



**ANALISIS SENTIMEN POSITIF DAN NEGATIF KOMENTAR VIDEO
YOUTUBE MENGGUNAKAN METODE *NAÏVE BAYES – SUPPORT
VECTOR MACHINE (NBSVM) CLASSIFIER***

SKRIPSI

Oleh

Abbi Nizar Muhammad

NIM 152410101150

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS JEMBER**

2019



**ANALISIS SENTIMEN POSITIF DAN NEGATIF KOMENTAR VIDEO
YOUTUBE MENGGUNAKAN METODE *NAÏVE BAYES – SUPPORT
VECTOR MACHINE (NBSVM) CLASSIFIER***

SKRIPSI

Diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat
untuk menyelesaikan pendidikan di Program Studi Sistem Informasi Universitas
Jember dan mendapat gelar Sarjana Sistem Informasi

Oleh

Abbi Nizar Muhammad

NIM 152410101150

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS JEMBER**

2019

PERSEMBAHAN

Skripsi ini saya persembahkan untuk :

1. Allah SWT yang senantiasa memberikan rahmat dan hidayah-Nya untuk mempermudah dan melancarkan dalam mengerjakan skripsi;
2. Ibunda tercinta Diantri Februana dan Ayahanda tercinta Suyoso;
3. Saudara laki-laki Affan Rizaldy Muhammad dan Saudara perempuan Anggun Annisa;
4. Guru-guru baik dari Pendidikan formal maupun informal.
5. Almamater Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Jember.

MOTO

“Kemuliaan orang adalah agamanya, harga dirinya (kehormatannya) adalah akalnya, sedangkan ketinggian kedudukannya adalah akhlaknya”

- Hadist Rasulullah SAW Ahmad –

“The great fights with your strongest rivals are always the bigger motivation when you easily it’s not the same taste”

- Valentino Rossi -

“Jangan berteriak meminta perubahan tetapi pimpin perubahan itu”

- Anonymous -

PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Abbi Nizar Muhammad

NIM : 152410101150

Menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Analisis Sentimen Positif Dan Negatif Komentar Video *YouTube* Menggunakan Metode *Naïve Bayes – Support Vector Machine (NBSVM) Classifier*” adalah benar-benar hasil karya sendiri, kecuali jika dalam pengutipan substansi disebutkan sumbernya, belum pernah diajukan pada institusi mana pun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa adanya tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika di kemudian hari pernyataan ini tidak benar.

Jember, 12 Juli 2019

Yang menyatakan,

Abbi Nizar Muhammad

NIM 152410101150

SKRIPSI

**ANALISIS SENTIMEN POSITIF DAN NEGATIF KOMENTAR VIDEO
YOUTUBE MENGGUNAKAN METODE *NAÏVE BAYES* – *SUPPORT
VECTOR MACHINE (NBSVM) CLASSIFIER***

Oleh

Abbi Nizar Muhammad

NIM 152410101150

Pembimbing

Dosen Pembimbing Utama : Prof. Dr. Saiful Bukhori, ST., M.Kom

Dosen Pembimbing Pendamping : Priza Pandunata, S.Kom., M.Sc

PENGESAHAN PEMBIMBING

Skripsi berjudul “Analisis Sentimen Positif dan Negatif Komentar Video *YouTube* Menggunakan Metode *Naïve Bayes – Support Vector Machine (NBSVM Classifier)*”, Telah diuji dan sisahkan pada:

hari, tanggal : Jumat, 12 Juli 2019

tempat : Program Studi Sistem Informasi Fakultas Ilmu Komputer
Universitas Jember

Disetujui Oleh,

Pembimbing I,

Pembimbing II,

Prof. Dr. Saiful Bukhori, ST., M.Kom

Priza Pandunata, S.Kom., M.Sc

NIP 196704201992011001

NIP 198301312015041001

PENGESAHAN PENGUJI

Skripsi berjudul “Analisis Sentimen Positif dan Negatif Komentar Video *YouTube* Menggunakan Metode *Naïve Bayes – Support Vector Machine (NBSVM) Classifier*”, Telah diuji dan sisahkan pada:

hari, tanggal : Jumat, 12 Juli 2019

tempat : Program Studi Sistem Informasi Fakultas Ilmu Komputer
Universitas Jember

Tim Penguji :

Penguji I,

Penguji II,

Oktalia Juwita, S.Kom., M.MT
NIP 198110202014042001

Diksy Media Firmansyah, S.Kom., M.Kom
NRP 760016853

Mengesahkan

a.n Dekan Fakultas Ilmu Komputer

Wakil Dekan I,

Drs. Antonius Cahya Prihandoko, M.App.Sc.,Ph.D.

NIP 196909281993021001

RINGKASAN

Analisis Sentimen Positif dan Negatif Komentar Video *YouTube* Menggunakan Metode *Naïve Bayes* – *Support Vector Machine (NBSVM) Classifier*; Abbi Nizar Muhammad, 152410101150; 2019, 111 halaman; Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Jember.

Sentimen analisis pada komentar video *YouTube* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapat informasi sentimen yang terkandung dalam suatu kalimat komentar video *YouTube*. Beragam reaksi opini dalam komentar pengguna ketika melihat konten video yang di unggah di *YouTube* sangat mempengaruhi reputasi konten video dan *channel* tersebut. Konten video didalam *YouTube* sangat beragam, salah satu kategori yang mengalami pertumbuhan adalah konten video pendidikan yang menjadi bagian terpenting dalam mengedukasi masyarakat.

Pendekatan *text mining* menjadi alternatif terbaik untuk mengartikan makna dari setiap komentar. Pengklasifikasian konten Positif dan Negatif menjadi sangat penting bagi pengguna *YouTube* untuk menilai seberapa bermakna konten yang telah diterbitkan tersebut berdasarkan opini pengguna. *Naïve Bayes* dan *Support Vector Machine* secara luas digunakan sebagai garis dasar dalam tugas-tugas yang berhubungan dengan teks tetapi kinerjanya bervariasi secara signifikan diseluruh varian, fitur dan jumlah kumpulan data. *Naïve Bayes* sangat baik dalam mengklasifikasikan teks dengan jumlah data yang kecil atau cuplikan dokumen dan *snippets* sedangkan *Support Vector Machine* sangat baik dalam mengklasifikasikan teks dengan jumlah data yang relative banyak atau dokumen yang panjang *full-length*. Kombinasi dari metode *Naïve Bayes* dan *Support Vector Machine* menghasilkan tingkat akurasi yang lebih baik dan performa yang lebih kuat dengan penggunaan skala data 7:3 yaitu 70% data *training* dan 30% data *testing*. Dengan menghasilkan nilai uji performansi terbesar yaitu *precision* sebesar 91%, *recall* sebesar 83% dan *f1score* sebesar 87%.

PRAKATA

Puji syukur kehadiran Allah SWT atas segala rahmad dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Sentimen Positif Dan Negatif Komentar Video *YouTube* Menggunakan Metode *Naïve Bayes – Support Vector Machine (NBSVM) Classifier*”, Skripsi ini disusun untuk memenuhi salah satu syarat menyelesaikan Pendidikan Strata Satu (S1) pada Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Jember.

Pengusunan skripsi ini tidak lepas dari dukungan berbagai pihak. Oleh karena itu penulis menyampaikan terima kasih kepada :

1. Allah SWT yang senantiasa memberikan rahmat dan hidayah-Nya untuk mempermudah dan melancarkan dalam mengerjakan skripsi;
2. Prof. Dr. Saiful Bukhori, ST., M.Kom selaku dosen pembimbing utama dan Priza Pandunata, S.Kom., M.Sc selaku dosen pembimbing pendamping yang telah meluangkan waktu, pikiran dan perhatian dalam penulisan skripsi;
3. Oktalia Juwita, S.Kom., M.MT selaku dosen penguji utama dan Diksy Media Firmansyah, S.Kom., M.Kom selaku penguji anggota yang telah berkenan untuk menguji skripsi ini dan memberikan masukan serta saran untuk pengembangan diri penulis dan skripsi ini;
4. Seluruh bapak dan ibu dosen beserta staf karyawan di Fakultas Ilmu Komputer Universitas Jember;
5. Ibunda tercinta Diantri Februana, Ayahanda tercinta Suyoso dan keluarga besar yang selalu mendukung dan mendoakan;
6. Saudara laki-laki Affan Rizaldy Muhammad dan Saudara perempuan Anggun Annisa;
7. Kepengurusan organisasi Himpunan Mahasiswa Sistem Informasi periode 2016/2017;
8. Kepengurusan komunitas *Developer Student Clubs Chapter* Universitas Jember periode 2018/2019;

9. Kepengurusan komunitas *Young On Top* Jember periode 2018/2019;
10. Kepengurusan asisten Laboratorium Rekayasa Perangkat Lunak periode 2018/2019;
11. Keluarga KKN 274 Jangur gelombang II 2018/2019;
12. Teman-teman dan karyawan di DOT Indonesia periode Januari - Maret 2019;
13. Sahabat seperjuangan Tim *One Step*, M Noor Dwi Eldianto dan Umroh Makhmudah yang telah membantu, memotivasi dan menemani;
14. Tim Lomba LAN Gemastik 2017.
15. Tim Latihan Kepemimpinan Manajemen Mahasiswa Tingkat Menengah dan Tingkat Lanjut Universitas Jember tahun 2017
16. Kepengurusan Divisi PSDM HIMASIF periode 2015/2016;
17. Keluarga SELECTION atas kekeluargaan yang selalu hangat dan menenangkan;
18. Teman-teman Kontrakan kreongan;
19. Teman-teman Program Studi Sistem Informasi di semua angkatan;
20. Teman-teman *Student Executive Board* 2013/2014;
21. Guru-guru baik dari Pendidikan formal maupun informal;
22. Semua pihak yang tidak dapat disebutkan satu persatu;

Dengan harapan bahwa penelitian ini nantinya terus berlanjut dan berkembang kelak, penulis juga menerima segala kritik dan saran dari semua pihak demi kesempurnaan skripsi ini. Akhirnya penulis berharap, semoga skripsi ini dapat bermanfaat bagi semua makhluk.

Jember, 22 Juli 2019

Penulis

DAFTAR ISI

PERSEMBAHAN	ii
MOTO	iii
PERNYATAAN.....	iv
SKRIPSI.....	v
PENGESAHAN PEMBIMBING.....	vi
PENGESAHAN PENGUJI.....	vii
RINGKASAN	viii
PRAKATA	ix
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii
BAB 1. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	4
BAB 2. TINJAUAN PUSTAKA	5
2.1. Text Mining.....	6
2.2. Natural Language Processing.....	6
2.3. Analisis Sentimen.....	8
2.4. Metode Klasifikasi	8

2.4.1.	Multinomial Naïve Bayes	9
2.4.2.	Support Vector Machine	10
2.4.3.	Support Vector Machine dengan fitur Naïve Bayes (NBSVM).....	10
2.5.	Uji Performansi	10
BAB 3. METODE PENELITIAN.....		12
3.1.	Jenis Penelitian	12
3.2.	Tahapan Penelitian	12
3.3.	Studi Literatur.....	13
3.4.	Populasi dan Sampel	13
3.5.	Jenis dan Sumber Data	13
3.6.	Variabel dan Skala Pengukurannya.....	14
3.7.	Perancangan Model Klasifikasi.....	15
3.8.	Implementasi Model Sistem.....	16
3.8.1.	Flow Chart Crawling.....	16
3.8.2.	Flow Chart Casefolding	17
3.8.3.	Flow Chart Tokenization.....	18
3.8.4.	Flow Chart Filtering.....	19
3.8.5.	Flow Chart Stemming	20
3.8.6.	Flow Chart Klasifikasi	21
3.9.	Pengujian Model dan Pengujian Sistem	23
3.10.	Gambaran Algoritma Sistem	23
BAB 4. HASIL DAN PEMBAHASAN.....		27
4.1.	Hasil Dataset.....	27
4.2.	Hasil Implementasi Model Sistem.....	28
4.2.1.	Implementasi Crawling	28

4.2.2.	Implementasi Text Preprocessing	28
4.2.3.	Implementasi Metode Klasifikasi	30
4.2.4.	Implementasi Sentimen Analisis.....	32
4.3.	Hasil Penerapan Crawling	33
4.4.	Hasil Text Preprocessing	33
4.5.	Hasil Perhitungan Naïve Bayes Support Vector Machine	38
4.6.	Hasil Klasifikasi Sentimen	56
4.7.	Hasil Uji Performansi	60
4.8.	Hasil Pengujian Sistem.....	61
4.9.	Visualisasi	64
4.9.1.	Tampilan Halaman Frontpage.....	64
4.9.2.	Tampilan Halaman Dashboard.....	65
4.9.3.	Tampilan Halaman Pencarian Video	66
4.9.4.	Tampilan Halaman Klasifikasi.....	66
4.9.5.	Tampilan Halaman Data Stopword.....	67
4.9.6.	Tampilan Halaman Data Rootword	68
4.9.7.	Tampilan Halaman Data Training.....	68
4.9.8.	Tampilan Halaman Data Kata Positif dan Negatif.....	69
4.9.9.	Tampilan Halaman Data User.....	70
BAB 5.	KESIMPULAN DAN SARAN	71
5.1.	Kesimpulan.....	71
5.2.	Saran.....	72
DAFTAR PUSTAKA	73
LAMPIRAN	75

