15 Nilai α_i dan b masing masing kelas

No.	Kelas	Nama Kelas	α_i	b_i
1	-1	Negatif	$\alpha_1 = 0,455$	$b = 0,125$
2	1	Positif	$\alpha_2 = 0,534$	$b = 0,125$
3	-1	Negatif	$\alpha_3 = 0,532$	$b = 0,125$
4	1	Positif	$\alpha_4 = 0,644$	$b = 0,125$
5	1	Positif	$\alpha_5 = 0,445$	$b = 0,125$
6	-1	Negatif	$\alpha_6 = 0,435$	$b = 0,125$

3.6.4 Bidang Pemisah

Hasil ini menunjukkan bahwa semua data training adalah *support vector*. Karena nilai $\alpha > 0$ sementara b didapatkan dari proses pelatihan yang dilakukan. Setelah semua α dan b didapat, maka model svm sudah siap digunakan untuk prediksi seperti pada persamaan 2.18, yaitu:

$$f(\phi(x)) = (\text{sign}(w \cdot \phi(x)) + b)$$

Dimana

$$= \sum_{i=1, x_i \in SV}^n (\text{sign}(\alpha_i y_i \phi(x) \cdot \phi(x_i)) + b)$$

$$= \sum_{i=1, x_i \in SV}^n (\text{sign}(\alpha_i y_i K(x, x_i)) + b)$$

dimana $i = 1, 2, 3, \dots, n = \text{jumlah support vector}$. Maka didapatkan bidang pemisah:

$$f(\phi(x)) = \sum_{i=1, x_i \in SV}^n (\text{sign}(\alpha_i y_i, K(x_i x_{testing})) + b)$$

$$f(x_{testing}) = (\text{sign} (0,255(0,645) + 0,122(0,05311) + 0,243(0,24446) - 0,060(0,2876) - 0,102(0,21687) - 0,123(0,24654) , K(x_i x_{testing})) + 0,125)$$

Data latih yang terbentuk adalah seperti pada Tabel 3.16.

Tabel 3.16 Data Hasil Pelatihan

Pernyataan	Data latih	Label
P1	wasit indonesia sulit fokus sudah dapat tekanan semoga mental wasit juga bisa diperbaiki padahal fase grup mantap	-1
P2	kejadian diprotes keras opiniku wasit sudah bertindak benar proses pinalti memang karena ada tekel keras belakang pemain lawan terima kasih wasit	1
P3	ketawa orang goblog tidakngerti bola marah semua sebenarnya saya bukan fans tim manapun	-1
P4	wasit salah coba penalti jelas pemain lawan sudah bermain sangat keras jangan hujat wasit	1

Pernyataan	Data latih	Label
P5	beberapa menit terakhir menyakitkan selamat persebaya wasit harus lebih baik ya kedepannya	1
P6	kenapa tidak ditonjok itu wasitnya perlu sekalian saja keroyok tahu	-1

3.7 Analisis Pengujian SVM

Setelah mendapatkan nilai α dan b sebagai model fitur, dan nilai parameter γ 0,5 (γ) dari proses pelatihan, selanjutnya menguji data uji ke dalam kelas +1 atau -1 dengan model fitur yang sudah di dapat. Data yang digunakan untuk dilakukan pengujian adalah data latih $P_1 - P_6$ sebagai hasil pelatihan data.

3.7.1 Mengubah ke Format SVM

Pada proses ini, data uji yang akan dilakukan pada sebuah komentar *testing* yang telah melalui tahapan preprocessing diubah kedalam data vektor ditampilkan pada Tabel 3.17.

