



**KLASIFIKASI KELUHAN MAHASISWA DI MEDIA SOSIAL
WHATSAPP MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT
VECTOR MACHINE**

TUGAS AKHIR

AMELIA VERGI KUSUMA
41517120101

UNIVERSITAS
MERCU BUANA
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021**



**KLASIFIKASI KELUHAN MAHASISWA DI MEDIA SOSIAL
WHATSAPP MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT
VECTOR MACHINE**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

UNIVERSITAS
MERCU BUANA
AMELIA VERGI KUSUMA
41517120101

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

LEMBAR PERNYATAAN ORISINALITAS

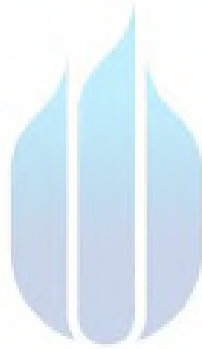
Yang bertanda tangan dibawah ini:

NIM : 41517120101

Nama : Amelia Vergi Kusuma

Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Amelia Vergi Kusuma
NIM : 41517120101
Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

MERCU BUANA

Jakarta, 23 Februari 2022



SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Amelia Vergi Kusuma
NIM : 41517120101
Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial
WhatsApp Menggunakan Metode *Naïve Bayes* dan
Support Vector Machine

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	✓
		Jurnal Nasional Terakreditasi	✓		
		Jurnal International Tidak Bereputasi		Diterima	
		Jurnal International Bereputasi			
Disubmit/dipublikasikan di :	Nama Jurnal	: Khazanah Informatika			
	ISSN	: E-ISSN 2477-698X			
	Link Jurnal	: https://journals.ums.ac.id/index.php/khif/index			
	Link File Jurnal Jika Sudah di Publish	:			

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 23 Februari 2022


3674DAKX015035258
Amelia Vergi Kusuma

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517120101
Nama : Amelia Vergi Kusuma
Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial
WhatsApp Menggunakan Metode *Naïve Bayes* dan
SVM Sehubungan Dengan Mutu Perguruan Tinggi
(Studi Kasus: Universitas Mercu Buana)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 23 Februari 2022

Dosen Penguji



(Achmad Kodar, Drs. MT)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517120101
Nama : Amelia Vergi Kusuma
Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial
WhatsApp Menggunakan Metode *Naïve Bayes* dan
SVM Sehubungan Dengan Mutu Perguruan Tinggi
(Studi Kasus: Universitas Mercu Buana)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 23 Februari 2022



(Anis Cherid, SE, MTI)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517120101
Nama : Amelia Vergi Kusuma
Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode *Naïve Bayes* dan SVM Sehubungan Dengan Mutu Perguruan Tinggi (Studi Kasus: Universitas Mercu Buana)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 23 Februari 2022


(Wawan Gunawan, S.Kom, MT)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41517120101
Nama : Amelia Vergi Kusuma
Judul Tugas Akhir : Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Agustus 2022

Menyetujui,

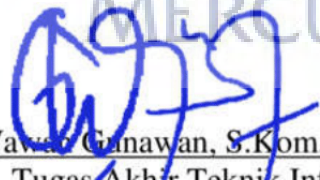


(Dr. Ir. Eliyani)
Dosen Pembimbing

Mengetahui,

UNIVERSITAS

MERCU BUANA



(Wawan Gunawan, S.Kom, MT)
Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan Allah SWT atas rahmat dan karunia-Nya sehingga Tugas Akhir yang berjudul “Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*” dapat diselesaikan dalam jangka waktu yang sudah di tentukan. Laporan tugas akhir ini dibuat sebagai syarat untuk LULUS sebagai sarjana Ilmu Komputer dari Universitas Mercu Buana.

Penulis menyadari bahwa Laporan Tugas Akhir ini masih jauh dari kata sempurna dikarenakan keterbatasan kemampuan dan pengetahuan yang penulis dapatkan. Dalam menyelesaikan Laporan Tugas Akhir ini, penulis mendapat banyak bantuan, bimbingan, dan dukungan Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Kedua orang tua, Ayah, Ibu, Kakak dan Adik, tugas akhir ini saya dedikasikan khusus untuk mereka yang tidak pernah lelah mendoakan, mendukung dan memberi semangat agar saya bisa menyelesaikan kuliah dengan tepat waktu untuk proses meraih gelar sarjana bagi saya.
2. Ibu Dr. Ir. Eliyani selaku Dosen Pembimbing Tugas Akhir yang telah memberikan masukan dan arahan saat serta meluangkan waktunya untuk melakukan bimbingan dalam penyusunan tugas akhir ini hingga selesai.
3. Bapak Emil R. Kaburuan, Ph.D. selaku Kepala Program Studi Teknik Informatika Universitas Mercu Buana.
4. Bapak Wawan Gunawan, S.Kom, MT selaku Koordinator Tugas Akhir Teknik Informatika Universitas Mercu Buana.
5. Dr. Mujiono Sadikin, M.T. selaku Dosen pembimbing akademik yang telah membimbing penulis sejak awal semester yang selalu memberi dukungan dan motivasi agar lulus tepat waktu.
6. Seluruh Dosen Program Studi Teknik Informatika yang sudah memberikan ilmu yang bermanfaat selama kuliah berlangsung. Memberi kesempatan untuk belajar, berkarya, dan juga berkembang.
7. Sahabat, teman - teman kontrakan dan teman - teman kuliah saya tercinta yang senantiasa memberi bantuan dan motivasi.

