

**PENDETEKSI KATEGORI EMOSI BERDASARKAN
STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA
NAIVE BAYES**

Skripsi



oleh
REVANY ARLITA
71110019

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2015

**PENDETEKSI KATEGORI EMOSI BERDASARKAN
STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA
NAIVE BAYES**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi
Informasi Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

REVANY ARLITA
71110019

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2015

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAIVE BAYES

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 15 Juni 2015



REVANY ARLITA

71110019

HALAMAN PERSETUJUAN

Judul Skripsi : PENDETEKSI KATEGORI EMOSI
BERDASARKAN STATUS MEDIA SOSIAL
MENERAPKAN ALGORITMA NAIVE BAYES

Nama Mahasiswa : REVANY ARLITA

N I M : 71110019

Matakuliah : Skripsi (Tugas Akhir)

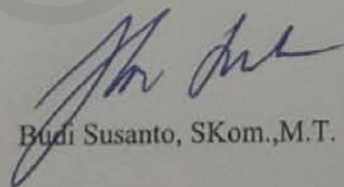
Kode : TIW276

Semester : Genap

Tahun Akademik : 2014/2015

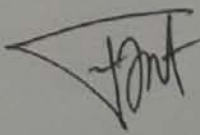
Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 15 Juni 2015

Dosen Pembimbing I



Budi Susanto, SKom.,M.T.

Dosen Pembimbing II



Antonius Rachmat C., S.Kom.,M.Cs.

HALAMAN PENGESAHAN

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAIVE BAYES


Oleh: REVANY ARLITA / 71110019

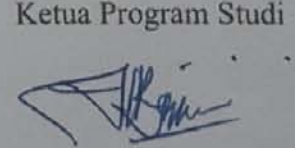
Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 12 Juni 2015

Yogyakarta, 15 Juni 2015
Mengesahkan,

Dewan Penguji:

1. Budi Susanto, SKom.,M.T.
2. Antonius Rachmat C., S.Kom.,M.Cs.
3. Gloria Virginia, S.Kom., MAI, Ph.D.
4. Aditya Wikan Mahastama, S.Kom


Dekan
(Budi Susanto, S.Kom., M.T.)


Ketua Program Studi
(Gloria Virginia, Ph.D.)

UCAPAN TERIMAKASIH

Puji syukur kepada Tuhan Yesus Kristus yang telah melimpahkan anugerah dan kasih karunianya, sehingga penulis dapat menyelesaikan tugas akhir dengan judul “Pendeteksi Kategori Emosi Berdasarkan Status Media Sosial Menerapkan Algoritma Naïve Bayes” ini dengan baik.

Penulis menyadari dalam penelitian ini masih banyak kekurangan, sehingga penulis menerima saran dan kritik dari berbagai pihak, secara langsung maupun tidak langsung. Selain itu, penulis juga menyampaikan ucapan terima kasih kepada :

1. Budi Susanto, SKom.,M.T. selaku dosen pembimbing serta Antonius Rachmat C, S.Kom., M.Cs. sebagai dosen pembimbing dan koordinator Tugas Akhir yang telah membimbing dan memberikan banyak masukan dalam menyelesaikan penelitian ini.
2. Keluarga dan sahabat terkasih yang selalu mendukung dan memberikan semangat kepada penulis selama pengerjaan penelitian ini.
3. Semua *volunteer* baik yang telah mengijinkan statusnya dipakai untuk bahan penelitian ini, maupun *volunteer* yang telah bersedia melakukan pelabelan dalam penelitian ini.
4. Semua pihak lain yang telah membantu dalam pengerjaan penelitian ini yang tidak dapat disebutkan satu per satu.

Akhir kata penulis mengucapkan permintaan maaf bila adanya kesalahan dalam penyusunan laporan ini. Terima kasih.

Yogyakarta, 19 Juni 2015

Penulis

ABSTRAKSI

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAÏVE BAYES

Dalam media sosial, status sering digunakan untuk mencurahkan perasaan penggunanya, seperti emosi sehari-hari senang, sedih, marah, dsb. Namun belum ada layanan selanjutnya untuk mengolah status yang sudah ada. Jikalau status tersebut dapat ditentukan emosinya, layanan dapat ditingkatkan kedepannya, baik untuk kepentingan pengguna maupun dari layanan media sosial tersebut.