Tabel 3.17 contoh data uji yang ditesting

Komentar	Data Uji	Format Vektor	(x_i)
$P_{testing}$	ada sedikit keberpihakan wasit buat tuan rumah Yang dari fase grup wasit nya pada bener eh pas udah quarter final malah jadi hancur	1:0,2:1,104,3:0,4:0,5:0,6:0,7:0,8:0,9:0,10:0 ,11:0,12:0,13:0,14:0,15:0,16:1,104,17:0,18 :0,19:0,20:1,104,21:0,22:0,23:0,24:0,25:0, 26:0,27:0,28:0,29:0,30:0,31:0,32:0,33:0,34 :0,35:0,36:0,37:0,38:0,39:0,40:0,41:0,42:0, 43:0,44:0,45:0,46:0,47:0,48:0,49:0,50:0,51 :1,104,52:0,53:1,104,54:0,55:0,56:0,57:0,5 8:0,59:0,544,60:0,61:1,104,62:0,845,63:0, 845,64:0,845,65:0,544,66:0,845,67:0,845, 68:0,845,69:0,845,70:1,104,71:0,845,72:0, 845,73:0,845:74:0,845,75:0,845,76:0,845, 77:0,845,78:1,104	$x_{testing}$

0	0	0	0	0	-0,736
0	0	90	0	0	0
$(x_i - x_{testing})^2$	0	5488,13	5018,41	3588,73	4773,81

Didapatkan hasil dari $\sum_{i=1, j=1}^n (x_i - x_{testing})^2$ dari tabel 3.10 masing-masing komentar, kemudian lakukan perhitungan kernelisasi dengan dimasukkan ke dalam persamaan 2.16, yaitu:

$$K(x_i x_{testing}) = \exp(-\gamma ||x_i - x_{testing}||^2), \gamma > 0$$

Nilai gamma menggunakan nilai gamma (γ) = 0,5. Berikut perhitungan kernelnya sebagai berikut:

$$\begin{aligned} K(x_1 x_{testing}) &= \exp(-\gamma ||x_1 - x_{testing}||^2) \\ &= \exp(-0,5(0)) \\ &= \exp(0) \\ &= 1 \end{aligned}$$

Lakukan sebanyak jumlah data, Setelah mendapatkan nilai kernelisasi RBF, kemudian proses selanjutnya adalah menentukan kelas mana komentar yang telah *testing* menggunakan fungsi hyperplane.

Sehingga

$$\begin{aligned} f(x_{testing}) &= (\text{sign} \sum_{i=1}^n \alpha_i y_i K(x_i x_{testing}) + b) \\ f(x_{testing}) &= (\text{sign} (\alpha_1 y_1 K(x_1 x_{testing}) + \alpha_2 y_2 K(x_2 x_{testing}) \\ &\quad + \alpha_3 y_3 K(x_3 x_{testing}) + \alpha_4 y_4 K(x_4 x_{testing}) \\ &\quad + \alpha_5 y_5 K(x_5 x_{testing}) + \alpha_6 y_6 K(x_6 x_{testing}) + b) \\ &= (\text{sign}(0,255(1) + 0,122(0,05311) + 0,243(0,24446) - 0,060(0,02876) - \\ &\quad 0,102(0,21687) - 0,123(0,24654) + 0,125) \\ &= \text{sign}(0,0819) \\ &= +1 \end{aligned}$$

Setelah salah satu kelas komentar melalui tahapan *testing*, menghasilkan fungsi hyperplane +1 yaitu (positif).

3.8 Analisis Kebutuhan Non-fungsional

Analisis kebutuhan non fungsional dilakukan menjadi tiga tahap, yaitu analisis kebutuhan perangkat keras(hardware), analisis kebutuhan perangkat lunak(software) minimal agar implementasi algoritma pada penelitian ini dapat berjalan dengan baik, dan juga analisis pengguna.

3.8.1. Analisis Kebutuhan Perangkat Keras (*Hardware*)

Spesifikasi perangkat keras (*Hardware*) yang dibutuhkan pada penelitian ini dapat dilihat pada 3.19

Tabel 3.19 Spesifikasi Kebutuhan Perangkat Keras (Hardware)

Perangkat Keras (<i>Hardware</i>)	Spesifikasi
<i>Processor</i>	<i>Dual Core</i> atau lebih tinggi
<i>RAM</i>	4 GB atau lebih tinggi
<i>Monitor</i>	Resolusi 1024x768 atau lebih tinggi
<i>Keyboard dan Mouse</i>	<i>Standard</i>

3.8.2. Analisis Kebutuhan Perangkat Lunak (*Software*)

Spesifikasi dari perangkat lunak (*software*) yang dibutuhkan pada penelitian ini dapat dilihat pada Tabel 3.20.