Akhir kata, penulis berharap segala kekurangan penulisan, eksperimen, cara penjelasan. Untuk itu, kritik dan saran pembaca sangat dihargai dan diharapkan semoga Tugas Akhir ini dapat memberikan manfaat bagi para pembaca

Jakarta, 21 Desember 2021
Penulis



DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... ..	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN PENGUJI	v
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT.....	x
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
NASKAH JURNAL	1
KERTAS KERJA.....	18
BAB 1. LITERATUR REVIEW.....	19
BAB 2. ANALISIS DAN PERANCANGAN	23
BAB 3. SOURCE CODE	28
BAB 4. DATASET.....	30
BAB 5. TAHAPAN EKSPERIMEN.....	31
BAB 6. HASIL SEMUA EKSPERIMEN.....	33
DAFTAR PUSTAKA	58
LAMPIRAN DOKUMEN HAKI.....	61
LAMPIRAN KORESPONDENSI	63

NASKAH JURNAL



Jurnal Ilmu Komputer dan Informatika

Klasifikasi Keluhan Mahasiswa di Media Sosial Whatsapp Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Amelia Vergi Kusuma¹, Eliyani²1Program Studi Teknik Informatika, Fakultas Ilmu Komputer
Universitas Mercu Buana

Jl. Meruya Selatan No. 1 Kembangan Jakarta Barat

*ameliavergi@gmail.com

Abstract- Salah satu indikator kepuasan pengguna jasa pendidikan adalah keluhan yang disampaikan atas pelayanan yang diberikan melalui media yang disediakan oleh lembaga maupun melalui media lainnya. Untuk meningkatkan kualitas pelayanan, perlu untuk diketahui keluhan yang disampaikan agar dapat segera tertangani. Salah satu sumber untuk mendapatkan keluhan ini yaitu dari media sosial yang bersifat tertutup seperti grup-grup Whatsapp. Penelitian ini ditujukan untuk melakukan klasifikasi teks dari grup-grup Whatsapp mahasiswa dengan memanfaatkan teknik *text mining* menggunakan metode *Naive Bayes* dan *Support Vector Machine* sebagai keluhan atau bukan keluhan. Terdapat 2047 data teks dari 15 grup mahasiswa. Pelabelan dilakukan secara otomatis menggunakan script. Hasil penelitian menunjukkan nilai *accuracy*, *precision* dan *f1-measure* algoritma *Support Vector Machine* memiliki performansi lebih baik dalam mengklasifikasikan keluhan dibanding algoritma *Naïve Bayes*. Nilai *accuracy*, *precision*, *recall* dan *f1-measure* algoritma *Support Vector Machine* masing-masing sebesar 0.938, 0.942, 0.990 dan 0.965. Sedangkan untuk algoritma *Naïve Bayes* menghasilkan nilai *accuracy* sebesar 0.882, nilai *precision* sebesar 0.880, nilai *recall* sebesar 1.000 dan nilai *f1-measure* sebesar 0.936. Berdasarkan hasil confusion matrix, keluhan yang disampaikan melalui obrolan di grup ini relatif sedikit, baik yang dideteksi melalui algoritma *Support Vector Machine* sebesar 8.5% maupun *Naive Bayes* yang sebesar 2%.

Kata Kunci: *Keluhan, Klasifikasi, Naive Bayes, Support Vector Machine, Text Mining.*

1. Pendahuluan

Mutu atau kualitas sebuah lembaga pendidikan menjadi salah satu alasan utama bagi setiap peserta didik yang akan melalui proses pendidikannya. Mutu dalam perguruan tinggi bisa dilihat dari rumusan visi dan misinya, yang kemudian diwujudkan dalam proses pendidikan yang akan dilakukan. Pendidikan tinggi yang berkualitas sangat penting bagi sebuah negara. Terutama bagi negara-negara berkembang seperti di Indonesia[1]. Hal ini dipicu oleh berbagai faktor, yaitu faktor eksternal seperti telah lahirnya era baru yaitu era revolusi industri 4.0 pada era ini banyak terjadi perkembangan teknologi dan faktor internal, yakni terkait dengan kualitas

Universitas Mercu Buana

perguruan tinggi di Indonesia yang pada umumnya masih tergolong rendah dibandingkan dengan Negara lain[2].

Dalam meningkatkan mutunya pihak perguruan tinggi melakukan strategi perubahan dari salah satu subsistem manusia dan teknologi yaitu mahasiswa yang dididik, dosen sebagai pendidik dan pengajar, sarana dan prasarana, kurikulum akademik, kerjasama di berbagai bidang, dan proses organisasi[3]. Akan tetapi dalam melakukan perubahan pada subsistem manusia dan teknologi masih belum maksimal sehingga tidak jarang mahasiswa menyampaikan keluhannya melalui media sosial whatsapp. Keluhan tersebut terjadi akibat adanya kejadian yang tidak diinginkan atau hal yang terjadi tidak sesuai harapan mahasiswa.

Akan tetapi keluhan tersebut dapat bersifat membangun apabila disampaikan dengan tepat. Namun dapat menjadi isu negatif dan penyebar kebencian apabila tidak disampaikan dengan tepat dan dikonsumsi langsung oleh pihak terkait[4]. Berdasarkan hal tersebut diperlukannya klasifikasi dokumen teks dengan memanfaatkan *text mining* teknik dalam melakukan pengkategorisasian dan pengelompokan teks keluhan mahasiswa.

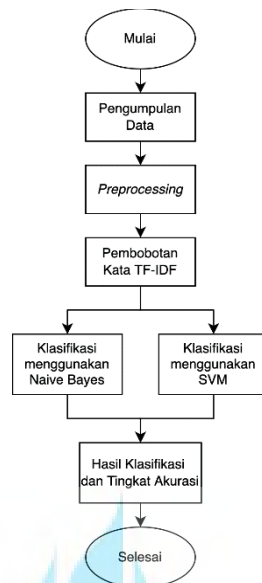
Text mining merupakan teknik mengolah data yang berbentuk teks, biasanya mengekstraksi sumber data dari dokumen dengan tujuan mendeteksi kata-kata yang bisa memperantarai isi artikel atau teks untuk melakukan analisis hubungan antar dokumen[5].

Penelitian sebelumnya terkait dengan klasifikasi keluhan adalah penelitian yang dilakukan oleh Hermanto, Ali Mustopa dan Antonius Yadi Kuntoro yaitu membahas tentang perbandingan metode *Naive Bayes* dan SVM (*Support Vector Machine*) dalam layanan complain mahasiswa menghasilkan bahwa *Support Vector Machine* lebih unggul disbanding dengan *Naive Bayes*. Untuk SVM nilai akurasi 84,45% dan nilai AUC 0,922, sedangkan untuk algoritma *Naive Bayes* nilai akurasi 69,75% dan nilai AUC 0,679[6]. Penelitian lain terkait dengan klasifikasi adalah penelitian yang dilakukan oleh Aan Rohanah, Dwi Latifah Rianti dan Betha Nurina Sari tentang perbandingan metode *Naive Bayes* dan SVM (*Support Vector Machine*) dalam klasifikasi ulasan pelanggan indihome dimana penelitian tersebut menghasilkan algoritme terbaik untuk mengklasifikasikan ulasan pengguna IndiHome pada media sosial Twitter adalah *Support Vector Machine* lebih baik dengan nilai accuracy 82,11%, Precision 76,44%, recall 88,01%, dan nilai AUC 0,909[7].