Dalam penelitian ini penulis menggunakan algoritma Naïve Bayes dalam melakukan klasifikasi status tersebut, menjadi tujuh kelas emosi. Status dalam media sosial berisi kalimat yang tidak baku sehingga status tersebut diolah terlebih dahulu menjadi dokumen status yang baku, setelah itu akan dilakukan tahap *preprocessing* yaitu *stopword* dan *tokenisasi* agar teks dapat terstruktur. Penelitian ini menggunakan status yang akan dilabeli secara manual sebagai acuan keakuratan sistem. Adapun penelitian ini menggunakan data ISEAR sebagai *feature set*, data ISEAR adalah data yang berisi kumpulan kalimat yang telah dikategorikan menjadi tujuh kelas emosi. Penelitian ini akan diuji keakuratan sistem dengan mengklasifikasi dokumen status uji dengan *feature selection* yang berbeda dan menggunakan data ISEAR maupun tidak.

Hasil dari sistem klasifikasi emosi berdasarkan status media sosial ini memang kurang baik, karena hasil akurasi tertinggi adalah 58,57% dengan *feature selection* 100% dan tanpa menggunakan data ISEAR. Hasil terendah adalah 44,28% dengan *feature selection* 10% dan menggunakan data ISEAR.

Kata kunci : klasifikasi, status, media sosial, Naïve Bayes

DAFTAR ISI

Halaman Judul.....	
Pernyataan Keaslian Skripsi.....	iii
Halaman Persetujuan.....	iv
Halaman Pengesahan.....	v
Ucapan Terimakasih.....	vi
Abstraksi.....	vii
Daftar Isi.....	viii
Daftar Tabel.....	xi
Daftar Gambar.....	xii
Daftar Lampiran.....	xiii
BAB 1 Pendahuluan.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Metode Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB 2 Tinjauan Pustaka.....	5
2.1 Tinjauan Pustaka.....	5
2.2 Landasan Teori.....	6
2.2.1 Facebook.....	6
2.2.2 Emosi.....	7
2.2.3 Teks Mining.....	8
2.2.4 Algoritma Naïve Bayes.....	9
BAB 3 Analisis dan Perancangan Sistem.....	11
3.1 Spesifikasi Sistem.....	11

3.1.1 Fungsional.....	11
3.1.2 Non Fungsional.....	12
3.2 <i>Usecase</i>	13
3.3 <i>Flowchart</i>	14
3.4 Arsitektur Sistem.....	21
3.5 Perancangan Basis Data.....	21
3.6 Perancangan <i>User Interface</i>	23
3.7 Perancangan Pengujian Sistem.....	26
BAB 4 Implementasi dan Analisis Sistem.....	27
4.1 Implementasi Sistem.....	27
4.1.1 Konfigurasi Awal.....	27
4.1.2 Antarmuka Sistem.....	28
4.1.3 <i>Pseudocode</i> dan Kode Sistem.....	31
4.1.3.1 <i>Stopword</i>	31
4.1.3.2 Penghitungan Frekuensi <i>Term</i>	32
4.1.3.3 Algoritma Naïve Bayes.....	33
4.2 Analisis Sistem.....	34
4.2.1 Analisis Secara Manual.....	34
4.2.2 Evaluasi Sistem.....	37
4.2.2.1 <i>Feature Selection</i> 10% tanpa data ISEAR.....	37
4.2.2.2 <i>Feature Selection</i> 20% tanpa data ISEAR.....	38
4.2.2.3 <i>Feature Selection</i> 40% tanpa data ISEAR.....	39
4.2.2.4 <i>Feature Selection</i> 60% tanpa data ISEAR.....	40
4.2.2.5 <i>Feature Selection</i> 100% tanpa data ISEAR.....	41
4.2.2.6 <i>Feature Selection</i> 10% dengan data ISEAR.....	42
4.2.2.7 <i>Feature Selection</i> 20% dengan data ISEAR.....	43
4.2.2.8 <i>Feature Selection</i> 40% dengan data ISEAR.....	44
4.2.2.9 <i>Feature Selection</i> 60% dengan data ISEAR.....	45
4.2.2.10 <i>Feature Selection</i> 100% dengan data ISEAR.....	46

4.2.2.11 Hasil Evaluasi.....	46
BAB 5 Kesimpulan dan Saran.....	50
5.1 Kesimpulan.....	50
5.2 Saran.....	50
Daftar Pustaka.....	
Lampiran.....	