Tabel 3.20 Spesifikasi Kebutuhan Perangkat Lunak (Software)

Perangkat Lunak (<i>Software</i>)	Spesifikasi
Sistem Operasi	Windows 10
Bahasa Pemrograman	Python versi 3
<i>Code Editor</i>	Visual Studio Code, Spyder

3.8.3. Analisis Pengguna

Karakteristik pengguna dalam menjalankan perangkat lunak yang akan dibangun hanya terdapat satu jenis pengguna yaitu seorang penguji. Penguji dapat menjalankan serta mengetahui hasil dari perangkat lunak yang dijalankan. Adapun spesifikasi pengguna yang dibutuhkan adalah :

1. Menguasai penggunaan komputer
2. Mengerti secara teknis *tools* dan perangkat lunak pendukung dalam menjalankan perangkat lunak yang telah dibangun.
3. Mengerti tahapan-tahapan dalam menjalankan perangkat lunak yang telah dibangun.

3.9 Spesifikasi Kebutuhan Perangkat Lunak Fungsional

Spesifikasi kebutuhan perangkat lunak dibagi menjadi dua kebutuhan yaitu kebutuhan non fungsional dan fungsional. Kebutuhan fungsional dapat dilihat pada Tabel 3.21.

Tabel 3.21 Spesifikasi Kebutuhan Fungsional

Kode SKPL	Spesifikasi Kebutuhan Perangkat Lunak
SKPL-F-001	Sistem dapat menambahkan data komentar dengan format .csv
SKPL-F-002	Sistem dapat melakukan tahap pre-processing
SKPL-F-003	Sistem dapat melakukan pembobotan <i>TF-Idf</i>
SKPL-F-004	Sistem dapat melakukan proses klasifikasi <i>Support Vector Machine</i>
SKPL-F-005	Sistem dapat mengukur akurasi klasifikasi

Kebutuhan non fungsional yang pada perangkat lunak yang dibangun dapat dilihat pada tabel 3.22.

Tabel 3.22 Spesifikasi Kebutuhan non fungsional

Kode SKPL	Spesifikasi Kebutuhan Perangkat Lunak
SKPL-NF-001	Sistem yang dibangun menggunakan python
SKPL-NF-002	Sistem menggunakan python 3.7
SKPL-NF-003	Tampilan antarmuka menggunakan flask

3.10 Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional merupakan spesifikasi kebutuhan yang dilakukan perangkat lunak ketika diimplementasikan. Analisis kebutuhan fungsional berisi analisis proses yang akan diterapkan di dalam sistem dan memberikan gambaran mengenai rencana desain dan struktur sistem agar lebih mudah dipahami. Analisis yang dilakukan akan dimodelkan dengan menggunakan DFD (*Data Flow Diagram*).