Berdasarkan latar belakang yang telah dipaparkan diatas, maka akan melakukan perbandingan metode *Naive Bayes* dan SVM (*Support Vector Machine*) terhadap klasifikasi keluhan mahasiswa sehingga mengetahui mana metode yang memiliki akurasi terbaik. Selain itu dengan adanya hasil dari klasifikasi keluhan tersebut diharapkan dapat membantu pihak perguruan tinggi dalam melakukan peningkatan mutu.

2. Metode

Proses dalam melakukan klasifikasi keluhan mahasiswa menggunakan metode *Naive Bayes* dan *Support Vector Machine* sehubungan dengan mutu perguruan tinggi yang dilakukan dalam penelitian ini disajikan pada Gambar 1.



Gambar 1. Alur Penelitian

Adapun penjelasan dari tahapan-tahapan diatas yaitu:

1. Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini, yaitu *scrapping* data. Data yang digunakan untuk melakukan klasifikasi keluhan mahasiswa dikumpulkan melalui *export chat whatsapp*. Data yang diambil berupa keluhan mahasiswa yang terdiri dari komplain SIA, komplain TU, komplain administrasi dan *Post error*.

2. Preprocessing

Tahap *preprocessing* adalah tahapan dimana aplikasi melakukan seleksi data yang akan diproses pada setiap dokumen. Proses *preprocessing* meliputi (1) *Case Folding* (menyeragamkan bentuk huruf) (2) *Tokenizing* (pemenggalan suku kata) (3) *Filtering* (menghilangkan kata yang tidak deskriptif) (4) *Stemming* (mengubah suku kata menjadi kata dasar)[8].

a.) Case Folding

Tahap *case folding* yaitu tahapan menyamakan kata. Tahapan ini dilakukan dengan mengubah huruf besar menjadi *lowercase* (huruf kecil)[9].

Tabel 1. Proses *Case Folding*

Sebelum	Sesudah
Menurut saya masih banyak sekali fasilitas kampus yang kurang, contohnya informasi akademik yang kurang <i>up to date</i> .	menurut saya masih banyak sekali fasilitas kampus yang kurang, contohnya informasi akademik yang kurang <i>up to date</i> .

b.) Tokenizing

Tokenizing merupakan tahapan pemotongan *string* untuk memisahkan setiap kata menjadi kata tunggal[9].

Tabel 2. Proses *Tokenizing*

Sebelum	Sesudah
Menurut saya masih banyak sekali fasilitas kampus yang kurang, contohnya informasi akademik yang kurang <i>up to date</i> .	“menurut”, “saya”, “masih”, “banyak”, “sekali”, “fasilitas”, “kampus”, “yang”, “kurang”, “contohnya”, “informasi”, “akademik”, “yang”, “kurang”, “ <i>up-to-date</i> ”

c.) Filtering

Filtering merupakan Kata-kata yang sering muncul secara umum dan kurang relevan dengan teks akan dihapus[10].

Tabel 3. Proses *Filtering*

Sebelum	Sesudah
“menurut”, “saya”, “masih”, “banyak”, “sekali”, “fasilitas”, “kampus”, “yang”, “kurang”, “contohnya”, “informasi”, “akademik”, “yang”, “kurang”, “ <i>up-to-date</i> ”	“fasilitas”, “kampus”, “banyak”, “informasi”, “contohnya”, “akademik”, “kurang”, “ <i>up-to-date</i> ”

d.) Stemming

Tahap *stemming* berfungsi untuk mengubah kata imbuhan seperti “ber-, di-, ke-, me-, -nya, -kan” menjadi kata dasar. Tahapan ini juga mengubah kata turunan menjadi kata baku[9].

Tabel 4. Proses *Stemming*

Sebelum	Sesudah
“menurut”, “saya”, “masih”, “banyak”, “sekali”, “fasilitas”, “kampus”, “yang”, “kurang”, “contohnya”, “informasi”, “akademik”, “yang”, “kurang”, “ <i>up-to-date</i> ”	“fasilitas”, “kampus”, “kurang”, “informasi”, “contoh”, “akademik”, “kurang”, “ <i>up-to-date</i> ”

3. Pembobotan Kata TF-IDF

Setelah melakukan preprocessing maka tahap selanjutnya adalah melakukan pembobotan kata menggunakan TF-IDF. TF-IDF merupakan salah satu metode yang digunakan untuk

melakukan pembobotan kata[11]. Adapun tahap pembobotan kata TF-IDF adalah menentukan banyak term atau kata pada setiap dokumen, kemudian menghitung df_i , idf_i , W_{tf} , W_{td} hingga normalisasi.

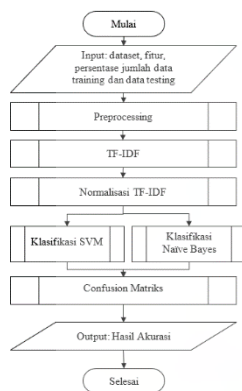
4. Klasifikasi

Pada tahap ini dilakukan klasifikasi menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) pada klasifikasi keluhan mahasiswa di perguruan tinggi.

5. Hasil Klasifikasi dan Tingkat Akurasi

Pada tahap ini merupakan tahapan yang terakhir yaitu penarikan kesimpulan dari proses yang sebelumnya dilakukan. Tahap ini dapat mengetahui hasil klasifikasi dan tingkat akurasi menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) pada klasifikasi keluhan mahasiswa di perguruan tinggi.

Keseluruhan tahapan tersebut disajikan pada Gambar 2.



Gambar 2. Flowchart algoritma *Naïve Bayes* dan SVM

3. Hasil dan Diskusi

Dalam penelitian ini menggunakan metode *Naïve Bayes* dan SVM untuk melakukan klasifikasi keluhan mahasiswa.

1. Dataset

Dataset yang digunakan berasal dari data keluhan mahasiswa perguruan tinggi yang diambil dari hasil export chat media social whatsapp ditunjukkan pada Tabel 5.