©UKDWN

DAFTAR TABEL

Tabel 2.1	Contoh Kalimat Emosi.....	8
Tabel 2.2	Pemetaan Dokumen Terambil atau Relevan.....	11
Tabel 4.1	Daftar Contoh Token.....	35
Tabel 4.2	Hasil pengujian FS 10% tanpa ISEAR.....	37
Tabel 4.3	Hasil pengujian FS 20% tanpa ISEAR.....	38
Tabel 4.4	Hasil pengujian FS 40% tanpa ISEAR.....	39
Tabel 4.5	Hasil pengujian FS 60% tanpa ISEAR.....	40
Tabel 4.6	Hasil pengujian FS 100% tanpa ISEAR.....	41
Tabel 4.7	Hasil pengujian FS 10% dengan ISEAR.....	42
Tabel 4.8	Hasil pengujian FS 20% dengan ISEAR.....	43
Tabel 4.9	Hasil pengujian FS 40% dengan ISEAR.....	44
Tabel 4.10	Hasil pengujian FS 60% dengan ISEAR.....	45
Tabel 4.11	Hasil pengujian FS 100% dengan ISEAR.....	46
Tabel 4.12	Akurasi Sistem.....	47

DAFTAR GAMBAR

Gambar 3.1	Diagram <i>use-case</i>	13
Gambar 3.2	<i>Flowchart Preprocessing</i> ISEAR.....	15
Gambar 3.3	<i>Flowchart</i> Sistem.....	16
Gambar 3.4	<i>Flowchart</i> Tambah Status.....	17
Gambar 3.5	<i>Flowchart Preprocessing</i> Status	18
Gambar 3.6	<i>Flowchart</i> Pembentukan <i>feature</i>	19
Gambar 3.7	<i>Flowchart</i> klasifikasi Status.....	20
Gambar 3.8	Arsitektur Sistem.....	21
Gambar 3.9	Skema Diagram <i>Database</i>	22
Gambar 3.10	Rancangan Tampilan Klasifikasi.....	24
Gambar 3.11	Rancangan Tampilan Hasil.....	24
Gambar 3.12	Rancangan Tampilan Lihat Status.....	25
Gambar 3.13	Rancangan Tampilan Tambah Status.....	25
Gambar 4.1	<i>Bootstrap</i> yang Dipasang.....	27
Gambar 4.2	<i>Bootstrap</i> yang Dipakai dalam <i>Source Code</i>	28
Gambar 4.3	Halaman Klasifikasi.....	28
Gambar 4.4	Halaman Hasil.....	29
Gambar 4.5	Halaman Lihat Status.....	30
Gambar 4.6	Halaman Tambah Status.....	30
Gambar 4.7	<i>Pseudocode</i> <i>Stopword</i>	31
Gambar 4.8	<i>Pseudocode</i> Penghitungan Frekuensi <i>Term</i>	32
Gambar 4.9	<i>Pseudocode</i> Algoritma Naïve Bayes.....	33
Gambar 4.10	Tabel <i>Feature</i> dalam <i>Database</i>	35
Gambar 4.11	Hasil Pengujian S5.....	36
Gambar 4.12	Grafik akurasi sistem.....	47

DAFTAR LAMPIRAN

Lampiran A : <i>Source Code</i>	Lampiran 1
Lampiran B : Tabel Hasil Pengujian.....	Lampiran 16
Lampiran C : Status Pelatihan.....	Lampiran

©UKDW

ABSTRAKSI

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAÏVE BAYES

Dalam media sosial, status sering digunakan untuk mencurahkan perasaan penggunanya, seperti emosi sehari-hari senang, sedih, marah, dsb. Namun belum ada layanan selanjutnya untuk mengolah status yang sudah ada. Jikalau status tersebut dapat ditentukan emosinya, layanan dapat ditingkatkan kedepannya, baik untuk kepentingan pengguna maupun dari layanan media sosial tersebut.