3.10.1 Diagram Konteks

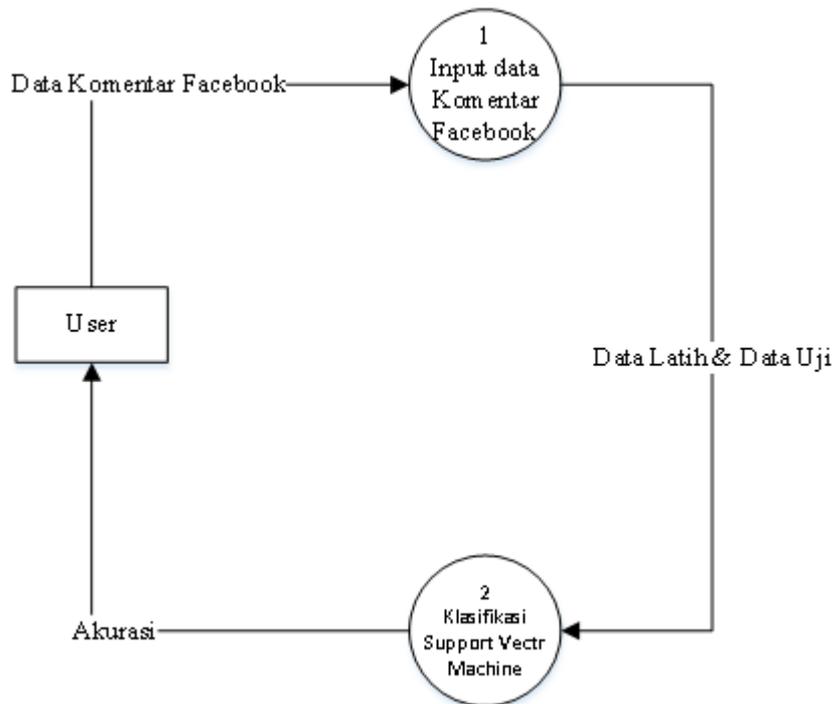
Diagram konteks menggambarkan sistem secara keseluruhan sistem analisis sentimen dengan user sebagai entitas eksternal atau pengguna sistem. *Diagram konteks* dapat dilihat pada gambar 3.8.



Gambar 3.8 Diagram Konteks

3.9.1. Data Flow Diagram (DFD) Level 1

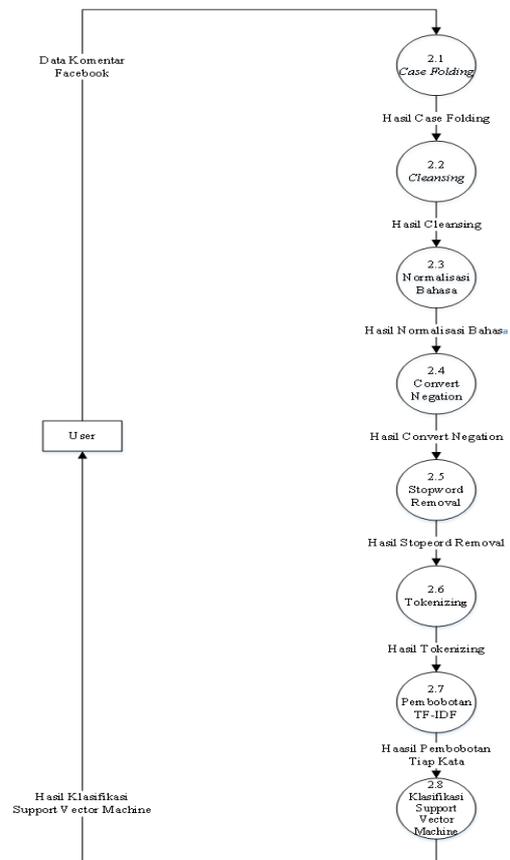
Data flow diagram level 1 merupakan pemecahan proses dari diagram konteks. Didalam diagram ini menjelaskan aliran data yang terdapat dalam sistem analisis sentimen. Pada diagram DFD level 1 terdapat 2 proses yaitu Input data komentar facebook, klasifikasi *Support Vector Machine*. DFD level 1 dapat dilihat pada.



Gambar 1.1 DFD Level 1

3.9.2. Data Flow Diagram (DFD) Level 2

Data flow diagram level 2 berisi pemecahan proses dari DFD level 1 pada proses klasifikasi *Support Vector Machine*. Didalam diagram ini menjelaskan aliran data yang terdapat pada proses klasifikasi *Support Vector Machine* yang didalamnya terdapat proses *case folding*, *cleansing*, normalisasi bahasa, *convert negation*, *stopword removal*, *tokenizing*, pembobotan *tf idf* dan klasifikasi *Support Vector Machine*. DFD level 2 proses klasifikasi *Support Vector Machine* dapat dilihat pada gambar 3.10.