Tabel 5. *Dataset* Keluhan Mahasiswa

Doc	Text	Kelas
1	Makasih mas rangka ngga ngambil	1
2	makas mas rangka engga ngambil	1
3	mahasiswa mas	1
4	oh iya upload file prestasi file upload gagal kemarin	1
5	ken upload surat janji isi email	1
6	krs sso nge down ga login	-1
7	email staf buruk	-1
	lambat	-1

8	error	-1
9	internal server error server ganggu	-1
10	Bayar titip pake bukopin sistem error	-1

Keterangan:

Kelas -1 artinya negatif (komplain)

Kelas 1 artinya positif (tidak komplain)

Penentuan “Komplain” ditentukan berdasarkan kata yang mengandung unsur negatif dalam suatu dokumen. Sedangkan penentuan “Tidak Komplain” ditentukan berdasarkan kata yang mengandung unsur positif dalam suatu dokumen.

Nilai parameter yang digunakan pada manualisasi algoritme adalah untuk SVM adalah kernel polynomial, dan Persentase jumlah data *training* dan data *testing* sebesar 80%:20% dan 50%:50%.

2. Preprocessing

Preprocessing dilakukan untuk mengubah *text* menjadi *term index* yang bisa mewakili dokumen. Adapun tahap yang dilakukan pada preprocessing adalah *tokenizing*, *filtering*, dan *stemming*[12]. *Tokenizing* akan memisahkan teks menjadi beberapa potongan berupa token, pada proses ini juga dilakukan pembuangan karakter yang tidak diperlukan seperti (.) (,) (). Kemudian *filtering* akan menghapus kata yang tidak penting atau masuk dalam daftar *stopword*, contohnya seperti dan, ini, itu, saya dll. Dan yang terakhir adalah *stemming*, yaitu proses untuk menghilangkan imbuhan pada kata, contohnya kata dipakai akan menjadi kata pakai setelah melalui proses stemming. Berikut adalah hasil tahap preprocessing ditunjukkan pada Tabel 6, sementara hasil *preprocessing* tanpa *stopword* disajikan pada Tabel 7.

Tabel 6. Hasil *Preprocessing*

Doc ke-i	Isi Doc	<i>Tokenizing</i>	<i>Stemming</i>
1	Makasih mas rangga ngga ngambil	makasih mas rangga ngga ngambil	makasih mas rangga
2	makasih mas rangga engga ngambil	makasih mas rangga engga ngambil	makasih mas rangga
3	mahasiswa mas oh iya upload file	mahasiswa mas oh iya upload	mahasiswa upload file
4	prestasi file upload gagal kemarin	file prestasi file upload gagal kemarin	prestasi gagal kemairn
5	ken upload surat janji isi krs sso nge down ga login email	ken upload surat janji isi krs	upload surat janji isi krs sso down

		sso nge down ga login email	
6	staf buruk	staf buruk	staf buruk
7	lambat	Lambat	lambat
8	error	Error	error
	internal server error	internal server	internal
9	server ganggu	error server ganggu	server error
	Bayar titip pake	bayar titip pake	sistem error
10	bukopin sistem error	bukopin sistem error	

Tabel 7. Hasil *Preprocessing* tanpa *Stopword*

Doc	Isi Doc	<i>Tokenizing</i>	<i>Stemming</i>
ke-i			
1	Makasih mas rangka ngga ngambil	makasih mas rangka ngga ngambil	makasih mas rangka
2	makasih mas rangka engga ngambil	makasih mas rangka engga ngambil	makasih mas rangka
3	mahasiswa mas	mahasiswa mas	mahasiswa
4	oh iya upload file prestasi file upload gagal kemarin	oh iya upload file prestasi file upload gagal kemarin	upload file prestasi gagal kemairn
5	ken upload surat janji isi krs sso nge down ga login email	ken upload surat janji isi krs sso nge down ga login email	upload surat janji isi krs sso down
6	staf buruk	staf buruk	staf buruk
7	lambat	Lambat	lambat
8	error	Error	error
9	internal server error server ganggu	internal server error server ganggu	internal server error
10	Bayar titip pake bukopin sistem error	bayar titip pake bukopin sistem error	sistem error

Sebelum dilakukan pembagian data *training* dan data *testing*[13], langkah yang perlu dilakukan adalah menentukan banyaknya kata atau *term t* pada setiap dokumen. Berikut adalah proses penentuan banyaknya *term t* pada setiap dokumen ditunjukkan pada Tabel 8.

Tabel 8. Hasil Perhitungan Term untuk Setiap Dokumen

IF Doc ke-i

No	Term	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	Buruk	0	0	0	0	0	1	0	0	0	0
2	Down	0	0	0	0	0	1	0	0	0	0
3	Error	0	0	0	0	0	0	0	1	1	0
4	File	0	0	0	1	0	0	0	0	0	0
5	Gagal	0	0	0	1	0	0	0	0	0	0
6	Internal	0	0	0	0	0	0	0	0	1	0
7	Isi	0	0	0	0	1	0	0	0	0	0
8	Janji	0	0	0	0	1	0	0	0	0	0

3. Perhitungan TF-IDF

Pada perhitungan TF-IDF data yang akan digunakan adalah data *training*. Sedangkan perhitungan TF-IDF untuk data testing akan menggunakan df_t dan idf_t yang dihasilkan pada data *training*[14]. df_t merupakan jumlah dokumen yang mengandung term t dan idf_t merupakan kebalikan dari df_t . Berikut adalah perhitungan df_t dan idf_t menggunakan data *training* ditunjukkan pada Tabel 9.

$$idf_t = \log_{10} \left(\frac{N}{df_t} \right) \quad (1)$$

$$idf_{internal} = \log_{10} \left(\frac{N}{df_{internal}} \right) \quad (2)$$

$$idf_{internal} = \log_{10} \left(\frac{10}{10} \right) \quad (3)$$

$$idf_{internal} = 1 \quad (4)$$

Tabel 9. Hasil Perhitungan df dan idf Menggunakan Data Training

TF No	Term	Doc ke-i					df	idf
		D1	D2	D3	D4	D5		
1	Buruk	0	0	0	0	0	1	1
2	Down	0	0	0	0	0	1	1
3	Error	0	0	0	0	0	2	0.69897
4	File	0	0	0	0	1	1	1
5	Gagal	0	0	0	0	1	1	1
6	Internal	0	0	0	0	0	1	1
7	Isi	0	0	0	0	0	1	1
8	Janji	0	0	0	0	0	1	1