DAFTAR TABEL

Tabel 2.1 Tabel perbandingan pustaka metode.....	5
Tabel 2.2 Tabel matriks <i>contingency</i>	11
Tabel 3.1 Operasional Variabel.....	14
Tabel 3.2 Penjelasan algoritma sistem.....	24
Tabel 4.1 Contoh Data <i>Training</i>	27
Tabel 4.2 Contoh hasil proses <i>casefolding</i>	34
Tabel 4.3 Contoh hasil proses <i>tokenization</i>	35
Tabel 4.4 Contoh hasil proses <i>filtering</i>	36
Tabel 4.5 Contoh hasil proses <i>stemming</i>	37
Tabel 4.6 Tabel Daftar Kata.....	38
Tabel 4.7 Perhitungan Probabilitas <i>Naïve Bayes</i>	39
Tabel 4.8 Perhitungan <i>Hyperplane Support Vector Machine</i>	56
Tabel 4.9 Pengujian model klasifikasi	56
Tabel 4.10 Hasil dari klasifikasi positif dan negatif	57
Tabel 4.11 Perhitungan uji performansi.....	60
Tabel 4.12 Uji performansi tiap skala data	61
Tabel 4.13 Uji Performansi <i>Confusion Matrix</i>	61
Tabel 4.14 Pengujian <i>Black Box</i> Sistem	61

DAFTAR GAMBAR

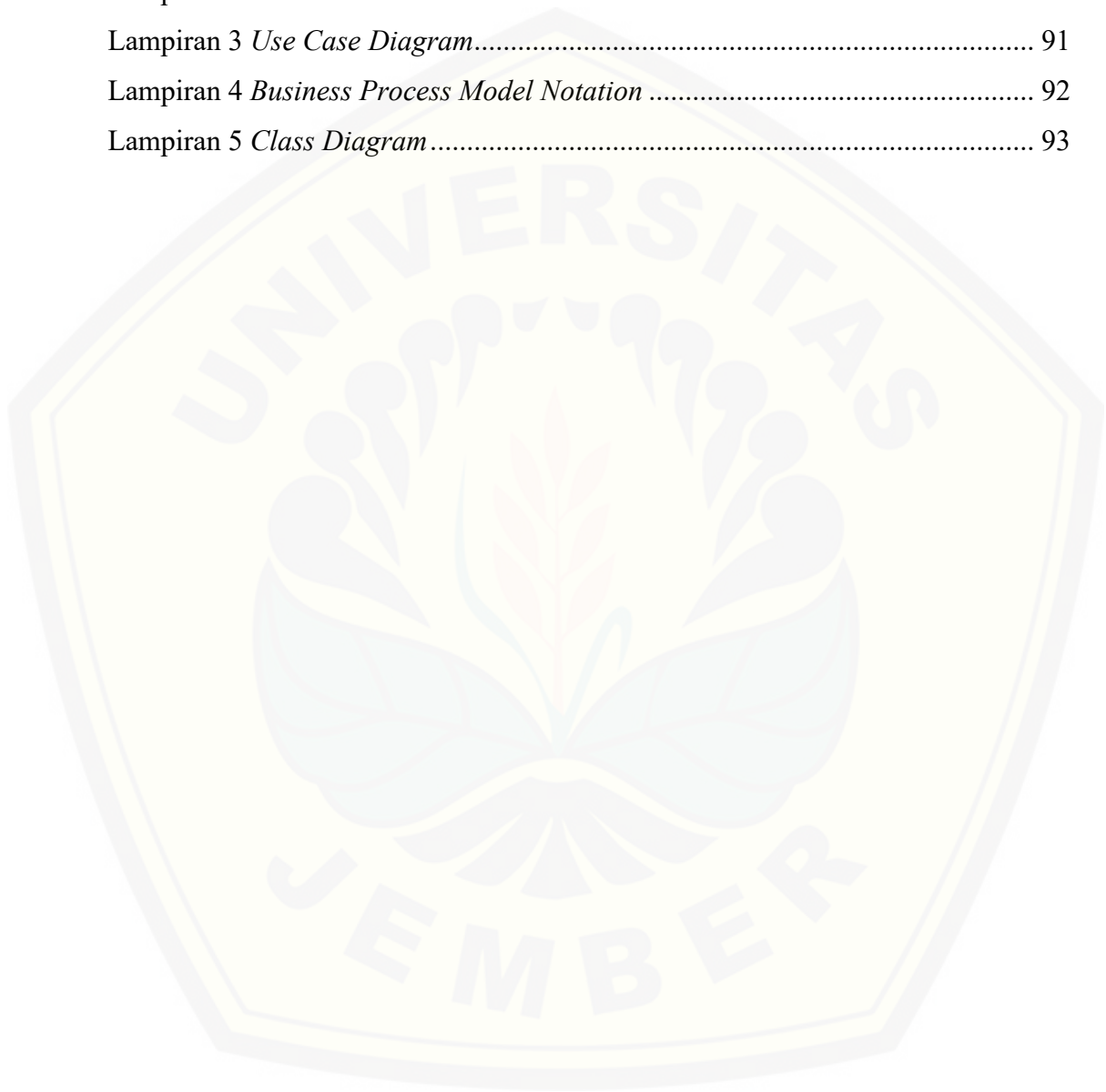
Gambar 3.1 Tahapan Penelitian	12
Gambar 3.2 Data <i>Flow Input Output</i>	14
Gambar 3.3 Gambaran <i>Hybrid Method NBSVM</i>	15
Gambar 3.4 <i>Flow Chart Crawling</i>	17
Gambar 3.5 <i>Flow Chart Casefolding</i>	18
Gambar 3.6 <i>Flow Chart Tokenization</i>	19
Gambar 3.7 <i>Flow Chart Filtering</i>	20
Gambar 3.8 <i>Flow Chart Stemming</i>	21
Gambar 3.9 <i>Flow Chart</i> Klasifikasi	22
Gambar 3.10 Gambaran Algoritma Sistem	23
Gambar 4.1 Implementasi kode <i>crawling</i>	28
Gambar 4.2 Implementasi kode <i>casefolding</i>	29
Gambar 4.3 Implementasi kode <i>tokenization</i>	29
Gambar 4.4 Implementasi kode <i>filtering</i>	29
Gambar 4.5 Implementasi kode <i>stemming</i>	30
Gambar 4.6 Implementasi kode <i>Naive Bayes</i>	30
Gambar 4.7 Implementasi kode <i>Support Vector Machine</i>	31
Gambar 4.8 Implementasi kode menghitung persamaan SVM	32
Gambar 4.9 Implementasi kode sentimen analisis NBSVM.....	32
Gambar 4.10 Alur Proses <i>Crawling</i>	33
Gambar 4.11 Halaman <i>frontpage</i>	65
Gambar 4.12 Halaman <i>Dashboard</i>	65
Gambar 4.13 Halaman pencarian video	66
Gambar 4.14 Halaman Klasifikasi	67
Gambar 4.15 Halaman Data Stopword	67
Gambar 4.16 Halaman Data <i>Rootword</i>	68
Gambar 4.17 Halaman Data <i>Training</i>	69
Gambar 4.18 Halaman Data Positif	69
Gambar 4.19 Halaman Data Negatif.....	70

Gambar 4.20 Halaman Data *User* 70



DAFTAR LAMPIRAN

Lampiran 1 Surat keterangan validasi data	75
Lampiran 2 Validasi data oleh ahli bahasa	76
Lampiran 3 <i>Use Case Diagram</i>	91
Lampiran 4 <i>Business Process Model Notation</i>	92
Lampiran 5 <i>Class Diagram</i>	93



BAB 1. PENDAHULUAN

1.1. Latar Belakang

Pada tahun 2018 pengguna YouTube mencapai 1,5 miliar dari seluruh dunia dan angka ini diproyeksikan akan tumbuh menjadi 1,86 miliar pengguna pada tahun 2021 (J. Clement, 2018). *YouTube* menjadi platform video terbesar di dunia dengan menampilkan beragam konten media yang dibuat oleh perusahaan atau individual yang mencakup video musik, video promosi produk, video blog, video review, video edukasi (Gordon, 2018). Meningkatnya pengguna berbanding lurus dengan bertambah banyaknya konten video yang diunggah di *YouTube*. Hal ini menjadi tempat semua orang untuk berlomba membuat konten video dan mendapat pendapatan dari video yang diunggah. Konten video yang baik akan menjadi alat untuk meningkatkan view like dan subscribe channel *YouTube*.

Konten video didalam *YouTube* sangat beragam, salah satu kategori yang mengalami pertumbuhan adalah konten video edukasi. Para konten *creator* berlomba lomba untuk membuat konten video edukasi yang dapat di terima oleh pengguna *YouTube*. Di Indonesia terdapat beberapa konten creator video edukasi, salah satunya adalah “kok bisa ?” yang menjadi *top ranking* konten edukasi di Indonesia dengan jumlah *subscriber* 1,416,935 ribu dan *views* sebesar 142,957,953 ribu (socialblade, 2019). Konten edukasi menjadi hal yang paling penting untuk diperhatikan karena data dan ilmu pengetahuan yang disampaikan kepada pengguna harus dapat dipelajari dengan baik dan tidak menyesatkan.

Beragam reaksi opini dalam komentar pengguna ketika melihat konten video yang di unggah di *YouTube* sangat mempengaruhi reputasi konten video dan *channel* tersebut. Dalam menilai konten video terdapat dua ulasan yaitu ulasan *like* dan *dislike* dan ulasan komentar. Ulasan *like* dan *dislike* menunjukkan evaluasi pengguna dalam melihat konten video dalam angka, namun komentar tekstual mampu bercerita lebih mendalam tentang pengalaman pengguna ketika melihat konten video. Dengan banyaknya komentar tekstual membuat konten kreator mengalami kesulitan dalam mencari tahu bagaimana konten video diterima baik oleh pengguna.

Bagian terpenting dari pengumpulan informasi adalah mencari tahu cara orang lain berfikir (Pang B. L., 2008). Memahami sesuatu yang difikirkan orang menjadi informasi yang sangat penting bagi konten kreator untuk membuat video yang bisa diterima pengguna dengan baik. Sehingga dapat dibuat sebuah pendekatan untuk mengetahui persepsi pengguna *YouTube* terhadap suatu konten video dengan menggunakan data analisis sentimen yang didapat dari komentar tekstual. Pendekatan *text mining* dengan *Natural Language Processing* menjadi alternative terbaik untuk mengartikan makna dari setiap komentar. Pengklasifikasian konten Positif dan Negatif menjadi sangat penting bagi pengguna *YouTube* untuk menilai seberapa bermakna konten yang telah diterbitkan tersebut berdasarkan opini pengguna dalam komentar (Tanesab, 2017). Dengan adanya pendekatan ini, banyak manfaat yang ditimbulkan dibandingkan menghitung jumlah like dan membaca satu per satu komentar yang ada, konten kreator dapat secara mudah mendapat hasil penilaian dari video yang telah diterbitkan. Pengurangan tenaga dan waktu yang dikeluarkan karena data dapat diambil secara cepat dan mudah melalui *crawling* di *YouTube*. Pengguna *YouTube* secara mudah dapat menikmati konten video yang bermanfaat dan dapat melihat apakah konten video edukasi diterima secara baik dari referensi komentar dan opini pengguna lain di *YouTube*.

Penelitian ini ingin memberikan kontribusi terhadap kemajuan pendekatan analisis sentimen menggunakan metode *Naïve Bayes - Support Vector Machine (NBSVM)* dengan pendekatan *Binary Classification*. Penggunaan metode ini dipilih karena metode klasifikasi ini bekerja sangat baik pada cuplikan dokumen atau *snippets* dan dokumen yang lebih panjang untuk analisis sentimen, topik dan klasifikasi subjektif dan seringkali lebih baik daripada hasil yang dipublikasikan sebelumnya (Sida Wang, 2012). *Naïve Bayes* dan *Support Vector Machine* secara luas digunakan sebagai garis dasar dalam tugas-tugas yang berhubungan dengan teks tetapi kinerjanya bervariasi secara signifikan diseluruh varian, fitur dan jumlah kumpulan data. *Naïve Bayes* sangat baik dalam mengklasifikasikan teks dengan jumlah data yang kecil atau cuplikan dokumen dan *snippets* sedangkan *Support Vector Machine* sangat baik dalam mengklasifikasikan teks dengan jumlah data

yang relative banyak atau dokumen yang panjang *full-length*. Kombinasi dari metode *Naïve Bayes* dan *Support Vector Machine* menghasilkan tingkat akurasi yang lebih baik dan performa yang lebih kuat. Akan tetapi Kombinasi dari beberapa pengklasifikasi tidak selalu meningkat akurasi klasifikasinya dibandingkan dengan klasifikasi yang hanya menggunakan satu metode (JAIN, 1998).

1.2. Rumusan Masalah

Berdasarkan uraian yang telah disampaikan dalam latar belakang. Adapun rumusan masalah dalam penelitian ini adalah :

1. Bagaimana menerapkan metode *Naïve Bayes - Support Vector Machine (NBSVM)* dalam mengklasifikasikan data komentar video *YouTube*?
2. Bagaimana mengetahui performansi klasifikasi dengan menerapkan metode *Naïve Bayes - Support Vector Machine (NBSVM)*?