Gambar 3.10 DFD Level 2

3.11 Spesifikasi Proses

Spesifikasi proses digunakan untuk menggambarkan proses model aliran yang terdapat pada DFD (*Data Flow Diagram*). Spesifikasi proses dari gambaran DFD tersebut akan dijelaskan pada tabel spesifikasi proses dibawah ini :

Tabel 3.19 Spesifikasi Proses Input Data Latih

No. Proses	1
Nama Proses	Input Data Latih
Sumber	<i>User</i>
<i>Input</i>	Data Latih
<i>Output</i>	Pesan berhasil melakukan penambahan data latih
Logika Proses	<ol style="list-style-type: none"> 1. <i>User</i> dapat melakukan input data data latih 2. <i>User</i> mendapatkan informasi banyaknya data latih

Tabel 3.20 Spesifikasi Proses Input Data Uji

No. Proses	1
Nama Proses	Input Data Uji
Sumber	<i>User</i>
<i>Input</i>	Data Uji
<i>Output</i>	Pesan berhasil melakukan penambahan data uji
Logika Proses	<ol style="list-style-type: none"> 1. <i>User</i> dapat melakukan input data uji 2. <i>User</i> mendapatkan informasi banyaknya data uji.

Tabel 3.21 Spesifikasi Proses Klasifikasi SVM

No. Proses	2
Nama Proses	Klasifikasi <i>Support Vector Machine</i>
Sumber	<i>User</i>
<i>Input</i>	Data latih & data Uji
<i>Output</i>	Nilai Akurasi
Logika Proses	<ol style="list-style-type: none"> 1. <i>User</i> melakukan input komentar facebook yang dapat dijadikan sebagai data latih dan data uji. 2. <i>User</i> mendapat informasi banyaknya data komentar facebook, data latih dan data uji 3. <i>System</i> melakukan klasifikasi SVM dengan menggunakan hasil data latih dan data uji 4. <i>Sistem</i> akan memberikan nilai akurasi.

3.11.1 Kamus Data

Kamus data ikut berperan dalam perancangan dan pembangunan sistem informasi. Berfungsi untuk menjelaskan semua data yang digunakan didalam sistem. Berikut adalah kamus data pada Analisis Sentimen Cyberbullying pada Komentar Facebook dengan Metode Klasifikasi *Support Vector Machine*:

Tabel 3.22 Kamus Data Dokumen Komentar

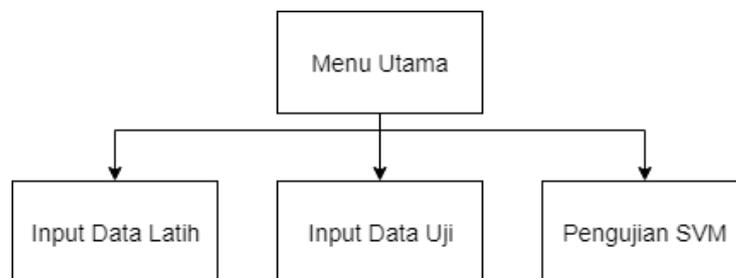
Nama Aliran Data	Data Dokumen Komentar
Digunakan pada	Proses 1,2,3,2.1,2.2,2.3,2.4,2.5,2.6
Deskripsi	Berisi Data Komentar Facebook
Struktur Data	Data,Label
Data	[A-Z][a-z][0-9]
Label	[0-9]

3.12 Perancangan Sistem

Pada bagian ini menjelaskan perancangan sistem yang akan dibuat untuk implementasi hasil analisis sebelumnya. Perancangan sistem yang dibuat terdiri dari perancangan struktur menu, perancangan antarmuka, perancangan pesan dan jaringan semantik.