Bobot term atau $w_{tf_{t,d}}$ merupakan logaritma dari frekuensi kemunculan term pada setiap dokumen. Logaritma berfungsi untuk mengurangi perbedaan term yang terlalu tinggi. Bobot term dapat dihitung sebagai berikut.

$$w_{tf_{t,d}} = \begin{cases} 1 + \log_{10}(tf_{t,d}), & \text{IF } tf_{t,d} > 0 \\ 0, & \text{IF } tf_{t,d} = 0 \end{cases} \quad (5)$$

$$w_{1,2} = 1 + \log_{10}(1) \quad (6)$$

$$w_{1,2} = 1 + 0 \quad (7)$$

$$w_{1,2} = 1 \quad (8)$$

Tabel 10. Hasil Perhitungan w_{tf} Menggunakan Data Training

TF	Doc ke-i	df	idf
----	----------	----	-----

No	Term	D1	D2	D3	D4	D5		
1	Buruk	0	0	0	0	0	1	1
2	Down	0	0	0	0	0	1	1
3	Error	0	0	0	0	0	2	0.69897
4	File	0	0	0	0	1	1	1
5	Gagal	0	0	0	0	1	1	1
6	Internal	0	0	0	0	0	1	1
7	Isi	0	0	0	0	0	1	1
8	Janji	0	0	0	0	0	1	1

Nilai TF-IDF dari sebuah term t merupakan perkalian antara $w_{tf_{t,d}}$ dan idf_t . TF-IDF.

$$w_{t,d} = w_{tf_{t,d}} \times idf_t \quad (9)$$

$$w_{1,1} = 0 \times 0 \quad (10)$$

$$w_{1,1} = 0 \quad (11)$$

Tabel 11. Hasil Perhitungan w_{td} Menggunakan Data *Training*

TF No	Term	Doc ke-i					df	idf
		D1	D2	D3	D4	D5		
1	Buruk	0	0	0	0	0	1	1
2	Down	0	0	0	0	0	1	1
3	Error	0	0	0	0	0	2	0.69897
4	File	0	0	0	0	1	1	1
5	Gagal	0	0	0	0	1	1	1
6	Internal	0	0	0	0	0	1	1
7	Isi	0	0	0	0	0	1	1
8	Janji	0	0	0	0	0	1	1

Berdasarkan Tabel 6.6, Untuk dilakukan perhitungan klasifikasi *text mining* menggunakan algoritme SVM data hasil perhitungan normalisasi $w_{t,d}$ perlu ditranspose. Term akan digunakan sebagai fitur. Berikut adalah transpose matriks normalisasi $w_{t,d}$ serta dokumen yang telah diberi label kelasnya ditunjukkan pada Tabel 12.

Tabel 12. Kelas Label

Doc	Fitur 1	Fitur 2	Fitur 3	Fitur 4	Fitur 5	Fitur 6	Fitur 7	Fitur 8	Kelas
1	0	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	1
4	0	0	0	0	0	0	0	0	1
5	0	0	0	0.009709	0.009709	0	0	0	1
6	0	0.009709	0	0	0	0	0.009709	0.009709	-1
7	0.009709	0	0	0	0	0	0	0	-1
8	0	0	0	0	0	0	0	0	-1
9	0	0	0.031758	0	0	0	0	0	-1
10	0	0	0.032799	0	0	0	0	0	-1

4. Melakukan Klasifikasi
 1. Naïve Bayes

a) Melakukan Pembobotan

Diperoleh : positif = 4,69897, negatif = 12,09691001.

b) Melakukan Pembobotan

Hitung probabilitas prior setiap kategori, yang menjadi 2 buah kategori yaitu kelas positif (tidak complain) dan negative (complain).

$$P(\text{positif}) = \frac{fx(\text{positif})}{|C|} = \frac{1}{2} \quad (12)$$

$$P(\text{negatif}) = \frac{fx(\text{negatif})}{|C|} = \frac{1}{2} \quad (13)$$

c) Melakukan jumlah

Menghitung probabilitas likelihood setiap *term* dari semua dokumen. Jumlah seluruh kata adalah 16.

$$P(w | \text{positif/negatif}) = \frac{(nk(\frac{\text{positif}}{\text{negatif}})+1)}{(n.\frac{\text{positif}}{\text{negatif}}+|T|)} \quad (14)$$

1. Probabilitas Kata “buruk”

$$P(\text{buruk}|\text{positif}) = \frac{0 + 1}{4,69897 + 16} = 16,2128196$$

$$P(\text{buruk}|\text{negatif}) = \frac{1 + 1}{12,09691001 + 16} = 17,08266574$$

2. Probabilitas Kata “Down”

$$P(\text{down}|\text{positif}) = \frac{0 + 1}{4,69897 + 16} = 16,2128196$$

$$P(\text{down}|\text{negatif}) = \frac{1 + 1}{12,09691001 + 16} = 17,08266574$$

d) Hitung Masing-Masing Dokumen

Hitung masing-masing dokumen untuk diklasifikasikan ke dalam kelas positif atau negatif dengan rumus.

$$P(\text{positif}) \times P(\text{term}|\text{positif}) \quad (15)$$

Contoh pada D9:

internal server error server ganggu

$P(\text{positif}) \times P(\text{term}|\text{positif})$

$$= 0,5 \times 17,08266574 \times 16,2128196 \times 16,2128196$$

$$= 138,479$$

$P(\text{negatif}) \times P(\text{term}|\text{negatif})$

$$= 0,5 \times 16,08266574 \times 17,08266574 \times 17,48060575$$

$$= 2.401,26$$

Maka sentiment D9 diklasifikasikan ke dalam kelas negative (komplain). Jika nilai yang diperoleh pada sentiment dokumen lebih besar untuk prior positif, maka dokumen tersebut nantinya akan diklasifikasikan ke dalam kelas positif, jika nilai sentiment dokumen lebih besar untuk prior negative, maka sentiment tersebut akan diklasifikasikan ke dalam kelas negative.