1.3. Batasan Masalah

Agar tidak terjadi penyimpangan dalam proses penelitian dan pengembangan sistem, maka ditetapkan Batasan masalah dalam penelitian ini. adapun Batasan masalahnya adalah:

1. Data diambil menggunakan *Application Programming Interface (API)* Data *YouTube* dengan *API key* penulis
2. Data komentas yang diambil menggunakan Bahasa Indonesia yang benar.
3. Komentar yang didapat dikelompokkan kedalam 2 analisis sentimen yaitu Positif dan Negatif
4. Algoritma yang digunakan dalam pengklasifikasian adalah *Naïve Bayes - Support Vector Machine (NBSVM)* tanpa membandingkannya dengan metode lain
5. Menggunakan konsep *Natural Language Processing* dalam proses *preprocessing text mining*
6. Dalam menjalankan sistem, apabila *API YouTube* yang disediakan *down* maka proses *crawling* juga berhenti

7. Kategori video yang dipilih adalah kategori Edukasi dengan melihat ranking dunia, jumlah *subscribe* dan jumlah *view* video dengan mengambil data dari *website Socialblade*

1.4. Tujuan Penelitian

Tujuan dalam penelitian ini merupakan jawaban dari rumusan masalah yang telah disampaikan. Adapun tujuan penelitian ini adalah:

1. Mengetahui cara penerapan metode *Naïve Bayes - Support Vector Machine (NBSVM)* dalam mengklasifikasikan data komentar *YouTube*
2. Mengetahui uji performansi klasifikasi dalam penerapan *Naïve Bayes - Support Vector Machine (NBSVM)*

1.5. Manfaat Penelitian

Manfaat dalam penelitian ini merupakan hasil yang dirasakan oleh semua pihak yang terlibat dalam penelitian ini. Adapun manfaat dari penelitian ini adalah:

1. Bagi pengguna *YouTube*
Penelitian ini memberikan informasi bagi pengguna *YouTube* untuk menilai sebuah konten video edukasi berdasarkan ekstraksi dan penambangan data dari komentar dan opini masyarakat di *YouTube*
2. Bagi konten *creator*
Penelitian ini memberikan informasi yang bermanfaat untuk konten kreator dalam memahami *viewers* di *YouTube*. Hasil dari analisis sentiment dapat mengurangi tenaga dan waktu untuk menghasilkan makna dari ribuan komentar. Konten kreator dapat mengetahui seberapa baik konten video yang di terbitkan diterima baik oleh pengguna *YouTube*
3. Bagi penulis
Penelitian ini memberikan wawasan dan kemampuan dalam pemecahan masalah text mining dalam analisis sentimen dalam *YouTube*. Penggunaan *Hybrid Method Naïve Bayes* dengan *Support Vector Machine* memberikan dampak positif bagi kontribusi ilmu dalam *machine learning* dengan meningkatkan akurasi dalam pemecahan masalah analisis sentiment.

BAB 2. TINJAUAN PUSTAKA

Berdasarkan penelitian yang dilakukan, acuan dari beberapa penelitian terdahulu menjadi sangat penting dalam melakukan sebuah penelitian dengan tujuan untuk mengetahui hubungan antara penelitian yang akan dilakukan dengan penelitian terdahulu. Fokus penelitian ini yaitu bagaimana sebuah algoritma *Naïve Bayes Support Vector Machine* dapat mengklasifikasikan komentar menjadi positif dan negatif. Perbandingan penelitian terdahulu dapat dilihat pada Tabel 2.1.

Tabel 2.1 Tabel perbandingan pustaka metode

No	Penulis	Metode	Keterangan
1.	Sida Wang and Christopher D. Manning (<i>Baselines and Bigrams: Simple, Good Sentiment and Topic Classification, 2012</i>)	<i>Naïve Bayes - Support Vector Machine (NBSVM)</i>	<i>Multinomial Naïve Bayes</i> sebesar 86,59, <i>Support Vector Machine</i> 89,16 dan <i>Naïve Bayes - Support Vector Machine (NBSVM)</i> sebesar 91,22.
2.	Lutfi Budi Ilmawan dan Edi Winarko (Aplikasi mobile untuk analisis sentimen pada <i>Google Play, 2015</i>)	<i>Naïve Bayes Classifier</i>	Nilai akurasi <i>Naïve Bayes Classifier</i> sebesar 83,87% lebih rendah jika dibandingkan dengan nilai akurasi dari <i>Support Vector Machine Linier Classifier</i> sebesar 89,49%.
3.	Fatimah Wulandari dan Anto Satriyo Nugroho	<i>Support Vector Machine</i>	Eksperiment dilakukan pada 3713

	(Text Classification Using Support Vector Machine for Webmining based spatio temporal analysis of the spread of tropical diseases, 2009)		features dan 360 instances. Menunjukkan Support Vector Machine menghasilkan nilai akurasi 92,5 % dibandingkan Naïve Bayes Classifier yang nilai akurasinya 90 %
--	--	--	--

Berdasarkan penelitian terdahulu yang dibandingkan metode *Naïve Bayes* dan *Support Vector Machine* saling mengungguli dan bekerja dengan baik. Perbedaan hasil dan akurasi setiap penelitian bergantung pada data yang di uji seberapa banyak varian, fitur dan jumlah kumpulan data yang diuji. Penelitian Sida wang yang berjudul *Baselines and Bigrams: Simple, Good Sentiment and Topic Classification* menggambarkan penggabungan dua metode *Naïve Bayes* dan *Support Vector Machine* untuk mencari nilai akurasi yang lebih tinggi. Hasilnya dengan data uji adalah *review film* menghasilkan nilai akurasi sebesar 91,22.

2.1. Text Mining

Text mining adalah proses ekstraksi pola (informasi dan pengetahuan) dari sejumlah besar sumber data tak terstruktur. Penambangan *text* memiliki tujuan dan menggunakan proses yang sama dengan penambangan data, namun memiliki masukan yang berbeda. Masukan untuk penambangan *text* adalah kata yang tidak terstruktur, seperti dokumen, word, pdf, kutikapan, komentar, *review* dan lain lain, sedangkan masukan untuk penambangan data adalah data yang terstruktur (Feldman, 2007)

2.2. Natural Language Processing

Natural Language Processing adalah bagian *computer science* dan *artificial intelligence* yang menangani *human languages*. Terdapat 2 bagian riset NLP yaitu *Natural Language Understanding* dan *Natural Language Generation*. *Natural*

Language Understanding adalah proses memetakan input yang diberikan kedalam Bahasa natural menjadi representasi yang bermanfaat dan menganalisa aspek-aspek dalam Bahasa tersebut, sedangkan *Natural Language Generation* adalah proses untuk menghasilkan frasa yang bermakna dan kalimat dengan bentuk NLP representasi dari dalam. Tahap paling penting melakukan *text mining* menggunakan NLP adalah tahap *preprocessing* karena pada umumnya *text* yang akan dilakukan proses text mining memiliki karakteristik yang berbeda yaitu memiliki dimensi yang tinggi, terdapat noise pada data, dan terdapat struktur *text* yang tidak baik (Steven Bird, 2009). Secara umum tahapan *preprocessing* pada dokumen adalah:

1. Case Folding

Case folding adalah proses mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf alfabet yang dapat diterima yaitu “a” sampai “z”. karakter selain huruf dihilangkan dan dianggap delimiter.

2. Tokenizing

Tokenization adalah proses untuk membagi teks yang dapat berupa kalimat, paragraph atau document menjadi token/bagian -bagian struktur atau unit yang bisa digunakan. Salah satunya bentuk membaginya menggunakan *Ngram*, yaitu token dari sejumlah kata yang tertulis secara berurutan.

3. Filtering

Filtering adalah proses mengambil kata-kata penting dari hasil *tokenizing*. Menggunakan algoritma *stoplist* atau *wordlist*. *Stoplist* atau *wordlist* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopwords* adalah “yang”, “dan”, “di” dan seterusnya.

4. Stemming

Stemming adalah proses mengubah kata yang terdapat imbuhan menjadi kata dasarnya. Tahap ini mencari *rootword* dari setiap kata hasil *filtering*. Resiko dari proses *stemming* adalah hilangnya informasi dari kata yang di-stem. Hal ini menghasilkan menurunnya akurasi atau presisi. Sedangkan untuk keuntungannya adalah, proses stemming bisa meningkatkan kemampuan untuk melakukan *recall*.

Tujuan dari *stemming* sebenarnya adalah untuk meningkatkan performa dan mengurangi penggunaan resource dari sistem dengan mengurangi jumlah *unique word* yang harus diakomodasikan oleh sistem

2.3. Analisis Sentimen

Analisis sentimen atau *opinion mining* merupakan proses memahami mengekstrak dan mengolah data tekstual secara otomatis untuk mendapat informasi sentimen yang terkandung dalam suatu kalimat opini. Sentimen analysis dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif (Pang B. , 2002). Berdasarkan sumber datanya, analisis sentimen dapat dibedakan menjadi dua kategori, yaitu:

1. Coarse grained sentimental analysis

Sentimen analisis dilakukan pada tingkat dokumen. *Coarse grained sentimental analysis* menganggap bahwa seluruh isi dokumen sebagai sebuah sentimen positif dan sentimen negatif.

2. Fined grained sentimental analysis

Fined grained sentimental analysis adalah sentiment analisis yang dilakukan pada tingkat kalimat. Pada jenis ini dikatakan bahwa setiap kalimat memiliki sentiment yang berbeda walaupun terdapat dalam satu dokumen.

2.4. Metode Klasifikasi

Naïve Bayes bekerja sangat bagus dalam tugas sentiment yang singkat sedangkan *Support Vector Machine* sangat bagus untuk dokumen yang lebih Panjang. *Hybrid Methode Naïve Bayes* dengan *Support Vector Machine* menghasilkan nilai akurasi yang lebih tinggi dengan varian *Support Vector Machine* yang menggunakan rasio penghitungan *log Naïve Bayes* sebagai nilai fitur secara konsisten berkinerja baik di seluruh tugas dan kumpulan data (Sida Wang, 2012).

Sebagai model pengklasifikasian *linier* varian model utama untuk memprediksi untuk kasus uji k adalah sebagai berikut:

$$y^{(k)} = \text{sign}(w^T x^{(k)} + b) \quad (1)$$

Jadi $F^{(i)}$ menjadi vektor fitur untuk kasus pelatihan i dengan label *binary* $y^{(i)} \in -1, 1$. Tentukan 2 hitungan vektor P dan Q sebagai berikut

$$P = a + \sum_{i:y^{(i)}=1} f(i) \quad (2)$$

$$q = a + \sum_{i:y^{(i)}=-1} f(i) \quad (3)$$

Dimana α adalah *smoothing parameter* dan *Log-count ratio* dapat didefinisikan sebagai berikut:

$$r = \log \left(\frac{P/\|p\|_1}{q/\|p\|_1} \right) \quad (4)$$

2.4.1. Multinomial Naïve Bayes

Dalam *Naïve Bayes* fitur vektor merepresentasikan frekuensi dengan *events* yang dihasilkan oleh distribusi *multinomial* $P = (p_1, \dots, p_n)$. Fitur vektor $x = (x_1, \dots, x_2)$ adalah *histogram*, dimana x_k adalah number dari k yang diamati dari beberapa *event* dalam contoh tertentu. Dengan asumsi *multinomial* dan asumsi *Naïve Bayes*, Likelihood ratio x bersyarat pada y dipengaruhi oleh

$$P(X|Y) = \frac{\sum_k x_k! + 1}{\prod_k x_k! + 1} \prod_k p_{y_k}^{x_k} \quad (5)$$

$$P(y|x) = \frac{p(y)p(x|y)}{p(x)} \alpha p(y)p(x|y) \quad (6)$$

$$P(y|x) = \log p(y|x) \alpha \log [p(y) \prod_k p_{y_k}^{x_k}] \quad (7)$$

$$P(y|x) = \log p(y) + \sum_k x_k \log p_{y_k} \quad (8)$$

Multinomial Naïve Bayes Classifier menjadi *linear classifier* ketika diekspresikan dalam *log-space*, oleh sebab itu persamaan 8 dapat berubah menjadi persamaan 1 menjadi

$$y = \operatorname{argmax} \log p(y) + \sum_k x_k \log p_k \quad (9)$$

$$y = \operatorname{sign} \left(\log \frac{p(y=1)}{p(y=-1)} + \sum_k x_k \log \frac{p_{1k}}{p_{-1k}} \right) \quad (10)$$

$$y = \operatorname{sign} \left(\log \frac{N_+}{N_-} + r^T f \right) \quad (11)$$

2.4.2. Support Vector Machine

Support Vector Machine adalah metode machine learning yang bekerja atas prinsip *structural risk minimization* (SRM) dengan tujuan menemukan hyperplane terbaik yang memisahkan dua buah class pada input space. $x^{(k)} = f^{(k)}$, and w, b diperoleh dengan meminimalkan *loss function*.

$$L(w, b) = w^T w + C \sum \max(0, 1 - y^{(i)} (w^T F^{(i)} + b))^2 \quad (12)$$

2.4.3. Support Vector Machine dengan fitur Naïve Bayes (NBSVM)

Pada *Multinomial Naïve Bayes* dan *Support Vector Machine* menggunakan Log likelihood ratio sebagai feature vector dengan $x = r \circ f$. Sehingga metode ini bekerja sangat baik untuk semua *document* dan penggunaan metode *Naïve Bayes – Support Vector Machine Machine (NBSVM)* menggunakan model berikut:

$$w^t = (1 - \beta) \omega + \beta w \quad (13)$$

Keterangan:

ω = $\|w\| / |V|$ adalah besarnya rata rata w

β = $[0,1]$ adalah parameter interpolasi

2.5. Uji Performansi

Pengujian dilakukan dengan mengukur keakuratan metode dan kesesuaian sistem yang dibangun. Pada pengujian performansi metode klasifikasi diukur menggunakan uji performansi yang meliputi *precision, recall dan f1 score*. Tabel mengenai uji performansi kelas prediksi dan actual (Ponilan, 2016) dapat dilihat pada Tabel 2.2.