Dalam penelitian ini penulis menggunakan algoritma Naïve Bayes dalam melakukan klasifikasi status tersebut, menjadi tujuh kelas emosi. Status dalam media sosial berisi kalimat yang tidak baku sehingga status tersebut diolah terlebih dahulu menjadi dokumen status yang baku, setelah itu akan dilakukan tahap *preprocessing* yaitu *stopword* dan *tokenisasi* agar teks dapat terstruktur. Penelitian ini menggunakan status yang akan dilabeli secara manual sebagai acuan keakuratan sistem. Adapun penelitian ini menggunakan data ISEAR sebagai *feature set*, data ISEAR adalah data yang berisi kumpulan kalimat yang telah dikategorikan menjadi tujuh kelas emosi. Penelitian ini akan diuji keakuratan sistem dengan mengklasifikasi dokumen status uji dengan *feature selection* yang berbeda dan menggunakan data ISEAR maupun tidak.

Hasil dari sistem klasifikasi emosi berdasarkan status media sosial ini memang kurang baik, karena hasil akurasi tertinggi adalah 58,57% dengan *feature selection* 100% dan tanpa menggunakan data ISEAR. Hasil terendah adalah 44,28% dengan *feature selection* 10% dan menggunakan data ISEAR.

Kata kunci : klasifikasi, status, media sosial, Naïve Bayes

**PENDETEKSI KATEGORI EMOSI BERDASARKAN
STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA
NAIVE BAYES**

Skripsi



oleh
REVANY ARLITA
71110019

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2015

**PENDETEKSI KATEGORI EMOSI BERDASARKAN
STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA
NAIVE BAYES**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi
Informasi Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

REVANY ARLITA
71110019

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2015

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAIVE BAYES

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 15 Juni 2015



REVANY ARLITA

71110019

HALAMAN PERSETUJUAN

Judul Skripsi : PENDETEKSI KATEGORI EMOSI
BERDASARKAN STATUS MEDIA SOSIAL
MENERAPKAN ALGORITMA NAIVE BAYES

Nama Mahasiswa : REVANY ARLITA

N I M : 71110019

Matakuliah : Skripsi (Tugas Akhir)

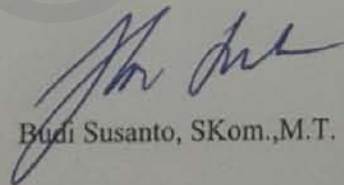
Kode : TIW276

Semester : Genap

Tahun Akademik : 2014/2015

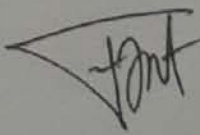
Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 15 Juni 2015

Dosen Pembimbing I



Budi Susanto, SKom.,M.T.

Dosen Pembimbing II



Antonius Rachmat C., S.Kom.,M.Cs.

HALAMAN PENGESAHAN

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAIVE BAYES


Oleh: REVANY ARLITA / 71110019

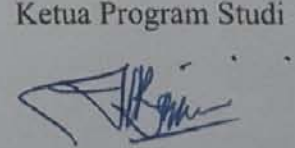
Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 12 Juni 2015

Yogyakarta, 15 Juni 2015
Mengesahkan,

Dewan Penguji:

1. Budi Susanto, SKom.,M.T.
2. Antonius Rachmat C., S.Kom.,M.Cs.
3. Gloria Virginia, S.Kom., MAI, Ph.D.
4. Aditya Wikan Mahastama, S.Kom


Dekan
(Budi Susanto, S.Kom., M.T.)


Ketua Program Studi
(Gloria Virginia, Ph.D.)

UCAPAN TERIMAKASIH

Puji syukur kepada Tuhan Yesus Kristus yang telah melimpahkan anugerah dan kasih karunianya, sehingga penulis dapat menyelesaikan tugas akhir dengan judul “Pendeteksi Kategori Emosi Berdasarkan Status Media Sosial Menerapkan Algoritma Naïve Bayes” ini dengan baik.