3.12.1 Perancangan Struktur Menu

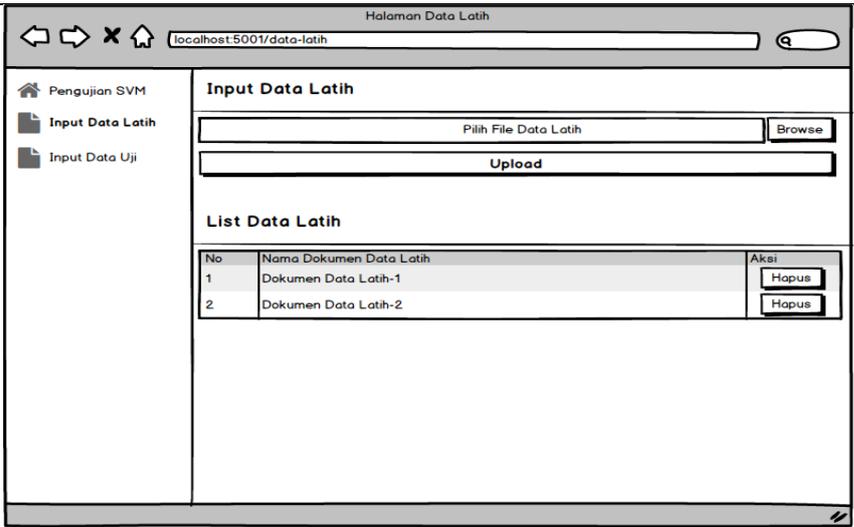
Perancangan struktur menu merupakan sebuah bentuk umum yang dapat menggambarkan alur sistem, sehingga sistem yang dibangun mudah di pahami dan digunakan. Adapun perancangan struktur menu dapat dilihat pada gambar 3.11.



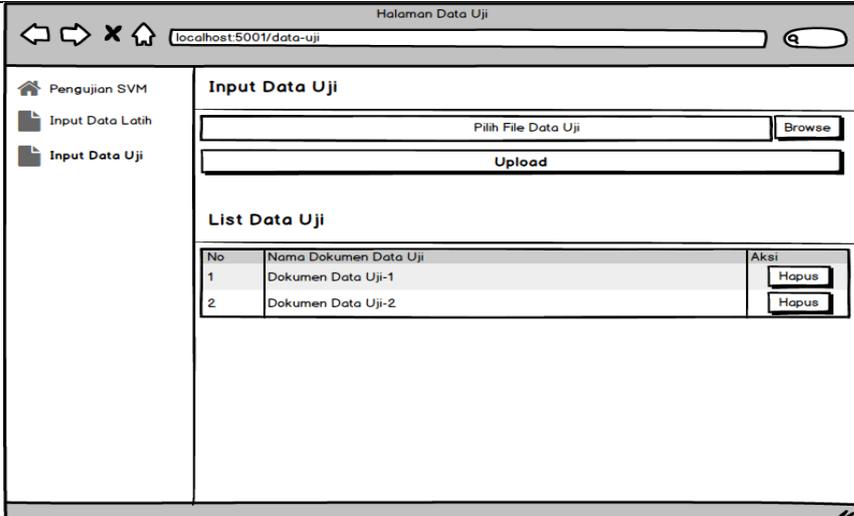
Gambar 3.11 Perancangan Struktur Menu

3.12.2 Perancangan Antarmuka

Perancangan antarmuka mendeskripsikan rencana tampilan yang akan digunakan pada sistem. Perancangan antarmuka pada penelitian ini terdiri dari halaman input data latih, input data uji, halaman pengujian *Support Vector Machine*.