Tabel 2.2 Tabel matriks *contingency*

		Predicted Class	
		Class = Yes	Class = No
Actual Class	Class = Yes	TP	FN
	Class = No	FP	TN

Keterangan:

TP (*True Positive*) adalah kelas yang diprediksi *yes* dan ternyata faktanya *yes*

TN (*True Negative*) adalah kelas yang diprediksi *no* dan ternyata faktanya *no*

FP (*False Positive*) adalah kelas yang diprediksi *yes* dan ternyata faktanya *no*

FN (*False Negative*) adalah kelas yang prediksi *no* dan ternyata faktanya *yes*

Precision adalah rasio jumlah ketepatan prediksi suatu kelas terhadap jumlah total prediksi yang diklasifikasikan dalam kelas tersebut. Rumus mencari *precision* dapat ditulis sebagai berikut

$$precision (p) = \frac{TP}{TP + FP} \quad (14)$$

Recall adalah rasio jumlah ketepatan prediksi suatu kelas terhadap jumlah total fakta yang diklasifikasikan dalam kelas tersebut. Rumus mencari *recall* dapat ditulis sebagai berikut

$$recall (r) = \frac{TP}{TP + FN} \quad (15)$$

Untuk menggabungkan keduanya maka digunakan perhitungan *f1score*. Rumus mencari *f1score* dapat ditulis sebagai berikut

$$f1\ score = \frac{2 * precision * recall}{precision + recall} \quad (16)$$

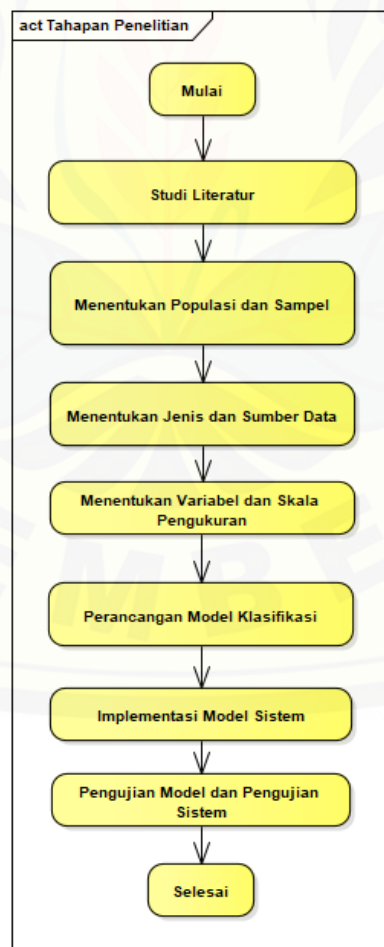
BAB 3. METODE PENELITIAN

3.1. Jenis Penelitian

Penelitian yang dilakukan merupakan penelitian Eksperimen. Penelitian eksperimen dapat dikatakan sebagai metode penelitian yang digunakan untuk mencari pengaruh perlakuan tertentu terhadap yang lain dalam kondisi yang terkendalikan (Sugiyono, 2008). Penelitian ini merupakan metode inti dari model penelitian yang mengandung pendekatan kuantitatif karena terdapat proses mengontrol, kegiatan manipulasi dan observasi.

3.2. Tahapan Penelitian

Tahapan penelitian yang dilakukan pada penelitian ini digambarkan melalui diagram yang ditunjukkan pada Gambar 3.1.



Gambar 3.1 Tahapan Penelitian

3.3. Studi Literatur

Studi literatur merupakan tahapan awal dalam penelitian ini, yang bertujuan untuk penyusunan dasar teori, dengan cara mengumpulkan dan membaca buku, jurnal, artikel ilmiah serta sumber lain yang mendukung penelitian ini.

3.4. Populasi dan Sampel

Populasi dari penelitian ini adalah komentar video *YouTube* dengan kategori Pendidikan. Data yang diperoleh untuk melakukan penelitian didapat langsung dari komentar di konten video *YouTube* menggunakan Teknik *crawling* dengan *YouTube API* dari penulis. *Channel* yang dipilih sudah ditentukan berdasarkan kategori pendidikan yaitu *Channel* “Kok Bisa ?” yang memiliki total grade B *Channel YouTube* di Indonesia dengan total mencapai 1.416.935 *subscriber* dan *views* sebesar 142.957.953 (socialblade, 2019).

Data ini diperoleh dari *website socialblade.com* pada 27 Juni 2019. Dalam mengekstraksi data dari hasil *crawling* menggunakan pendekatan *Natural Language Processing* untuk mendapatkan data yang akurat.

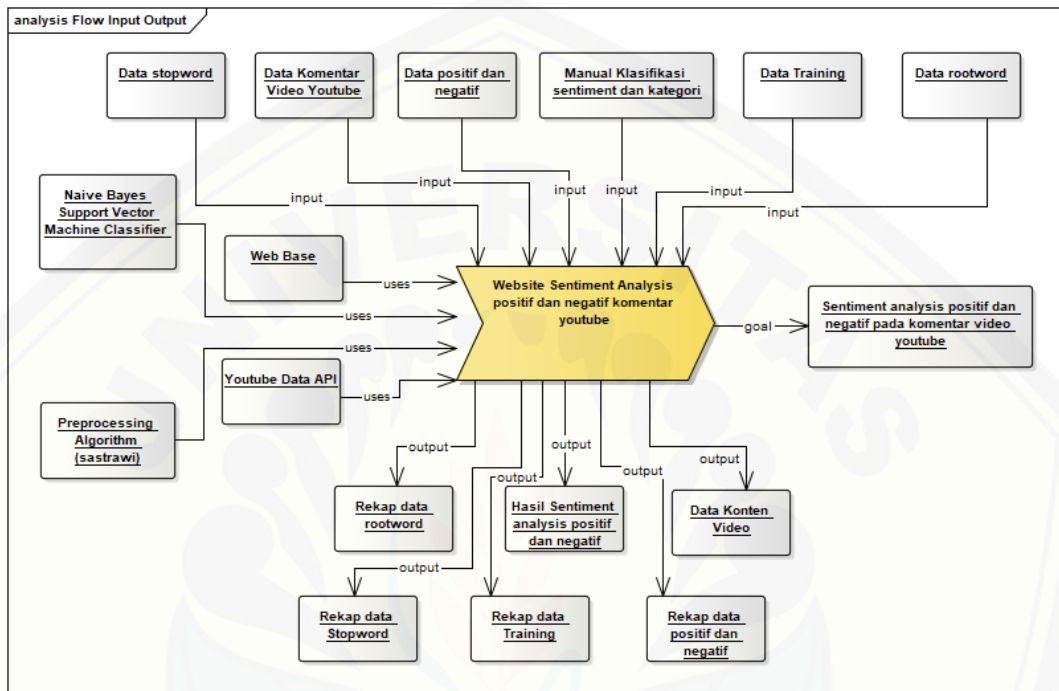
3.5. Jenis dan Sumber Data

Jenis dan sumber data dalam penelitian ini diperoleh dari penelitian sebelumnya yang divalidasi oleh ahli bahasa. Pada tahap preprocessing pada penggunaan data *stopword* dalam Bahasa Indonesia pada proses *stemming* menggunakan sumber dari Tala dengan jumlah 758 data (Tala, 2003). Sedangkan untuk kata dasar (*rootword*) menggunakan data dari kamus besar Bahasa Indonesia yang sudah terdapat dalam library sastrawi dengan jumlah 28.533 data dan untuk referensi kata positif dan negatif menggunakan sumber data dari penelitian dari wahid dengan jumlah 3.587 data (Wahid, 2016) .

Data *training* yang digunakan untuk membangun model klasifikasi diambil dari komentar *YouTube* yang sudah diklasifikasikan secara manual menjadi dua kelas yaitu positif dan negatif. Sedangkan data *testing* merupakan data komentar yang diperoleh hasil *crawling* menggunakan *YouTube* data API.

Website client melakukan *request* untuk mendapatkan komentar dari *YouTube* kemudian *YouTube* mengembalikan data ke *website client* berupa komentar-komentar yang akan dijadikan sebagai data. Data ini digunakan untuk membangun

model klasifikasi melalui proses *training*. Komentar tersebut selanjutnya akan melalui tahap *preprocessing* kemudian dikirimkan untuk diklasifikasikan, data yang dikembalikan merupakan hasil klasifikasi. Proses input dan output data yang digunakan digambarkan dalam Gambar 3.2.



Gambar 3.2 Data Flow Input Output

3.6. Variabel dan Skala Pengukurannya

Variabel positif dan negatif ditentukan untuk mengelompokkan data kedalam dua kelas. Pada data training menentukan skala pengukuran positif dan negative dilakukan secara manual. *Labeling* manual dilakukan berdasarkan bentuk kata oleh ahli Bahasa yaitu lulusan mahasiswa jurusan Bahasa Indonesia Fakultas Keilmuan dan Keguruan Universitas Jember. Penjelasan dan definisi operasional variabel dalam penelitian ini digambarkan dalam Tabel 3.1.

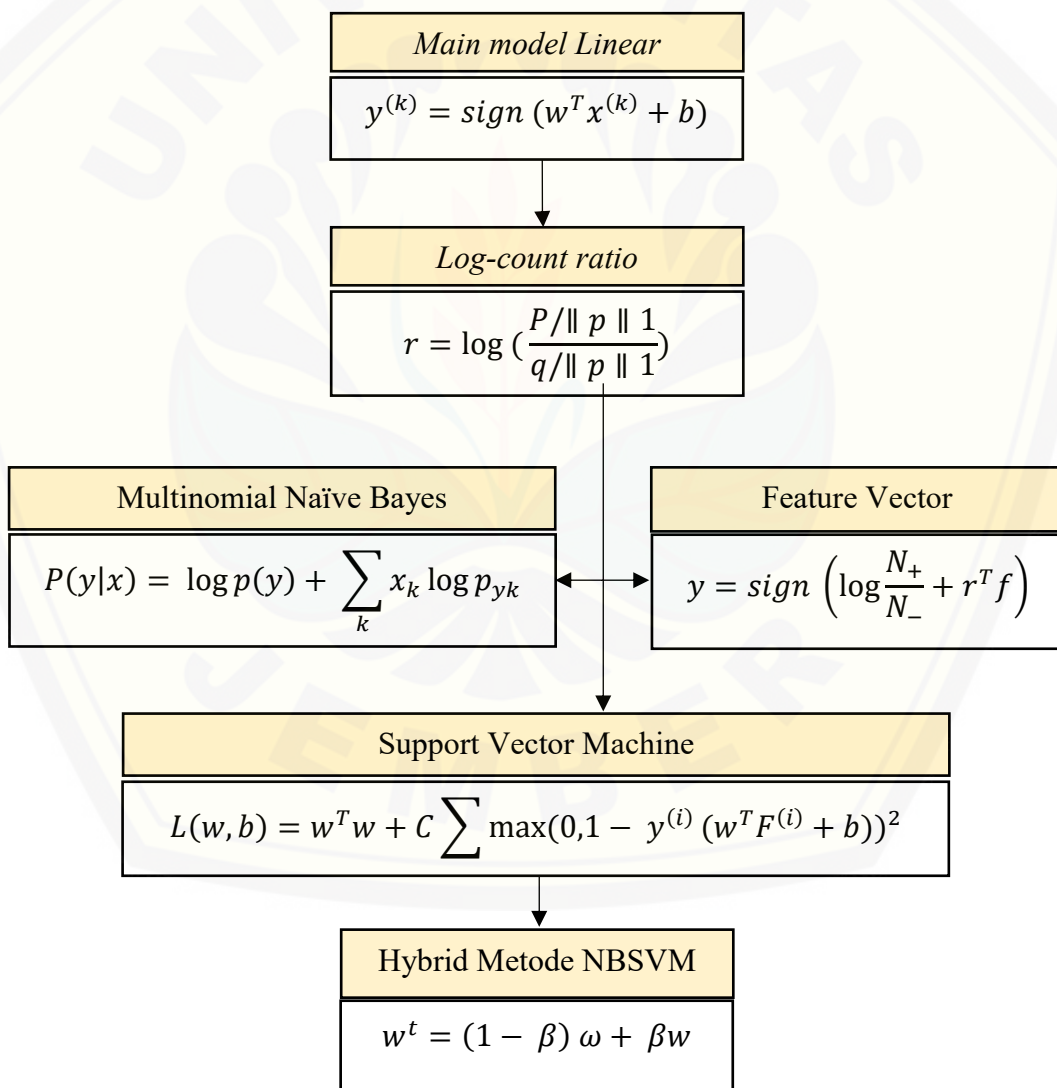
Tabel 3.1 Operasional Variabel.

Variable	Operasional Variabel	Nilai variabel
Positif	konten yang menarik dan bisa dibagikan	1
Positif	channel edukasi terhebat di indonesia semoga kedepan bisa banyak channel bermanfaat	1

Negatif	semua membosankan dan buruk	-1
Negatif	orang orang malas membaca berita jadi kurang tertarik sama berita	-1

3.7. Perancangan Model Klasifikasi

Pembangunan model penelitian ini menggunakan website dengan Bahasa pemrograman PHP. *Hybrid Method* penelitian ini adalah menggabungkan metode *Naïve Bayes* dengan *Support Vector Machine* dan *Log-count ratio* sebagai nilai fitur untuk mengelompokkan data menjadi kelas positif dan negative (Sida Wang, 2012). Perancangan alur model klasifikasi dijelaskan pada Gambar 3.3.