Penulis menyadari dalam penelitian ini masih banyak kekurangan, sehingga penulis menerima saran dan kritik dari berbagai pihak, secara langsung maupun tidak langsung. Selain itu, penulis juga menyampaikan ucapan terima kasih kepada :

1. Budi Susanto, SKom.,M.T. selaku dosen pembimbing serta Antonius Rachmat C, S.Kom., M.Cs. sebagai dosen pembimbing dan koordinator Tugas Akhir yang telah membimbing dan memberikan banyak masukan dalam menyelesaikan penelitian ini.
2. Keluarga dan sahabat terkasih yang selalu mendukung dan memberikan semangat kepada penulis selama pengerjaan penelitian ini.
3. Semua *volunteer* baik yang telah mengijinkan statusnya dipakai untuk bahan penelitian ini, maupun *volunteer* yang telah bersedia melakukan pelabelan dalam penelitian ini.
4. Semua pihak lain yang telah membantu dalam pengerjaan penelitian ini yang tidak dapat disebutkan satu per satu.

Akhir kata penulis mengucapkan permintaan maaf bila adanya kesalahan dalam penyusunan laporan ini. Terima kasih.

Yogyakarta, 19 Juni 2015

Penulis

ABSTRAKSI

PENDETEKSI KATEGORI EMOSI BERDASARKAN STATUS MEDIA SOSIAL MENERAPKAN ALGORITMA NAÏVE BAYES

Dalam media sosial, status sering digunakan untuk mencurahkan perasaan penggunanya, seperti emosi sehari-hari senang, sedih, marah, dsb. Namun belum ada layanan selanjutnya untuk mengolah status yang sudah ada. Jikalau status tersebut dapat ditentukan emosinya, layanan dapat ditingkatkan kedepannya, baik untuk kepentingan pengguna maupun dari layanan media sosial tersebut.

Dalam penelitian ini penulis menggunakan algoritma Naïve Bayes dalam melakukan klasifikasi status tersebut, menjadi tujuh kelas emosi. Status dalam media sosial berisi kalimat yang tidak baku sehingga status tersebut diolah terlebih dahulu menjadi dokumen status yang baku, setelah itu akan dilakukan tahap *preprocessing* yaitu *stopword* dan *tokenisasi* agar teks dapat terstruktur. Penelitian ini menggunakan status yang akan dilabeli secara manual sebagai acuan keakuratan sistem. Adapun penelitian ini menggunakan data ISEAR sebagai *feature set*, data ISEAR adalah data yang berisi kumpulan kalimat yang telah dikategorikan menjadi tujuh kelas emosi. Penelitian ini akan diuji keakuratan sistem dengan mengklasifikasi dokumen status uji dengan *feature selection* yang berbeda dan menggunakan data ISEAR maupun tidak.

Hasil dari sistem klasifikasi emosi berdasarkan status media sosial ini memang kurang baik, karena hasil akurasi tertinggi adalah 58,57% dengan *feature selection* 100% dan tanpa menggunakan data ISEAR. Hasil terendah adalah 44,28% dengan *feature selection* 10% dan menggunakan data ISEAR.

Kata kunci : klasifikasi, status, media sosial, Naïve Bayes

DAFTAR ISI

Halaman Judul.....	
Pernyataan Keaslian Skripsi.....	iii
Halaman Persetujuan.....	iv
Halaman Pengesahan.....	v
Ucapan Terimakasih.....	vi
Abstraksi.....	vii
Daftar Isi.....	viii
Daftar Tabel.....	xi
Daftar Gambar.....	xii
Daftar Lampiran.....	xiii
BAB 1 Pendahuluan.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Metode Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB 2 Tinjauan Pustaka.....	5
2.1 Tinjauan Pustaka.....	5
2.2 Landasan Teori.....	6
2.2.1 Facebook.....	6
2.2.2 Emosi.....	7
2.2.3 Teks Mining.....	8
2.2.4 Algoritma Naïve Bayes.....	9
BAB 3 Analisis dan Perancangan Sistem.....	11
3.1 Spesifikasi Sistem.....	11