2. Support Vector Machine

a. Matriks Hessian

Melakukan perhitungan matriks hessian menggunakan fungsi kernel polynomial of degree up to d.

Dalam perhitungan matriks hessian terdapat perhitungan kernel. Berikut adalah perhitungan kernel menggunakan kernel polynomial of degree up to d ditunjukkan pada Tabel 13.

$$D_{ij} = (y_i x y_j) x (K(x_i, x_j) + \lambda^2) \quad (16)$$

$$K(x_i, x_j) = (x_i \cdot x_j + c)^d \quad (17)$$

Tabel 13. Hasil Perhitungan Kernel Polynomial Of Degree Up to d Pada Data Training

	1	2	3	4	5
1	1	1	1	1	1
2	1	8	1	1	1
3	1.0002888	8.001153	1.0002888	1.0002888	1.0002888
4	1	8	1	1	1
5	1.003596	8.014374	1.003596	1.003596	1.003312

Menghitung matriks hessian menggunakan hasil perhitungan kernel *polynomial*, data target dan lambda.

$$D_{ij} = (y_i x y_j) x (K(x_i, x_j) + \lambda^2) \quad (18)$$

Tabel 14. Hasil Perhitungan Matriks Hessian Pada Data Training

	1	2	3	4	5
1	2	2	2	2	2
2	2	9	2	2	2
3	2.000288	9.001153	2.000288	2.000288	2.000288
4	2	9	2	2	2
5	2.003596	9.0014374	2.003596	2.003596	2.003029
6	-0.36078	-8.36915	-0.36043	-0.36043	0
7	-0.00057	-7.00113	-0.00331	-0.00028	-0.00028
8	0	-7	0	0	0

b. Melakukan perulangan i sebanyak iterasi yang telah ditentukan.

c. Melakukan perhitungan nilai Error untuk setiap data pada matriks hessian.

Sequential learning merupakan proses untuk mendapatkan nilai lagrange multiplier. Perhitungan yang terdapat dalam sequential learning adalah menghitung nilai error dan update nilai lagrange multiplier atau update alfa. Dalam sequential learning proses update alfa akan terus dilakukan sampai maksimum iterasi yang telah ditentukan. Sebelum menghitung nilai error dan memulai proses update alfa perlu dilakukan inisialisasi nilai alfa sebanyak kolom matriks hessian pada data *training*. Berikut adalah inisialisasi nilai alfa ditunjukkan pada Tabel 14.

Tabel 15. Inisialisasi Nilai *Alfa*

	Alfa
1	0
2	0
3	0
4	0
5	0

Menghitung nilai error dengan mengkalikan nilai alfa dengan matriks hessian.

$$E_i = \sum_{j=1}^n \alpha_j D_{ij} \quad (19)$$

- d. Melakukan perhitungan untuk mengupdate nilai Lagrange Multiplier.

Perhitungan nilai alfa membutuhkan nilai alfa awal, cLR, *gamma*, dan *complexity*. Langkah pertama pada perhitungan nilai alfa adalah menentukan nilai *gamma* menggunakan nilai cLR dan matriks *hessian*. Berikut adalah perhitungan untuk menentukan nilai *gamma*.

$$\gamma = \frac{cLR}{\max(\text{Matriks Hessian})}$$

$$\gamma = \frac{0,1}{9,014374}$$

$$\gamma = 0,011093$$

Langkah kedua adalah menghitung perubahan nilai alfa. Berikut adalah perhitungan update alfa untuk alfa pertama.

$$\delta\alpha_i = \min\{\max[\gamma(1 - E_i), \alpha_i], C - \alpha_i\}$$

$$\delta\alpha_1 = \min\{\max[0,011093(1 - 0), 0], 1 - 0\}$$

$$\delta\alpha_1 = \min\{\max[0,011093(1), 0], 1\}$$

$$\delta\alpha_1 = \min\{\max[0,011093, 0], 1\}$$

$$\delta\alpha_1 = \min\{0,011093, 1\}$$

$$\delta\alpha_1 = 0,011093$$

Langkah ketiga adalah mengupdate nilai alfa menggunakan hasil perhitungan sebelumnya. Berikut adalah update nilai alfa untuk alfa pertama.

$$\alpha_1 = \alpha_1 + \delta\alpha_1$$

$$\alpha_1 = 0 + 0,011093$$

$$\alpha_1 = 0,011093$$

Perhitungan nilai bias digunakan pada proses testing agar dapat menghasilkan suatu kelas pada data baru atau data

yang akan diklasifikasikan. Perhitungan bias akan mengambil tiga data dari kelas yang berbeda.

$$bias_i = -\frac{1}{2} \cdot (\sum_{j=1}^n \alpha_j y_j K(x_i, x^-) + \sum_{j=1}^n \alpha_j y_j K(x_i, x^+))$$

$$bias_1 = -\frac{1}{2} \cdot ((1 \times (1) \times 29,40547) + (1 \times 1 \times 23,06719))$$

$$bias_1 = -\frac{1}{2} \cdot (29,40547 + 23,06719)$$

$$bias_1 = -0,29102$$

Tabel 16. Hasil Perhitungan Bias

No	Bias
1	-0.29102
2	-0.96597
3	-0.96611
4	-0.96597
5	-0.96797

e. Proses testing

Melakukan proses *testing* untuk mendapatkan hasil klasifikasi menggunakan nilai bias pada perhitungan sebelumnya dan menggunakan metode one against all untuk mendapatkan kelas. Proses *testing* digunakan untuk menghasilkan kelas baru menggunakan sebuah fungsi sign untuk menentukan kelas negatif atau kelas positif. Nilai yang digunakan untuk menghitung nilai testing adalah nilai alfa, data target, perhitungan kernel untuk data prediksi dan nilai bias pada perhitungan sebelumnya. Berikut adalah perhitungan proses *testing* untuk menghasilkan kelas baru.