Gambar 3.3 Gambaran Hybrid Method NBSVM

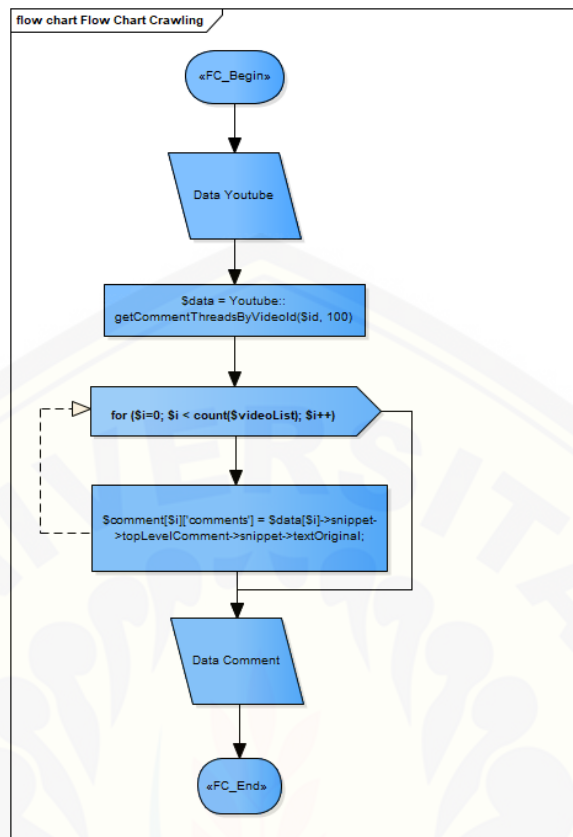
3.8. Implementasi Model Sistem

Implementasi model sistem adalah proses penulisan kode model *text preprocessing* dan model klasifikasi yang menggunakan metode *Naïve Bayes - Support Vector Machine* kedalam program komputer. Dalam penelitian ini menggunakan Bahasa pemrograman *PHP: Hypertext Preprocessor* dengan *framework Laravel 5.8*.

Implementasi model sistem dapat digambarkan melalui *Flow chart*. *Flow chart* adalah suatu bagan dengan symbol – symbol tertentu yang menggambarkan urutan proses secara mendetail dan hubungan antar suatu proses dengan proses lainnya dalam suatu program.

3.8.1.Flow Chart Crawling

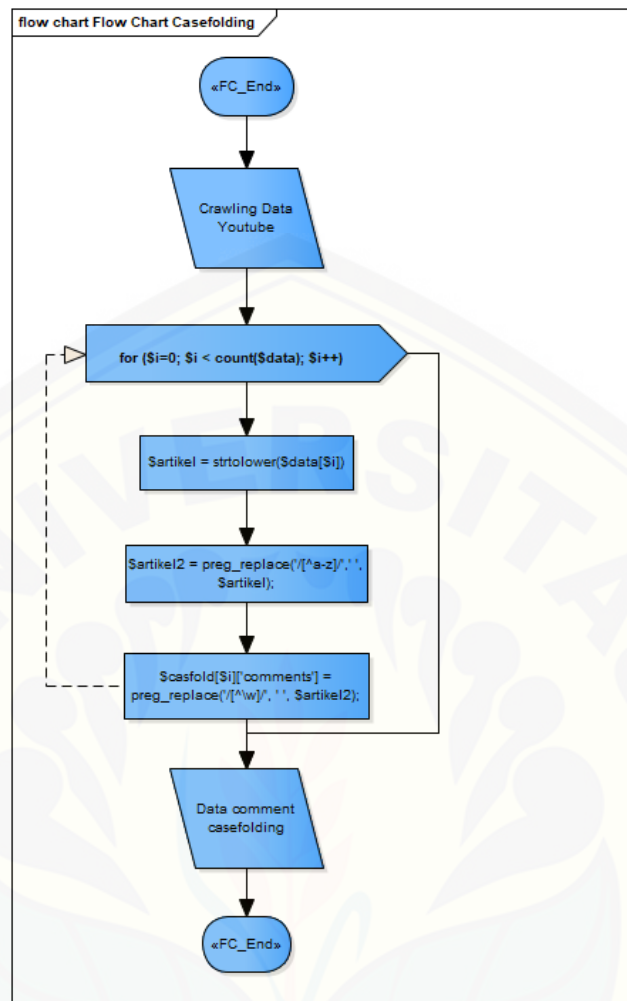
Flow Chart Crawling menggambarkan proses pengambilan atau *mining* data dari *resourcenya* yaitu sosial media *YouTube*. Inti dari proses ini adalah pada pengambilan data komentar yang berjumlah 100 data teratas *YouTube*. *Flow Chart Crawling* dapat dilihat pada Gambar 3.4



Gambar 3.4 Flow Chart Crawling

3.8.2. Flow Chart Casefolding

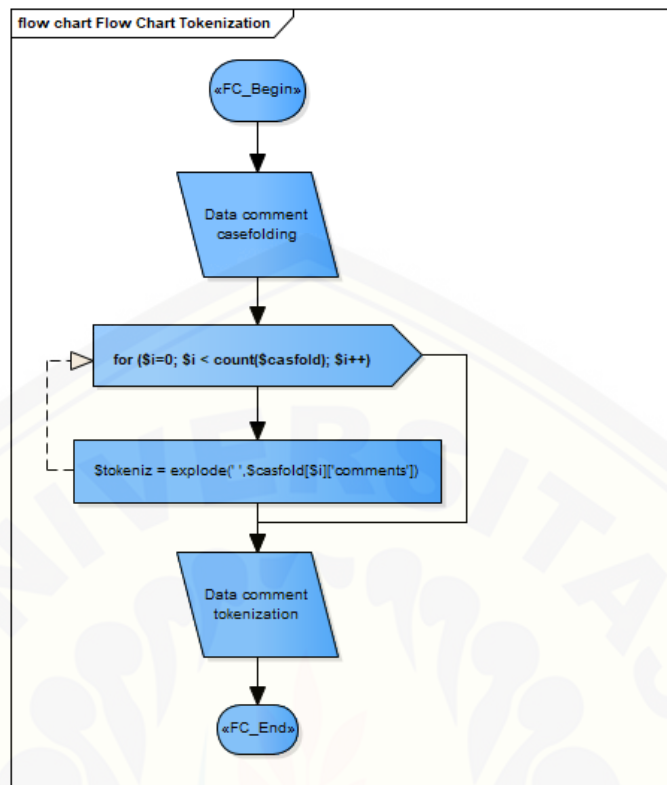
Flow Chart Casefolding menggambarkan proses pertama dalam tahap *preprocessing* yaitu mengubah huruf menjadi *lowercase* semua dan menghapus tanda baca yang tidak bisa diterjemahkan. *Flow Chart Casefolding* dapat dilihat pada Gambar 3.5



Gambar 3.5 Flow Chart Casefolding

3.8.3. Flow Chart Tokenization

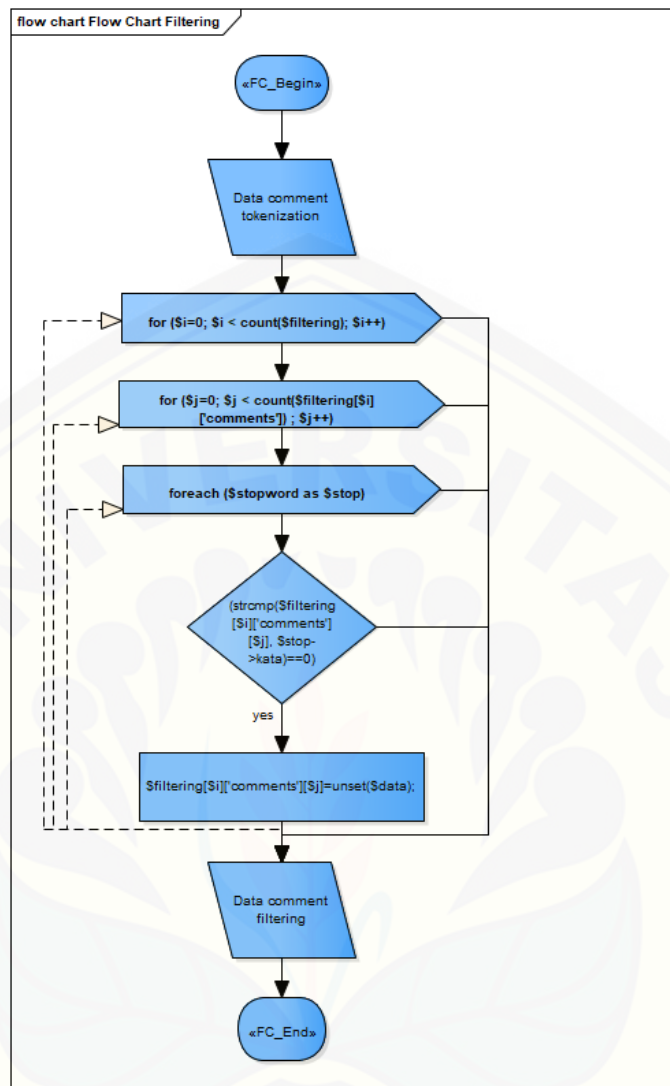
Flow Chart Tokenization menggambarkan proses kedua dalam tahap *preprocessing* yaitu mengubah kalimat menjadi token/bagian kecil untuk selanjutnya digunakan dalam proses *filtering* dan *stemming*. *Flow Chart Tokenization* dapat dilihat pada Gambar 3.6



Gambar 3.6 *Flow Chart Tokenization*

3.8.4. Flow Chart Filtering

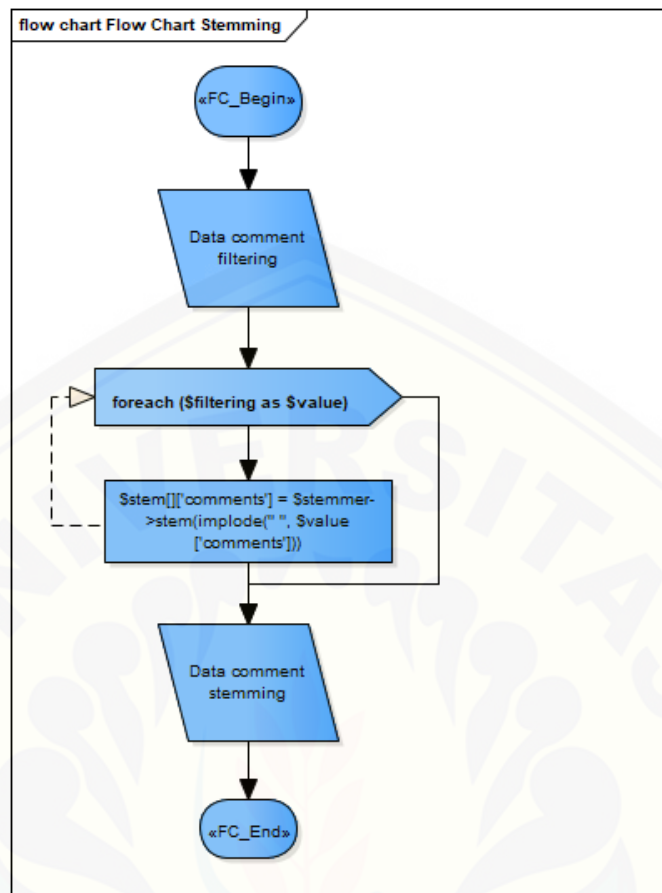
Flow Chart Filtering menggambarkan proses ketiga dalam tahap *preprocessing* yaitu menghilangkan kata-kata yang tidak penting yang dianggap tidak mewakili sentiment positif dan negatif. *Flow Chart Filtering* dapat dilihat pada Gambar 3.7



Gambar 3.7 *Flow Chart Filtering*

3.8.5. Flow Chart Stemming

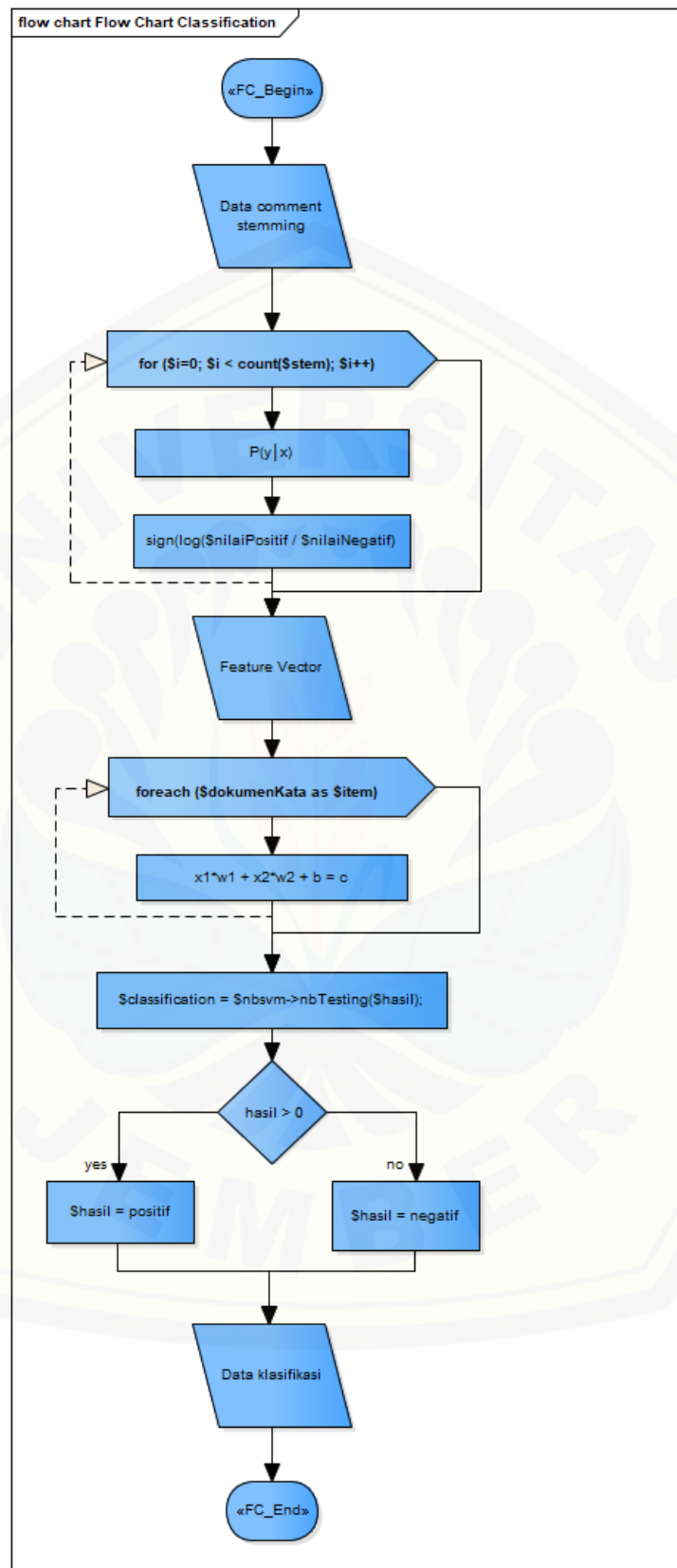
Flow Chart Stemming menggambarkan proses terakhir dalam tahapan *text preprocessing* yaitu mengubah kata kata menjadi kata baku dan menghirangkan imbuhan dan akhiran dalam suatu kata. *Flow Chart Stemming* dapat dilihat dalam Gambar 3.8