3.1.1 Fungsional.....	11
3.1.2 Non Fungsional.....	12
3.2 <i>Usecase</i>	13
3.3 <i>Flowchart</i>	14
3.4 Arsitektur Sistem.....	21
3.5 Perancangan Basis Data.....	21
3.6 Perancangan <i>User Interface</i>	23
3.7 Perancangan Pengujian Sistem.....	26
BAB 4 Implementasi dan Analisis Sistem.....	27
4.1 Implementasi Sistem.....	27
4.1.1 Konfigurasi Awal.....	27
4.1.2 Antarmuka Sistem.....	28
4.1.3 <i>Pseudocode</i> dan Kode Sistem.....	31
4.1.3.1 <i>Stopword</i>	31
4.1.3.2 Penghitungan Frekuensi <i>Term</i>	32
4.1.3.3 Algoritma Naïve Bayes.....	33
4.2 Analisis Sistem.....	34
4.2.1 Analisis Secara Manual.....	34
4.2.2 Evaluasi Sistem.....	37
4.2.2.1 <i>Feature Selection</i> 10% tanpa data ISEAR.....	37
4.2.2.2 <i>Feature Selection</i> 20% tanpa data ISEAR.....	38
4.2.2.3 <i>Feature Selection</i> 40% tanpa data ISEAR.....	39
4.2.2.4 <i>Feature Selection</i> 60% tanpa data ISEAR.....	40
4.2.2.5 <i>Feature Selection</i> 100% tanpa data ISEAR.....	41
4.2.2.6 <i>Feature Selection</i> 10% dengan data ISEAR.....	42
4.2.2.7 <i>Feature Selection</i> 20% dengan data ISEAR.....	43
4.2.2.8 <i>Feature Selection</i> 40% dengan data ISEAR.....	44
4.2.2.9 <i>Feature Selection</i> 60% dengan data ISEAR.....	45
4.2.2.10 <i>Feature Selection</i> 100% dengan data ISEAR.....	46

4.2.2.11 Hasil Evaluasi.....	46
BAB 5 Kesimpulan dan Saran.....	50
5.1 Kesimpulan.....	50
5.2 Saran.....	50
Daftar Pustaka.....	
Lampiran.....	

©UKDWN

DAFTAR TABEL

Tabel 2.1	Contoh Kalimat Emosi.....	8
Tabel 2.2	Pemetaan Dokumen Terambil atau Relevan.....	11
Tabel 4.1	Daftar Contoh Token.....	35
Tabel 4.2	Hasil pengujian FS 10% tanpa ISEAR.....	37
Tabel 4.3	Hasil pengujian FS 20% tanpa ISEAR.....	38
Tabel 4.4	Hasil pengujian FS 40% tanpa ISEAR.....	39
Tabel 4.5	Hasil pengujian FS 60% tanpa ISEAR.....	40
Tabel 4.6	Hasil pengujian FS 100% tanpa ISEAR.....	41
Tabel 4.7	Hasil pengujian FS 10% dengan ISEAR.....	42
Tabel 4.8	Hasil pengujian FS 20% dengan ISEAR.....	43
Tabel 4.9	Hasil pengujian FS 40% dengan ISEAR.....	44
Tabel 4.10	Hasil pengujian FS 60% dengan ISEAR.....	45
Tabel 4.11	Hasil pengujian FS 100% dengan ISEAR.....	46
Tabel 4.12	Akurasi Sistem.....	47

DAFTAR GAMBAR

Gambar 3.1	Diagram <i>use-case</i>	13
Gambar 3.2	<i>Flowchart Preprocessing</i> ISEAR.....	15
Gambar 3.3	<i>Flowchart</i> Sistem.....	16
Gambar 3.4	<i>Flowchart</i> Tambah Status.....	17
Gambar 3.5	<i>Flowchart Preprocessing</i> Status	18
Gambar 3.6	<i>Flowchart</i> Pembentukan <i>feature</i>	19
Gambar 3.7	<i>Flowchart</i> klasifikasi Status.....	20
Gambar 3.8	Arsitektur Sistem.....	21
Gambar 3.9	Skema Diagram <i>Database</i>	22
Gambar 3.10	Rancangan Tampilan Klasifikasi.....	24
Gambar 3.11	Rancangan Tampilan Hasil.....	24
Gambar 3.12	Rancangan Tampilan Lihat Status.....	25
Gambar 3.13	Rancangan Tampilan Tambah Status.....	25
Gambar 4.1	<i>Bootstrap</i> yang Dipasang.....	27
Gambar 4.2	<i>Bootstrap</i> yang Dipakai dalam <i>Source Code</i>	28
Gambar 4.3	Halaman Klasifikasi.....	28
Gambar 4.4	Halaman Hasil.....	29
Gambar 4.5	Halaman Lihat Status.....	30
Gambar 4.6	Halaman Tambah Status.....	30
Gambar 4.7	<i>Pseudocode</i> <i>Stopword</i>	31
Gambar 4.8	<i>Pseudocode</i> Penghitungan Frekuensi <i>Term</i>	32
Gambar 4.9	<i>Pseudocode</i> Algoritma Naïve Bayes.....	33
Gambar 4.10	Tabel <i>Feature</i> dalam <i>Database</i>	35
Gambar 4.11	Hasil Pengujian S5.....	36
Gambar 4.12	Grafik akurasi sistem.....	47