$$f(x_{Prediksi}) = \text{Sign}(\sum_{i=1}^m \alpha_i y_i K(x_i, x_{Pred}) + b_i) \quad (20)$$

$$x_{pred_i} = \alpha_i y_i K(x_i, x_{Pred}) \quad (21)$$

Tabel 17. Before Prediksi

No	Xpred ke-i
1	0.011093394
2	0.011093394
3	0.011096593
4	0.011093394
5	0.011133281

$$f(x_6) = \text{Sign}((0,011093394 + (-0,29102)) + (0,011093394 + (-0,96597)) + (0,011096593 + (-0,96611)) + (0,011093394 + (-0,96597)) + (0,011133281 + (-0,96597)))$$

$$f(x_6) = -1$$

Tabel 18. Hasil Klasifikasi Menggunakan Data *Testing*

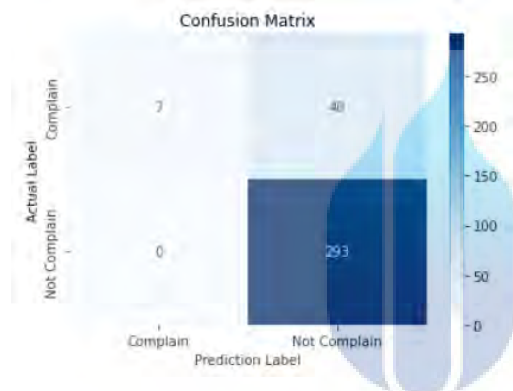
Dokumen	Hasil Klasifikasi	Kelas
6	-1	Kelas -1
7	-1	Kelas -1
8	-1	Kelas -1
9	-1	Kelas -1
10	-1	Kelas -1

5. Melakukan Evaluasi

Melakukan evaluasi hasil klasifikasi menggunakan perhitungan confusion matriks. *Confusion matrix* merupakan perhitungan matriks yang digunakan untuk mengukur berapa jumlah keberhasilan algoritme dalam mengklasifikasikan kelas yang sesuai dengan kelas yang sebenarnya[15].

a.) Naïve Bayes

Berikut ini hasil perhitungan *confusion matrix* dengan algoritma *Naïve Bayes*.

Gambar 3. Hasil *Confusion Matrik* Pada Algoritma *Naïve Bayes*

Berikut ini merupakan perhitungan nilai *accuracy*, *precision*, *recall* dan *F1-Measure* yang dihasilkan dari algoritma *Naïve Bayes* yaitu:

$$\text{Akurasi} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\%$$

$$\text{Akurasi} = \frac{(7 + 293)}{(7 + 293 + 40 + 0)} \times 100\%$$

$$\text{Akurasi} = \frac{300}{340} \times 100\%$$

$$\text{Akurasi} = 0,88 \times 100\%$$

$$\text{Akurasi} = 88\%$$

$$\text{Recall} = \frac{(TP)}{(TP + FN)}$$

$$\text{Recall} = \frac{(7)}{(7 + 0)}$$

$$\text{Recall} = \frac{(7)}{(7)}$$

$$\text{Recall} = 1 \times 100\%$$

$$\text{Recall} = 100\%$$

$$\text{Precision} = \frac{(\text{TP})}{(\text{TP} + \text{FP})}$$

$$\text{Precision} = \frac{(7)}{(7 + 40)}$$

$$\text{Precision} = \frac{(7)}{(47)}$$

$$\text{Precision} = 0,14 * 100\%$$

$$\text{Precision} = 14\%$$

Berikut ini merupakan hasil perhitungan nilai *accuracy*, *precision recall* dan *F1-Measure* yang dihasilkan dari algoritma *Naïve Bayes* yaitu:

$$\text{Accuracy Score} = 0.882$$

$$\text{Precision} = 0.880$$

$$\text{Recall} = 1.000$$

$$\text{F-Score} = 0.936$$

b.) Support Vector Machine

Berikut ini merupakan hasil perhitungan *confusion matrix* pada algoritma SVM.



Gambar 4. Hasil *Confusion Matrik* Pada Algoritma SVM

Berikut ini merupakan perhitungan nilai *accuracy*, *precision recall* dan *F1-Measure* yang dihasilkan dari algoritma *Support Vector Machine* yaitu:

$$\text{Akurasi} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \times 100\%$$

$$\text{Akurasi} = \frac{(29 + 290)}{(29 + 290 + 18 + 3)} \times 100\%$$

$$\text{Akurasi} = \frac{319}{340} \times 100\%$$

$$\text{Akurasi} = 0,938 * 100\%$$

$$\text{Akurasi} = 94\%$$

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})}$$

$$\text{Recall} = \frac{(29)}{(29 + 3)}$$

$$\text{Recall} = \frac{(29)}{(32)}$$

$$\text{Recall} = 0,90 * 100\%$$

$$\text{Recall} = 90\%$$

$$\text{Precision} = \frac{(TP)}{(TP + FP)}$$

$$\text{Precision} = \frac{(29)}{(29 + 18)}$$

$$\text{Precision} = \frac{(29)}{(47)}$$

$$\text{Precision} = 0,61 * 100\%$$

$$\text{Precision} = 61\%$$

Berikut ini merupakan hasil perhitungan nilai *accuracy*, *precision*, *recall* dan *F1-Measure* yang dihasilkan dari algoritma *Support Vector Machine* yaitu:

$$\text{Accuracy Score} = 0.938$$

$$\text{Precision} = 0.942$$

$$\text{Recall} = 0.990$$

$$\text{F-Score} = 0.965$$

Berdasarkan pengujian pada data *training* untuk klasifikasi keluhan mahasiswa menggunakan dua algoritma yaitu algoritma *Naïve Bayes* dan SVM. Dihasilkan bahwa perhitungan pada algoritma SVM menghasilkan nilai *accuracy*, *precision* dan *f1-measure* lebih tinggi namun nilai *recall* yang dihasilkan lebih rendah dari algoritma *Naïve Bayes*. Adapun hasil tersebut dapat dilihat pada tabel 12.