Gambar 3.8 *Flow Chart Stemming*

3.8.6. Flow Chart Klasifikasi

Flow Chart Klasifikasi menggambarkan proses penggabungan metode dan pengolahan kata dari tahap *text preprocessing* diolah dalam model training menggunakan metode *Naïve Bayes Support Vector Machine*. *Flow Chart* Klasifikasi dapat dilihat pada Gambar 3.9



Gambar 3.9 Flow Chart Klasifikasi

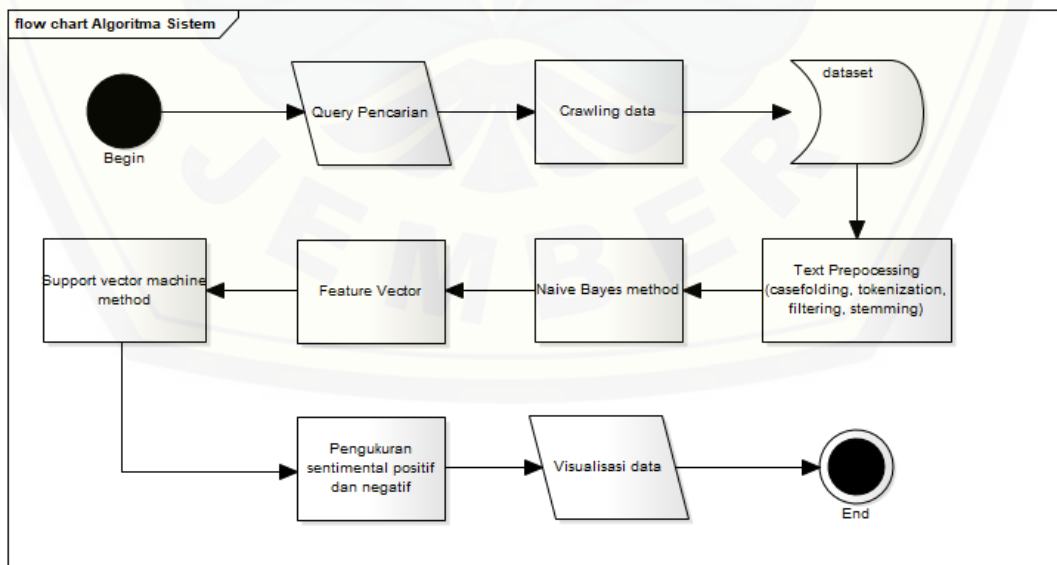
3.9. Pengujian Model dan Pengujian Sistem

Pengujian model adalah tahap dimana peneliti melakukan pengujian terhadap model klasifikasi *Naïve Bayes - Support Vector Machine* yang telah dibangun. Pengujian dilakukan dengan melakukan perubahan pada skala data agar dapat memberikan hasil yang maksimal. Dalam penelitian ini menggunakan 3 skenario skala data yang berbeda. Masing-masing scenario tersebut yaitu: skenario data 60:40, skenario data 70:30 dan skenario data 80:20. Pengujian model selanjutnya akan membandingkan akurasi dari masing-masing model menggunakan *confusion matrix* yang nantinya akan dipilih model dengan akurasi tertinggi.

Pengujian sistem adalah tahap dimana peneliti melakukan pengujian terhadap sistem yang telah dibuat dengan menggunakan pengujian black box. Pengujian black box adalah pengujian dengan cara melakukan pengujian terhadap sistem yang telah berjalan. Pengujian dilakukan dengan mencoba sistem dengan berbagai kemungkinan kesalahan yang ada.

3.10. Gambaran Algoritma Sistem

Sistem analisis sentimen yang dibangun adalah sistem berbasis web yang dapat menganalisis komentar video *YouTube*. Algoritma sistem yang diterapkan dapat dilihat pada Gambar 3.10.



Gambar 3.10 Gambaran Algoritma Sistem

Pada diagram alur algoritma sistem dimulai dengan pengambilan data sampai pada visualisasi data. Penjelasan tiap prosesnya dapat dilihat pada Tabel 3.2.

Tabel 3.2 Penjelasan algoritma sistem

No	Tahap	Input	Proses	Output
1	<i>Query</i> pencarian		pada tahap ini <i>query</i> pencarian dataset dimasukkan untuk menampilkan data video	<i>Query</i>
2	<i>Crawling</i> data	<i>Query</i> YouTube Search API	<i>Crawling</i> data dilakukan dengan menggunakan <i>YouTube Search API</i> dengan menggunakan parameter <i>commentThreads</i>	<i>Dataset</i>
3	<i>Text preprocessing</i>	<i>Dataset</i>	Mengambil komentar video dengan perulangan 100 data komentar	Dataset komentar video
			<i>Casefolding</i> : pengubahan dataset menjadi <i>lowercase</i> dan penghilangan tanda baca titik, koma dan tanda baca lainnya yang tidak perlu	Dataset komentar yang menjadi <i>lowercase</i> dan dihilangkan tanda bacanya
			<i>Tokenization</i> : melakukan pemotongan	Dataset komentar yang telah

			kalimat menjadi kata dan dipisahkan oleh spasi	menjadi kata yang dipisahkan spasi atau <i>bag of words</i>
			<i>Filtering</i> : penghilangan kata yang bukan kata dasar. Apabila kata dalam <i>dataset</i> mengandung kata <i>stopword</i> akan dihilangkan karena tidak termasuk kata sentimen	Dataset komentar yang telah menjadi kata yang mengandung sentimen yang sudah terfilter oleh data <i>stopword</i>
			<i>Stemming</i> : penghilangan imbuhan, baik <i>prefix</i> (kata depan), sufiks (kata akhir) dan konfiks. Dalam stemming dilakukan agar tidak terjadi redundansi data dan kata dapat dimaknai. Proses ini menggunakan library sastrawi	Dataset komentar yang telah menjadi kata baku sesuai data <i>rootword</i>
4	Naïve Bayes Classifier	Dataset yang sudah	Melakukan proses penghitungan nilai probabilitas setiap kata	Data feature vector

		melewati <i>text preprocessing</i>	dari komentar dan pelabelan kelas dengan metode <i>Naïve Bayes</i>	
5	<i>Feature vector</i>	Data feature vector	Melakukan proses penghitungan data dengan cara <i>log count ratio</i>	Data <i>vector</i>
6	Support Vector Machine	Data <i>vector</i>	Melakukan proses pengklasifikasian domain menjadi 2 kelas data menggunakan <i>Support Vector Machine</i>	Data klasifikasi sentimen
7	Pengukuran sentimen	Hasil klasifikasi sentimen	Proses penghitungan indeks kalimat menggunakan perbandingan jumlah sentiment positif dan negatif	Sentimen negatif dan positif
8	Visualisasi data	Sentimen positif dan negatif	Semua data yang telah dilakukan divisualisasikan dalam bentuk statistik	Visualisasi data

BAB 5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan analisis dan pengujian yang dilakukan pada bab sebelumnya, maka kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Metode *Naïve Bayes – Support Vector Machine* untuk klasifikasi data komentar video di *YouTube* telah dilakukan dengan tahapan pembangunan model dan *hybrid* metode *Naïve Bayes* dengan *Support Vector Machine*. Data komentar yang sudah melewati tahapan *text preprocessing* selanjutnya dihitung menggunakan nilai probabilitasnya menggunakan *Multinomial Naïve Bayes* sehingga menghasilkan nilai *feature vector*. Selanjutnya *feature vector* diproses kedalam persamaan metode *Support Vector Machine*. Dimana untuk mencari *hyperplane* terbaik harus mencari jarak antar *support vector* dengan *hyperplane* sehingga ditemukan persamaan nilai W_1 , W_2 dan b dalam penelitian ini adalah $X_1 - 88.869565217391 + X_2 110.125 + 1.0450852922821 = 0$.
2. Penerapan metode *Naïve Bayes - Support Vector Machine* menjadi sangat optimal ketika data training yang digunakan memiliki jumlah data yang banyak dan data yang bervariasi. *Naïve Bayes* sangat baik dalam mengklasifikasikan teks dengan jumlah data yang kecil atau cuplikan dokumen dan *snippets* sedangkan *Support Vector Machine* sangat baik dalam mengklasifikasikan teks dengan jumlah data yang relative banyak atau dokumen yang panjang *full-length*. Kombinasi dari metode *Naïve Bayes* dan *Support Vector Machine* menghasilkan tingkat akurasi yang lebih baik dan performa yang lebih kuat dengan penggunaan skala data 7:3 yaitu 70% data *training* dan 30% data *testing*. Dengan menghasilkan nilai uji performansi terbesar yaitu *precision* sebesar 91%, *recall* sebesar 83% dan *f1score* sebesar 87%.

5.2. Saran

Penulis menyarankan pengembangan penelitian lebih lanjut sentimen analisis positif dan negatif komentar video *YouTube* menggunakan metode *Naïve Bayes - Support Vector Machine* sebagai berikut:

1. Untuk mendapat nilai akurasi yang lebih tinggi lebih baik, diharapkan menambah proses dalam pengolahan data. Penggunaan *Ngram* menjadikan hasil pengolahan data menjadi lebih akurat.
2. Penggunaan variasi data dengan jumlah besar menjadikan proses training menjadi lebih beragam, nilai yang dihasilkan untuk mencari persamaan *hyperplane* terbaik akan semakin akurat jika data training bervariasi.
3. Penelitian selanjutnya dapat menggunakan *Facebook, Twitter, Instagram* dan sosial media lainnya untuk menjangkau data lebih banyak.

DAFTAR PUSTAKA

- Feldman, R. (2007). *Advanced Approaches in Analyzing Unstructured Data*. United States of America: Cambridge University Press.
- Gordon, K. (2018). *YouTube: Statistics & Data*. Retrieved from statista: <https://www.statista.com/topics/2019/youtube/>
- J. Clement. (2018, February 13). *Statista*. Retrieved from [www.statista.com: https://www.statista.com/statistics/805656/number-youtube-viewers-worldwide/](https://www.statista.com/statistics/805656/number-youtube-viewers-worldwide/)
- JAIN, Y. H. (1998). Classification of Text Documents. *THE COMPUTER JOURNAL*, Vol. 41, No. 8.
- Pang, B. (2002). Thumbs up? Sentiment Classification using Machine Learning. *Association for Computational Linguistics*, 79-86.
- Pang, B. L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 1-135.
- Ponilan, I. R. (2016). Pengukuran Happiness Index Masyarakat Kota Bandung pada Media Sosial. *Ind. Symposium on Computing*, 17-22.
- Sida Wang, d. C. (2012). Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. *Department of Computer Science Stanford University*.
- socialblade. (2019, July 1). *Socialblade*. Retrieved from [socialblade.com: https://socialblade.com/youtube/channel/UCu0yQD7NFMylu_-TmKa4Hqg](https://socialblade.com/youtube/channel/UCu0yQD7NFMylu_-TmKa4Hqg)
- Steven Bird, E. K. (2009). *Natural Language Processing in Python*. United states of america: O'reilly media.
- Sugiyono. (2008). *Metode penelitian pendidikan: (pendekatan kuantitatif, kualitatif dan R & D)*. Bandung: Alfabeta.

- Tala, F. Z. (2003). A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. . *M.Sc. Thesis. Master of Logic Project. Institute for Logic, Language and Computation.* .
- Tanesab, F. I. (2017). Sentiment Analysis Model Based On Youtube Comment Using Support Vector Machine. *Fakultas Teknologi Informasi.*
- Wahid, D. H. (2016). Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 10(2), 207-218.
- Wibisono, Y. (2005). Klasifikasi berita bahasa indonesia menggunakan Naive Bayes Classifier.

LAMPIRAN

Lampiran 1 Surat keterangan validasi data

SURAT KETERANGAN

Yang bertanda tangan di bawah ini, menerangkan bahwa:

Nama : Abbi Nizar Muhammad
Nomor Induk Mahasiswa : 152410101150
Program Studi : Sistem Informasi
Fakultas : Ilmu Komputer
Instansi : Universitas Jember
Judul Skripsi : Analisis Sentimen Positif Dan Negatif Komentar Video
*Youtube Menggunakan Metode Naïve Bayes – Support Vector
Machine (NBSVM) Classifier*

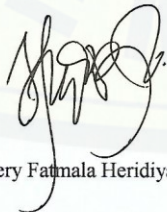
telah melaksanakan tahap validasi data dalam penyusunan skripsi untuk keperluan penggunaan kata Bahasa Indonesia yang baik dan benar pada sistem yang akan dibangun.