DAFTAR LAMPIRAN

Lampiran A : <i>Source Code</i>	Lampiran 1
Lampiran B : Tabel Hasil Pengujian.....	Lampiran 16
Lampiran C : Status Pelatihan.....	Lampiran

©UKDWN

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Media sosial adalah sebuah kelompok aplikasi berbasis Internet yang membangun di atas dasar ideologi dan teknologi Web 2.0, dan yang memungkinkan penciptaan dan pertukaran *user-generated content* (Sindang, 2013). Generasi sekarang ini menggunakan media sosial sebagai alat berkomunikasi. Tak jarang para pengguna media sosial juga mengungkapkan isi hatinya ke media sosial, yang sering disebut *update* status, misalnya ketika mereka sedang sedih hingga marah. Status dari pengguna media sosial ini mencerminkan perasaan pemilik akun tersebut, dengan berbagai alasan pengguna media sosial mengunggah status tersebut.

Semakin harinya belum ada kelanjutan tentang status yang dibuat oleh pengguna media sosial tersebut. Tidak jarang ada pengguna media sosial yang terbiasa atau bahkan membuat status yang lebih buruk, sehingga merugikan dirinya sendiri. Jika dari sistem dapat mengenali emosi yang sedang dialami oleh pengguna, maka bisa lebih berguna untuk meningkatkan layanan media sosial dan kebaikan untuk pengguna media sosial tersebut. Sebagai contoh sistem dapat memberikan pesan khusus yang berguna kepada pengguna sesuai dengan emosi yang dialaminya.

Oleh sebab itu, penulis tertarik untuk membangun sistem yang dapat mengklasifikasi status media sosial sehingga dapat mendeteksi emosi dari pengguna media sosial tersebut, menggunakan algoritma klasifikasi Naïve Bayes. Dengan mengetahui kondisi emosi dari pengunggah status, diharapkan dapat berguna untuk layanan media sosial dan bagi pemilik akun tersebut.

1.2 Perumusan Masalah

- Seberapa tepat algoritma Naïve Bayes dalam melakukan klasifikasi kategori emosi berdasarkan status-status seseorang di layanan jejaring media sosial?
- Bagaimana perbandingan hasil klasifikasi emosi menggunakan data ISEAR dengan tanpa menggunakan data ISEAR?

1.3 Batasan Masalah

Sistem yang dibangun akan memiliki beberapa batasan antara lain :

- Status akan diklasifikasi menjadi tujuh kategori emosi yaitu marah, jijik, takut, senang, sedih, malu, dan bersalah. Kategori ini didapat dari dataset ISEAR(*International Survey on Emotion Antecedents and Reaction*) yang akan diterjemahkan menjadi bahasa Indonesia, sekaligus dipakai menjadi *feature set* dalam proses klasifikasinya. Diambil dari (<http://emotion-research.net/toolbox/toolboxdatabase.2006-10-13.2581092615>). Status yang diambil akan dilakukan pelabelan menjadi tujuh kategori oleh beberapa *volunteer*.
- Status yang akan diklasifikasi tidak diambil dari semua media sosial tapi hanya diambil dari salah satu media sosial saja yaitu jejaring sosial *facebook* (FB).
- Status diambil dari pengguna *facebook* (FB), status yang diambil berjumlah 420 status, terdiri dari 70 status tiap kelasnya.
- Status yang diambil akan diklasifikasi dalam bentuk teks.
- Hanya terbatas dengan bahasa Indonesia saja, dan status yang bukan merupakan sebuah kutipan seperti dari sebuah lagu atau ayat alkitab dan bukan mengandung bahasa selain bahasa Indonesia (bahasa daerah, bahasa asing).