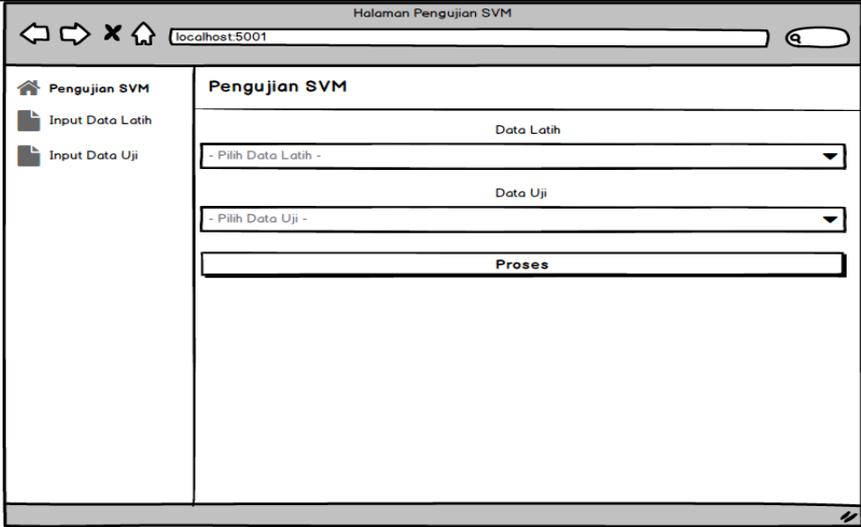
<p>No T01</p> 	<ul style="list-style-type: none"> • Klik tombol “Browse” untuk memilih data latih yang akan diproses • Klik tombol “Upload” setelah memilih data latih maka menampilkan pesan P03 • Klik tombol “Upload” belum memilih data latih maka menampilkan pesan P01 • Klik tombol “Hapus” maka menampilkan pesan P02
<p>Keterangan: Ukuran: 1366x768 pixel Warna Layout: Putih Deskripsi: Form Input Data Latih</p>	

Gambar 3.12 Perancangan antarmuka menu utama input data latih

<p>No T02</p> 	<ul style="list-style-type: none"> • Klik tombol “Browse” untuk memilih data latih yang akan diproses • Klik tombol “Upload” setelah memilih data latih maka menampilkan pesan P06 • Klik tombol “Upload” belum memilih data latih maka menampilkan pesan P04
<p>Keterangan:</p>	

Ukuran: 1366x768 pixel Warna Layout: Putih Deskripsi: Form Input Data Uji	<ul style="list-style-type: none"> • Klik tombol “Hapus” maka menampilkan pesan P05
---	--

Gambar 3.13 Perancangan antarmuka menu utama input data uji

<p>No T03</p> 	<ul style="list-style-type: none"> • Klik drop-down list “Pilih Data Latih” untuk memilih data latih yang akan diproses • Klik drop-down list “Pilih Data Uji” untuk memilih data uji yang akan diproses • Klik tombol “Proses” belum ada data latih dan data uji menampilkan pesan P07 • Klik tombol “Proses” belum ada data latih menampilkan pesan P08 • Klik tombol “Proses” belum ada data uji menampilkan pesan P09
<p>Keterangan: Ukuran: 1366x768 pixel Warna Layout: Putih Deskripsi: Form Input Data Uji</p>	

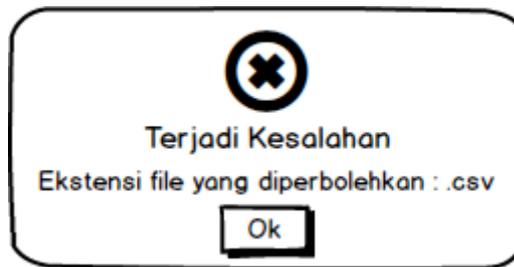
Gambar 3.14 Perancangan antarmuka Pengujian SVM

3.12.3 Perancangan Pesan

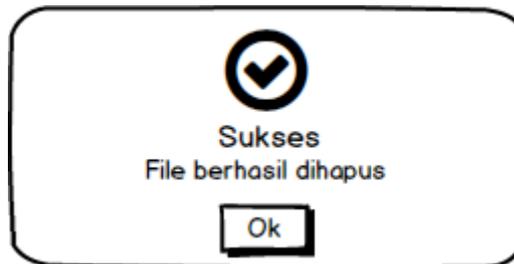
Berikut ini adalah perancangan pesan yang terdapat pada aplikasi yang akan dibangun.