Profil validator data sebagai ahli bahasa :

Nama : Mery Fatmala Heridiyanto
Nomor Induk Mahasiswa : 150210402026
Instansi : Pendidikan Bahasa Indonesia Fakultas Keguruan dan Ilmu
Pendidikan Universitas Jember

Jember, 30 Juni 2019

Ahli bahasa,



Mery Fatmala Heridiyanto

Peneliti,



Abbi Nizar Muhammad

Lampiran 2 Validasi data oleh ahli bahasa

No	Kata	Kelas
1	videonya bagus	positif
2	konten yang menarik dan bisa dibagikan	positif
3	semakin rajin bikin animasi, kami penonton setia selalu sabar menunggu	positif
4	keren inget banget dulu pertama kali subscribe video masih sederhana banget	positif
5	channel edukasi terhebat di indonesia semoga kedepan bisa banyak channel bermanfaat	positif
6	datanya akurat dan baik untuk konten pendidikan	positif
7	saya suka video ini	positif
8	Itu video bagus, video keren	positif
9	konten bagus, itu video keren	positif
10	bagus banget bagaimana membuat kartun seperti itu	positif
11	ini channel paling mantap	positif
12	pembahasan yang menarik	positif
13	sangat bermanfaat dan menambah wawasan	positif
14	info keren	positif
15	konten edukasi dengan animasi walaupun masih sangat pemula, silahkan mampir bila berkenan terima saran dan kritikan	positif
16	mantap mencerdaskan bangsa	positif
17	ini channel paling cocok	positif
18	setuju untuk dukung konten pendidikan	positif
19	bagus sekali mimpi untuk membangun Indonesia	positif
20	semoga bisa berkolaborasi	positif
21	saya akan membuat channel edukasi tahun depan	positif
22	ayo konten edukasi perbanyak lagi	positif
23	sangat mendidik	positif

24	ambil positif aja	positif
25	setuju banget sama gagasannya	positif
26	semangat terus	positif
27	untung ketemu channel ini	positif
28	sangat mengapresiasi channel ini	positif
29	setuju sekali	positif
30	mantap konten bermanfaat	positif
31	mantap banget saya dukung	positif
32	hidup channel edukasi indonesia	positif
33	ingin berkolaborasi	positif
34	bener banget	positif
35	jenius gambar indonesia	positif
36	pendidikan indonesia bagus	positif
37	video cerdas ulas sampai akhir	positif
38	beri inspirasi bagi penikmat video ini	positif
39	setuju sama video ini	positif
40	keren kok bisa	positif
41	semakin memotivasi channel video	positif
42	semoga maju terus pendidikan indonesia	positif
43	pertama suka	positif
44	keren nonton youtube sambil belajar juga	positif
45	meningkatkan pengetahuan	positif
46	dukung channel konten edukasi indonesia	positif
47	gokil banget	positif
48	kualitas video yang baik	positif
49	terima kasih sudah mengedukasi masyarakat	positif
50	bermanfaat bagi generasi muda indonesia	positif
51	menginspirasi pemuda untuk belajar terus	positif
52	penyampaian materi yang sangat enak dipelajari	positif
53	serius pengen berkolaborasi	positif

54	kalo begini pendidikan indonesia semakin maju	positif
55	ayo kita bagikan video ini ke keluarga kita	positif
56	pemateri ramah	positif
57	berkesan setelah menonton video ini	positif
58	memberi dampak positif bagi yang melihat	positif
59	membantu anak anak belajar	positif
60	mengurangi beban pikiran	positif
61	setelah melihat video ini saya semakin suka	positif
62	berguna bagi pelajar	positif
63	semakin tahu tentang pengetahuan	positif
64	setuju dengan isi video	positif
65	bisa diterima dan masuk dalam otak	positif
66	bagikan dan suka	positif
67	meningkatkan minat baca orang orang itu keren	positif
68	menarik untuk diceritakan ke teman	positif
69	puncak karir bagus	positif
70	rasa hormat buat pemateri dan bikin konten	positif
71	riang gembira melihatnya	positif
72	sabar pasti jadi trending	positif
73	selamat sudah mencapai trending nomor satu	positif
74	sukses terus kok bisa	positif
75	tegas dalam penyampaian dan teguh dalam pendapat	positif
76	pemateri sopan menyampaikan isi video	positif
77	visioner pendidikan indonesia akan baik	positif
78	wow keren sekali	positif
79	harus diapresiasi setinggi tingginya	positif
80	bagus terus tingkatkan	positif
81	gila keren banget	positif
82	terus berkarya sampai pendidikan indonesia baik	positif
83	pasti banyak manfaat	positif

84	semakin banyak yang suka	positif
85	channel besar memberi manfaat	positif
86	bagus tingkatkan	positif
87	bijaksana dalam berfikir	positif
88	terus berjuang mencerdaskan rakyat	positif
89	pemuda butuh konten edukasi	positif
90	bakat terpendam	positif
91	semua orang juga bisa	positif
92	serius ini keren	positif
93	tunggu karya berikutnya	positif
94	wah bagus	positif
95	patut ditiru	positif
96	data yang tampil akurat dan bisa percaya	positif
97	visual data bagus	positif
98	grafik data bisa dibaca	positif
99	deskripsi jelas dan bisa temukan	positif
100	ahli sudah berbicara	positif
101	beri kesan dan makna bagus	positif
102	kenapa pendidikan indonesia penting	positif
103	pikiran yang brilian	positif
104	terima kasih sudah memberi inspirasi	positif
105	senang sekali melihat anak muda aktif berkarya	positif
106	anak muda sadar pendidikan	positif
107	edukasi mengajarkan kita untuk terus belajar dan eksplor kemampuan diri	positif
108	semakin sadar bahwa pendidikan penting bagi kehidupan	positif
109	kadang sering fikir kenapa orang harus berbuat baik	positif
110	harus tingkatkan ilmu untuk mencapai cita cita	positif
111	pendidikan formal itu penting pendidikan non formal bisa tambah ilmu	positif

112	semakin banyak orang sadar melihat video edukasi	positif
113	bagus buat referensi	positif
114	keren tingkatkan lagi	positif
115	cantik pandang	positif
116	elegan untuk dilihat	positif
117	enak di dengar saat lagi belajar	positif
118	channel favorit yang selalu memberi wawasan tentang ilmu alam	positif
119	gratis bisa belajar kapan saja dan dimana saja	positif
120	hasil karya terbaik	positif
121	imajinasi yang indah bisa dilihat dari cara mengolah konten	positif
122	jujur video ini bagus	positif
123	kagum dengan isi konten dan proses produksi video	positif
124	totalitas tanpa batas	positif
125	serasa belajar dan bermain	positif
126	kerja belajar bermain	positif
127	setelah belajar disini mendapat juara kelas	positif
128	bahagia bisa melihat anak muda aktif berkarya	positif
129	konsep kurikulum baik akan jadi landasan siswa untuk berpikir kreatif	positif
130	kurikulum pendidikan sekolah sudah baik	positif
131	pendidikan dinamis bisa berubah karena teknologi terus maju	positif
132	tips untuk ujian akhir	positif
133	kemas secara menarik	positif
134	sangat niat untuk membuat konten kualitas	positif
135	video bagus membuat saya untuk masuk sekolah	positif
136	belajar disini membuat pintar	positif
137	ada yang sekali belajar dia ngerti itu namanya passion	positif

138	ada yang usaha belajar berkali berkali baru mengerti dan paham itu namanya cari skill	positif
139	fasilitas sekolah sudah bagus	positif
140	naik kelas karena sering nonton video ini	positif
141	tanamkan pendidikan masyarakat	positif
142	sekolah jadi pendidikan moral dan karakter	positif
143	akademik dilakukan secara lebih efisien dan menarik seperti belajar interaktif	positif
144	guru memberi projek dimana guru juga memberi arah terkait kompetensi dasar	positif
145	menjadi kompetensi jurusan yang mencakup lebih banyak jurusan daripada sekolah sekarang dan lebih mengarah ke gaya kuliah	positif
146	suruh anak magang	positif
147	lihat tempat kerja atau ikutan dapat membuat peserta didik familiar	positif
148	tingkat kemampuan karakter dan akademik	positif
149	baca buku	positif
150	siswa belajar sendiri dari buku dan internet	positif
151	materi akademis adalah yang terpenting meski itu bukan ahli dan potensi siswa	positif
152	tempat les adalah tempat untuk mencari ilmu	positif
153	pemikiran kritis harus ditingkatkan	positif
154	studi lapangan projek dan bermain	positif
155	jujur dalam menjawab soal	positif
156	daya saing masuk sekolah jadi ketat	positif
157	kenapa video ini menarik	positif
158	konten yang menarik dan bagus	positif
159	channel kok bisa edukasi	positif
160	aplikasi ruang guru bantu siswa paham pelajaran	positif

161	teknologi rubah cara pandang manusia untuk belajar	positif
162	belajar dimana saja pasti bisa	positif
163	video interaktif itu menarik	positif
164	karena sudah biasa belajar lewat musik	positif
165	lebih nyaman lihat video daripada lihat buku	positif
166	baca buku setiap hari agar bisa	positif
167	buku pelajaran di perpustakaan sekolah banyak yang bagus	positif
168	baca berita setiap pagi	positif
169	kreatif dalam mengolah kata dan kalimat	positif
170	bagus untuk jadi rujukan materi	positif
171	belajar jadi lebih asik	positif
172	suka dengan gaya bahasa	positif
173	konten menarik	positif
174	visualisasi gambar bagus dan interaktif	positif
175	mudah di cerna otak	positif
176	pendidikan harus difungsikan sebagaimana mestinya	positif
177	wow keren banget	positif
178	mantap abis	positif
179	iya saya setuju	positif
180	suka dengan bahasanya	positif
181	selera orang indonesia yang kontennya kreatif	positif
182	channel yang banyak memberi manfaat	positif
183	saya dukung seratus persen	positif
184	karya anak bangsa	positif
185	patut diapresiasi	positif
186	lebih keren dari saingannya	positif
187	bagikan ke teman	positif
188	edukasi adalah segalanya	positif
189	generasi muda harus sadar politik	positif
190	pemuda bawa perubahan	positif

191	sukses di usia muda	positif
192	sukses bawa perubahan	positif
193	simpan data untuk belajar	positif
194	data bisa dipertanggungjawabkan	positif
195	kontribusi pada kebaruan	positif
196	kontribusi ilmu	positif
197	bersih banget	positif
198	dampak positif	positif
199	bertumbuh pendidikan indonesia bagus	positif
200	mengubah mindset orang	positif
201	mengubah perilaku siswa menjadi baik	positif
202	melanjutkan membuat konten video	positif
203	harus mampu menahan emosi	positif
204	tahap pembuatan video susah	positif
205	harus fokus dan tetap berdoa	positif
206	video berbeda dengan konten lain	positif
207	sebuah kemajuan yang pesat	positif
208	patut diberi penghargaan	positif
209	selalu mendukung demi edukasi indonesia	positif
210	pasti selalu jadi nomor satu	positif
211	harus dibagikan ke teman biar tau	positif
212	kenapa suka ini konten	positif
213	harus kerja keras	positif
214	pasti dapat yang baik	positif
215	pasti memberi manfaat bagi lingkungan	positif
216	lingkungan mendukung tumbuh kembang anak	positif
217	otak harus beri bernutrisi	positif
218	agar pikiran jernih	positif
219	juara satu setelah nonton video ini	positif
220	nilai bagus karena rajin lihat konten disini	positif

221	saya suka	positif
222	tanda tanda kiamat	positif
223	indonesia kapan maju	positif
224	sekarang sudah banyak tapi isinya banyak melenceng	positif
225	kenapa orang lebih sering menonton televisi dibanding nonton youtube	positif
226	channel edukasi masih menerima pesan iklan dalam videonya	positif
227	channel edukasi apa channel promosi	positif
228	mayoritas masyarakat masih senang dengan tayangan hiburan	positif
229	pertanyaan paling membingungkan	positif
230	permasalahan membuat kita marah	positif
231	pasti admin kehabisan ide	positif
232	jaman sekarang channel lebay	positif
233	yang penonton banyak biasanya gaming	positif
234	waspada semuanya	positif
235	jangan disebar	positif
236	ironis lihat video konten edukasi	positif
237	sarkasme terhadap pemerintah	positif
238	tipu daya setan	positif
239	cerewet sekali	positif
240	ancaman bagi konten game	positif
241	tolong kepada pembuat konten perhatikan jam upload	positif
242	tolong hilangkan deskripsi panjang	positif
243	tingkatkan kualitas video	positif
244	tolong tambahkan arti	positif
245	siswa wajib datang tepat waktu	positif
246	nilai adalah segalanya	positif
247	sekolah adalah tempat mencari nilai	positif
248	anak indonesia punya penyakit bisa lupa waktu dan belajar	positif

249	siswa dituntut biaya mahal	positif
250	pelajar demo	positif
251	sekolah tempat bermain	positif
252	tidak sia sia bikin video edukasi lanjutkan	negatif
253	tidak ada masalah	negatif
254	setiap anak juga berbeda nalar tapi bukan bodoh	negatif
255	sekolah tidak mengedepankan eksistensi dan harga tapi kembali ke fungsi awal yaitu pendidikan untuk masyarakat	negatif
256	tidak peduli nilai tapi usaha untuk dapat nilai	negatif
257	saya benci video ini	negatif
258	konten buruk	negatif
259	mencuri konten video	negatif
260	tidak mungkin	negatif
261	youtube jadi kurang manfaat	negatif
262	tidak suka video ini	negatif
263	makanya orang Indonesia goblok goblok	negatif
264	tidak populer	negatif
265	kok alamat jelek baget pencet malah situs porno	negatif
266	ambil negatif aja	negatif
267	salah tinggal di indonesia	negatif
268	sudah lama muak sama konten youtube indonesia tidak berkualitas	negatif
269	karena di indonesia banyak orang gabut	negatif
270	bukan channel edukasi	negatif
271	jangan terlalu dominan kontennya	negatif
272	youtube indonesia gak seru	negatif
273	berawal dari tontonan televisi yang belum sepenuhnya memiliki tayangan edukasi	negatif
274	konten paling tidak penting	negatif
275	mungkin channel belum ada yang menarik	negatif