Tabel 19. Hasil Keseluruhan

Metode	Accuracy	Precision	Recall	F-Score
<i>Naïve Bayes</i>	0.882	0.880	1.000	0.936
<i>Support Vector Machine</i>	0.938	0.942	0.990	0.965

4. Kesimpulan

Berdasarkan penelitian yang dilakukan untuk Klasifikasi Keluhan Mahasiswa di Media Sosial Whatsapp Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine* didapatkan kesimpulan sebagai berikut:

1. Pada penelitian ini dapat diterapkan algoritma *naïve bayes* dan SVM untuk melakukan klasifikasi keluhan mahasiswa pada perguruan tinggi dengan jumlah dataset yang digunakan yaitu sebanyak 2047 data setelah melalui tahapan *preprocessing* menjadi 1699 data keluhan untuk data training dan data testing serta label yang digunakan yaitu label *complain* dan *not complain*.
2. Penelitian ini nilai *accuracy*, *precision* dan *f1-measure* yang dihasilkan oleh algoritma SVM lebih tinggi dibandingkan dengan algoritma *naïve bayes* sehingga dapat digunakan untuk memberikan solusi untuk perguruan tinggi dalam hal meningkatkan mutu perguruan tinggi.
3. Pada algoritma SVM dihasilkan nilai *accuracy* sebesar 0.938, nilai *precision* sebesar 0.942, nilai *recall* sebesar 0.990 dan nilai *f1-measure* sebesar 0.965. Sedangkan, untuk algoritma *naïve bayes* menghasilkan nilai *accuracy* sebesar 0.882, nilai *precision* sebesar 0.880, nilai *recall* sebesar 1.000 dan nilai *f1-measure* sebesar 0.936.

Daftar Pustaka

- [1] M. Fadhli, "Sistem Penjaminan Mutu Internal Dan Eksternal Pada Lembaga Pendidikan Tinggi," *AL-TANZIM J. Manaj. Pendidik. Islam*, vol. 4, no. 2, pp. 53–65, 2020, doi: 10.33650/al-tanzim.v4i2.1148.
- [2] A. Online and J. Page, "Available Online: <https://dinastirev.org/JIMT> Page 416," vol. 1, no. 5, pp. 416–424, 2020, doi: 10.31933/JIMT.
- [3] S. Angkotasan and S. Watianan, "Faktor-Faktor Yang Mempengaruhi Peningkatan Mutu Pendidikan Di Kampus Stia Alazka Ambon," *KOMUNITAS J. Ilmu Sosiol.*, vol. 4, no. 2, pp. 42–50, 2021, doi: 10.30598/komunitasvol4issue2page42-50.
- [4] H. Hermanto, A. Mustopa, and A. Y. Kuntoro, "Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 5, no. 2, pp. 211–220, 2020, doi: 10.33480/jitk.v5i2.1181.
- [5] D. Ariyanti and K. Iswardani, "Teks Mining untuk Klasifikasi Keluhan Masyarakat Pada Pemkot Probolinggo Menggunakan Algoritma Naive Bayes," *J. IKRA-ITH Inform.*, vol. 4, no. 3, pp. 125–132, 2020.
- [6] H. P. Doloksaribu and Y. T. Samuel, "Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, pp. 1–11, 2022.
- [7] A. Rohanah, D. L. Rianti, and B. N. Sari, "Perbandingan Naive Bayes dan Support Vector Machine untuk Klasifikasi Ulasan Pelanggan Indihome," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 6, no. 1, p. 23, 2021, doi: 10.30998/string.v6i1.9232.
- [8] I. Wahyudi, S. Bahri, and P. Handayani, "Aplikasi Pembelajaran Pengenalan Budaya Indonesia," vol. V, no. 1, pp. 135–138, 2019, doi: 10.31294/jtk.v4i2.
- [9] H. Tuhuteru and A. Iriani, "Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier," *J. Inform. J. Pengemb. IT*, vol. 3, no. 3, pp. 394–401, 2018, doi: 10.30591/jpit.v3i3.977.
- [10] Y. A. Rohman, "Pengenalan NumPy, Pandas, Matplotlib," *Medium*. <https://medium.com/@yasirabd/pengenalan-numpy-pandas-matplotlib-b90bafd36c0> (accessed Jan. 05, 2022).
- [11] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naive Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *Smatika J.*, vol. 10, no. 02, pp. 71–76, 2020, doi: 10.32664/smatika.v10i02.455.
- [12] T. T. Widowati and M. Sadikin, "Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naive Bayes dan Support Vector Machine," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 2, pp. 626–636, 2021, doi: 10.24176/simet.v11i2.4568.
- [13] S. Dwiasnati and Y. Devianto, "Optimasi Prediksi Bencana Banjir menggunakan Algoritma SVM untuk penentuan Daerah Rawan Bencana Banjir," *Pros. SISFOTEK*, pp. 202–207, 2021, [Online]. Available: <http://seminar.iaii.or.id/index.php/SISFOTEK/article/view/283>
- [14] W. Feng, "2016 IEEE 35th International Performance Computing and Communications Conference, IPCCC 2016," *2016 IEEE 35th Int. Perform. Comput. Commun. Conf. IPCCC 2016*, 2017.
- [15] A. N. Muhammad, S. Bukhori, and P. Pandunata, "Sentiment Analysis of Positive and Negative of YouTube Comments Using Naive Bayes-Support Vector Machine (NBSVM) Classifier," *Proc. - 2019 Int. Conf. Comput. Sci. Inf. Technol. Electr. Eng. ICOMITEE 2019*, vol. 1, pp. 199–205, 2019, doi: 10.1109/ICOMITEE.2019.8920923.

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul di atas. Kertas kerja berisi semua material hasil penelitian Tugas Akhir yang tidak dimuat/atau disertakan di artikel jurnal. Di dalam kertas kerja ini disajikan: literature review, source code, dataset yang digunakan, tahapan eksperimen dan hasil eksperimen secara keseluruhan.

Pada bagian Literature Review ini ditampilkan hasil review terhadap beberapa literatur atau jurnal ilmiah yang terkait dengan penelitian ini yaitu Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode Naïve Bayes dan *Support Vector Machine*. Literature Review ini terdiri dari 10 Artikel jurnal umum Nasional ataupun Internasional.

Analisis dan Perancangan

Source Code

Dataset berisi asset data yang akan digunakan dalam penelitian Klasifikasi Keluhan Mahasiswa di Media Sosial WhatsApp Menggunakan Metode Naïve Bayes dan *Support Vector Machine*.

Tahapan eksperimen seluruhnya, merupakan penjelasan tahapan – tahapan eksperimen pada laporan ini yang menggunakan rekayasa perangkat lunak yaitu dengan Aplikasi Rapid Miner.

Hasil eksperimen bagian ini berisi hasil semua eksperimen baik yang berhasil maupun yang gagal. Sesuai metode maupun jenis test yang digunakan pada laporan ini yaitu Metode *Naïve Bayes* dan *Support Vector Machine*.