Tabel 3.23 Perancangan Pesan

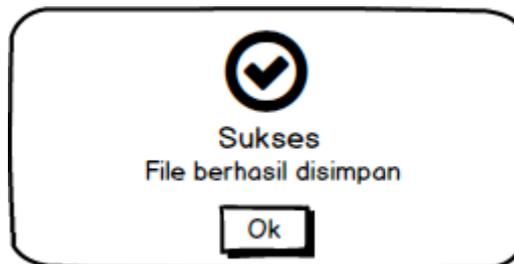
No	Keterangan Pesan
P01	Tampil pada antarmuka T01 apabila menekan tombol “Upload” tapi belum ada data latih yang dipilih
P02	Tampil pada antarmuka T01 apabila berhasil menghapus data latih
P03	Tampil pada antarmuka T01 apabila berhasil mengupload data latih
P04	Tampil pada antarmuka T02 apabila menekan tombol “Upload” tapi belum ada data uji yang dipilih
P05	Tampil pada antarmuka T02 apabila berhasil menghapus data uji
P06	Tampil pada antarmuka T02 apabila berhasil mengupload data uji
P07	Tampil pada antarmuka T03 apabila belum ada data latih dan data uji yang dipilih untuk di proses
P08	Tampil pada antarmuka T03 apabila belum ada data latih yang dipilih untuk di proses
P09	Tampil pada antarmuka T03 apabila belum ada data uji yang dipilih untuk di proses



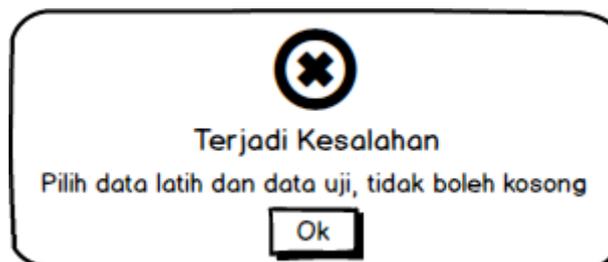
Gambar 3.15 Perancangan Pesan P01 dan P04



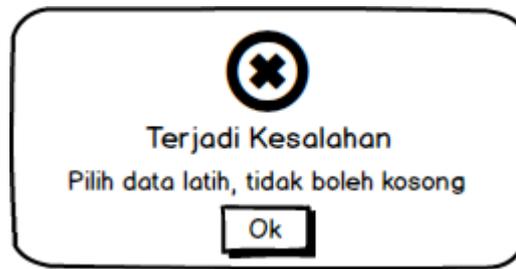
Gambar 3.16 Perancangan Pesan P02 dan P05



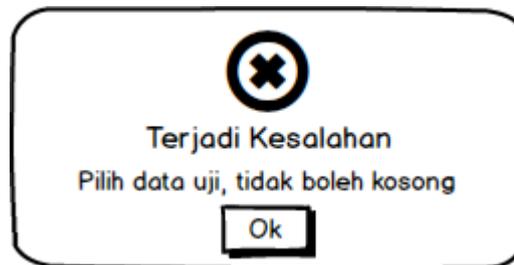
Gambar 3.17 Perancangan Pesan P03 dan P06



Gambar 3.18 Perancangan Pesan P07



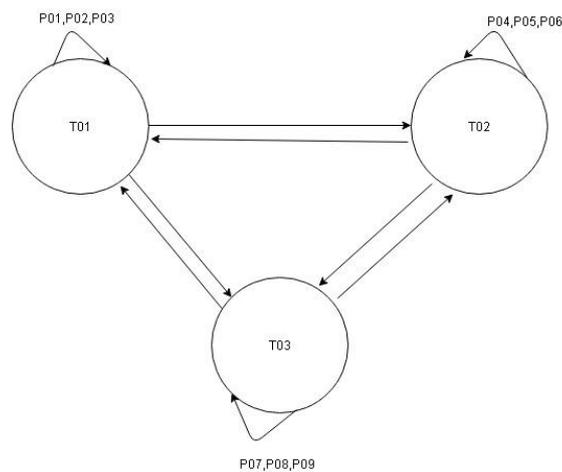
Gambar 3.19 Perancangan P08



Gambar 3.20 Perancangan P09

4.1.1 Jaringan Semantik

Jaringan Semantik memberikan gambaran mengenai keterhubungan dari satu antarmuka ke antarmuka lainnya. Jaringan semantik yang terbentuk dari sistem ini dapat dilihat pada gambar 3.21.



Gambar 3.21 Jaringan Semantik