276	bocah kurang edukasi	negatif
277	jelek banget videonya	negatif
278	bodoh banget orang	negatif
279	tidak setuju dengan channel ini	negatif
280	absurd banget tidak ada kejelasan video	negatif
281	biasa saja tidak menarik	negatif
282	beli kelas online jadi boros	negatif
283	presiden diktator	negatif
284	rugi banget lihat konten video ini	negatif
285	sayang tidak begitu mengerti	negatif
286	jarang ada channel edukasi	negatif
287	tidak peduli sama konten edukasi	negatif
288	semua membosankan dan buruk	negatif
289	komentar sok pinter	negatif
290	tidak pantas ditonton	negatif
291	kecewa lihat pendidikan indonesia	negatif
292	terpaksa sekali orang kalo mebuat konten	negatif
293	konten sampah	negatif
294	konten sampah	negatif
295	kurang menarik	negatif
296	penjelasan yang ribet	negatif
297	tidak memberi makna	negatif
298	memancing provokasi	negatif
299	tingkatkan kualitas suara	negatif
300	suara dan grafik buruk	negatif
301	tolong kurangi suara yang tidak penting	negatif
302	lebih baik buat konten porno daripada konten seperti ini	negatif
303	tidak berguna dilihat	negatif
304	tidak setuju dengan isi video	negatif
305	konten menyesatkan	negatif

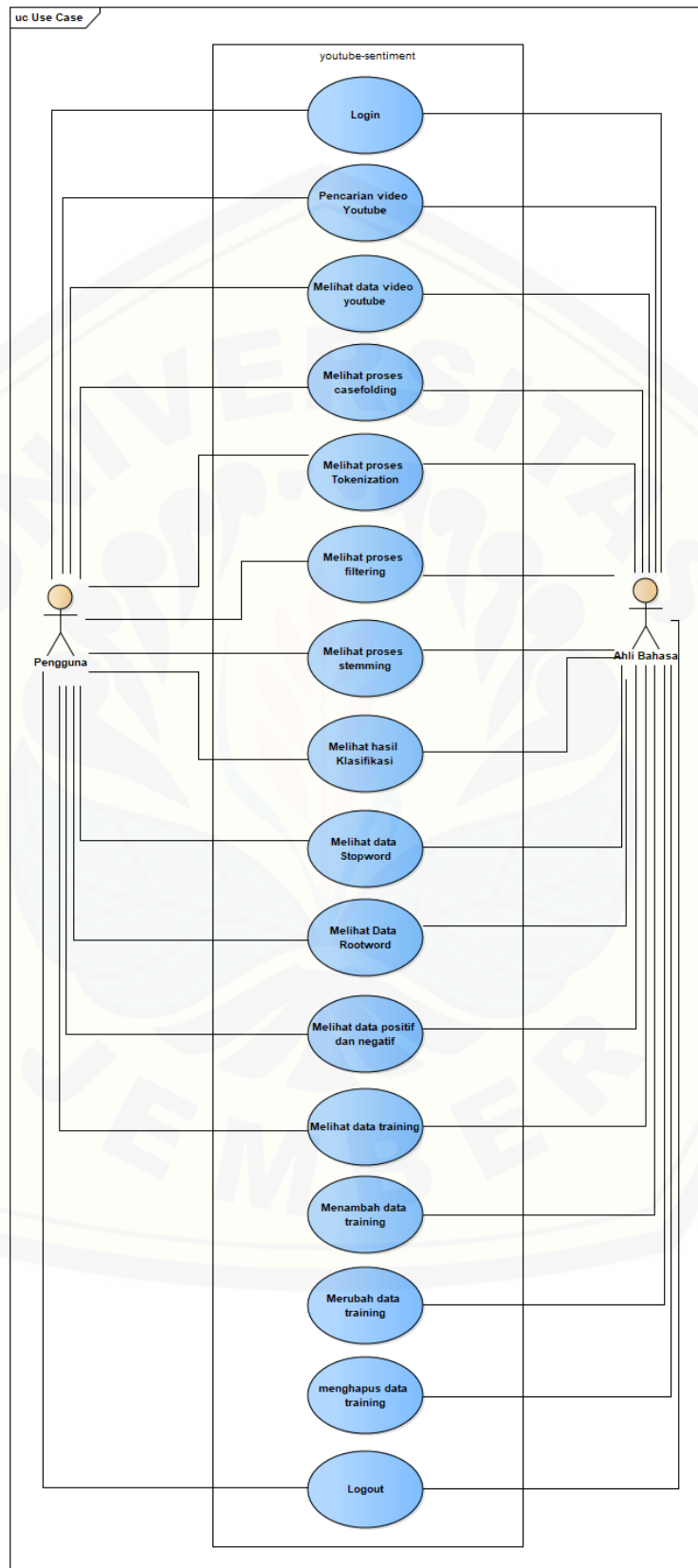
306	tidak bisa dipahami	negatif
307	sulit dimengerti	negatif
308	orang orang malas membaca berita jadi kurang tertarik sama berita	negatif
309	bodo amat sama berita macam gini	negatif
310	mending lihat konten hiburan daripada konten edukasi	negatif
311	rusak semua	negatif
312	materi salah dengan praktek	negatif
313	sedih melihat youtube indonesia	negatif
314	pemikiran yang sempit untuk jadi konten	negatif
315	sulit bikin video	negatif
316	takut menyebarkan video	negatif
317	teror bagi pendidikan	negatif
318	pemateri tidak sopan menyampaikan isi video	negatif
319	cacat jelek	negatif
320	malas lihat konten yang kurang menarik	negatif
321	hanya bisa bacot	negatif
322	bodoh itu konten yang jelek tidak bermakna	negatif
323	kecewa sama anda	negatif
324	orang orang tolol	negatif
325	basi tidak ada arti	negatif
326	anjing apa maksudnya	negatif
327	durasi kurang panjang	negatif
328	kalimat yang diucapkan kotor	negatif
329	tidak bisa diterima karena banyak kata yang tidak ada artinya	negatif
330	kurang berkesan baik	negatif
331	kenapa tidak sewa orang untuk bicara	negatif
332	kadang suka tidak setuju yang dikatakan oleh video	negatif
333	kurang mampu mengolah kata	negatif

334	bangsat kurang mendidik	negatif
335	kurang jelas suaranya	negatif
336	menjadi pemuda semakin malas membaca	negatif
337	tidak ada guru jangan percaya hoax	negatif
338	sekolah kurang fasilitas	negatif
339	jalan jauh sekolah susah	negatif
340	tergantung orang tua benar mau anak cerdas atau sekedar nerima ijazah sebagai tanda sudah selesai	negatif
341	kenapa edukasi kurang rata	negatif
342	konsep belajar kuno	negatif
343	guru malas buat materi di kelas	negatif
344	video profokasi	negatif
345	kadang saya lihat semua konten edukasi jelek	negatif
346	visual kurang menarik dan suara kurang jelas	negatif
347	sekolah jangan buat sistem sekolah favorit tapi jadi sekolah lokal	negatif
348	kualitas guru sangat rendah	negatif
349	murid mencontek satu sama lain	negatif
350	siswa tidak dapat ilmu apa apa	negatif
351	kemampuan non akademis sangat diabaikan	negatif
352	ekstrakurikuler sekolah hanya jadi kelompok tidak dianggap sebagai potensi	negatif
353	penghargaan dan dukungan sekolah untuk murid ikut lomba sangat minim	negatif
354	anak kurang mampu untuk ikut les kurang mendapat ilmu yang harus didapat	negatif
355	masalah fasilitas sangat miris	negatif
356	bobot belajar yang tidak sesuai	negatif
357	pemerintah salah arah dalam membentuk sistem pendidikan	negatif

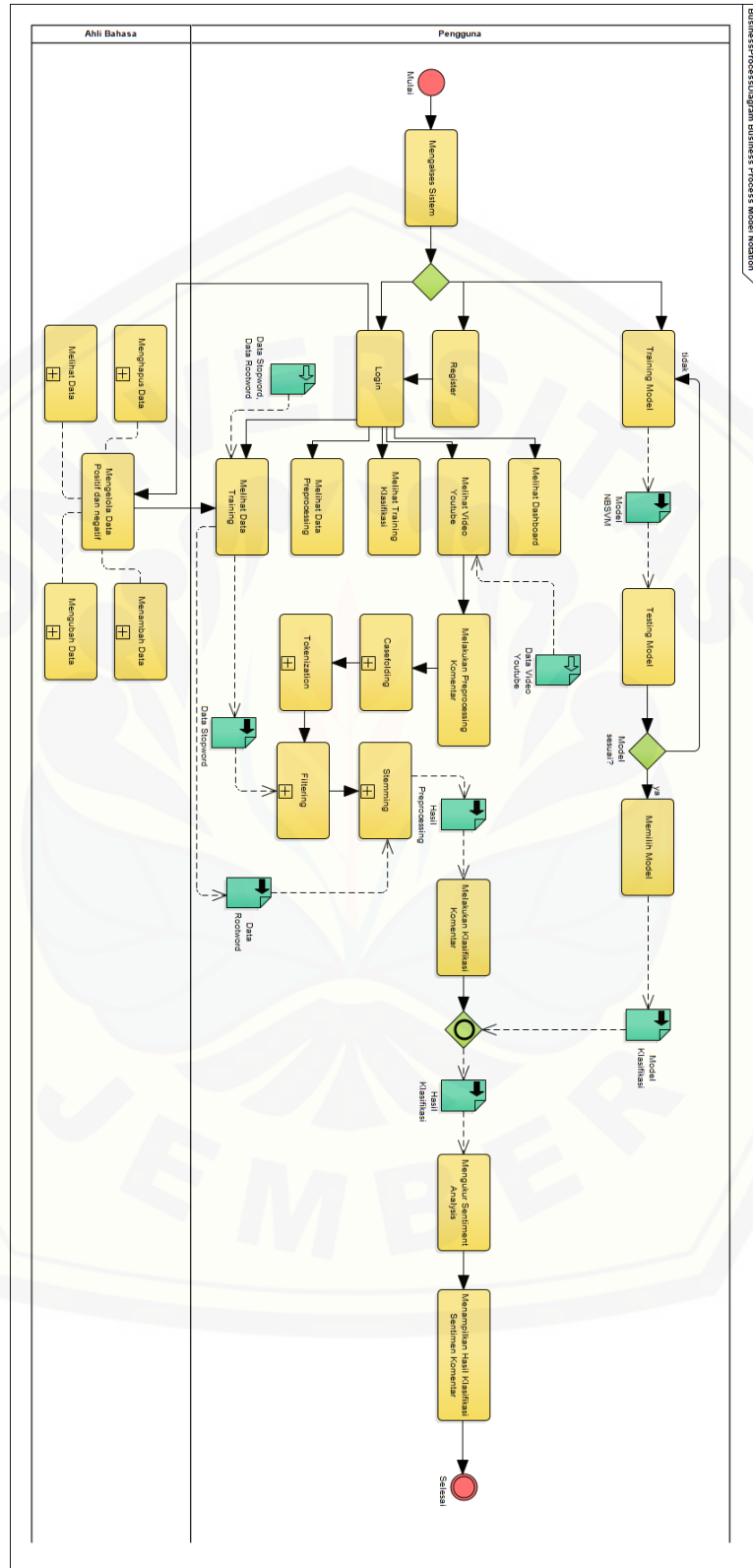
358	tidak efektif sama sekali dalam mengembangkan minat dan bakat siswa	negatif
359	salah jurusan	negatif
360	dalam artian banyak nilai dikontrol dan mencontek masal	negatif
361	kualitas guru dan fasilitas belajar sangat minim	negatif
362	guru tidak belajar dan tidak memberi ilmu sama sekali	negatif
363	waktu belajar sendiri dari buku dan medium	negatif
364	murid yang takut salah	negatif
365	ujian kualitas rendah berbasis pilihan ganda	negatif
366	budaya mencontek dan nilai itu penting	negatif
367	sistem pendidikan tidak efisien	negatif
368	sistem belajar murid harus bisa menguasai mata pelajaran padahal itu salah	negatif
369	video kurang informatif	negatif
370	kurang mantap	negatif
371	materi kurang berfaedah	negatif
372	kadang suka iri dengan konten edukasi	negatif
373	buku lebih bagus dari materi di video	negatif
374	tidak setuju dengan opini pemateri	negatif
375	tidak menarik untuk dilihat	negatif
376	tidak suka dengan cara ngomongnya	negatif
377	bagus kenapa tidak jadi trending	negatif
378	konten memecah persatuan dan kesatuan	negatif
379	menyebarkan hoax	negatif
380	membohongi publik itu tidak pantas	negatif
381	kurang bisa menghargai orang	negatif
382	data bohong	negatif
383	buruk banget	negatif
384	tidak ada akal	negatif
385	tidak ada fungsi	negatif

386	tidak ada kontribusi	negatif
387	fasilitas minim	negatif
388	tidak ada fasilitas yang memadai	negatif
389	kotor banget	negatif
390	dampak negatif	negatif
391	kurang kerja	negatif
392	tidak mengakomodasi pelajar dalam belajar	negatif
393	pelajar mulai malas dengan drama pemerintah	negatif
394	tidak mencontohkan yang baik	negatif
395	kurang bisa menahan emosi	negatif
396	menguubah menjadi buruk	negatif
397	tidak mudah	negatif
398	tidak elegan	negatif
399	buruk dan tidak berfungsi	negatif
400	tidak ada yang bisa diambil	negatif

Lampiran 3 Use Case Diagram



Lampiran 4 Business Process Model Notation



Lampiran 5 Class Diagram