1.4 Tujuan Penelitian

Tujuan utama dari penelitian ini adalah menerapkan algoritma klasifikasi Naïve Bayes untuk mendeteksi emosi berdasarkan status media sosial, yang dideteksi menjadi beberapa klasifikasi emosi. Selain itu penelitian ini bertujuan untuk mengetahui ketepatan algoritma Naïve Bayes dalam melakukan klasifikasi kategori emosi berdasarkan status-status seseorang di layanan jejaring media sosial. Penelitian ini juga ingin menemukan besar(banyaknya) *feature set* untuk membangun model probabilitas dalam algoritma tersebut.

1.5 Metode Penelitian

Tahapan penelitian yang dilakukan adalah sebagai berikut :

1. Metode observasi langsung

Peneliti akan mengobservasi pengguna media sosial, seperti status-status yang ada di media sosial tersebut, yang memenuhi syarat untuk dipakai sebagai bahan penelitian seperti yang tercantum dalam batasan masalah.

2. Pengumpulan data

Dalam tahap ini akan dilakukan pengambilan data yaitu status-status *facebook*. Lalu akan dilakukan metode kuesioner yang ditujukan pada lima orang *volunteer*, baik dari usia muda hingga dewasa, hal ini dilakukan untuk melakukan pelabelan emosi status yang akan dipakai sebagai *training data set*.

3. Membangun sistem

Tahap ini adalah tahap pembuatan sistem yang telah dirancang, dengan menuliskan *source code*, membuat *feature set* sesuai kebutuhan, mencari kesalahan dalam sistem, dan memperbaiki sistem hingga dapat berjalan sesuai dengan rancangan awal.

4. *Testing*

Sistem yang telah dibangun akan dilakukan uji coba dengan menggunakan status uji(diambil 70 status dari seluruh status yang telah dikumpulkan) yang terdiri dari 10 status tiap kelasnya.. Data yang masuk akan dilakukan *preprocessing* dan proses klasifikasi.

5. Evaluasi klasifikasi

Tahapan ini dilakukan penghitungan untuk mengevaluasi hasil klasifikasi, dengan menghitung nilai akurasi. Hal ini dilakukan untuk mengetahui tingkat ketepatan hasil klasifikasi emosi pada status tersebut.

1.6 Sistematika Penulisan

Penjelasan tentang kasus penelitian yang akan dilakukan akan berada di bab pertama, bab dua dan bab tiga. Bab pertama adalah bab pendahuluan yang berisi penjelasan tentang gambaran penelitian yang akan dilakukan, seperti alasan melakukan penelitian ini, pentingnya dari penelitian, serta permasalahan yang akan diselesaikan dengan penelitian ini. Bab kedua merupakan bab tinjauan pustaka, menjelaskan tentang uraian teori yang dipakai sebagai sumber pustaka, dan menjelaskan persamaan, model matematis yang berkaitan dengan penelitian. Bab tiga yaitu bab analisis dan perancangan sistem, mencakup analisis teori yang digunakan dan bagaimana menterjemahkannya ke dalam suatu sistem yang hendak dibuat.

Bab empat dan bab lima akan menjelaskan tentang hasil dari penelitian yang akan dilakukan. Bab empat adalah implementasi dan analisis sistem, yang berisi hasil riset/implementasi, dan pembahasan/analisis riset tersebut. Terakhir, bab lima yakni bab kesimpulan dan saran, berisi kesimpulan dari hasil penelitian ini dan saran untuk memperbaiki sistem yang dapat dilakukan di penelitian berikutnya.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil implementasi dan analisis sistem yang telah dibahas di bab sebelumnya, maka dapat disimpulkan sebagai berikut :

1. Dalam penelitian ini hasil terbaik adalah 58,57 %, sedangkan hasil tersebut kurang optimal. Hal ini dikarenakan banyak faktor antara lain tahap *preprocessing*, data ISEAR yang dipakai, status yang memiliki arti yang ambigu. Sehingga hasil klasifikasi menggunakan algoritma Naïve Bayes dalam penelitian ini kurang baik.
2. Sistem klasifikasi emosi berdasarkan status media sosial menerapkan algoritma Naïve Bayes menghasilkan nilai akurasi tertinggi sebesar 58,57 % untuk *feature selection* 100% tanpa menggunakan data ISEAR. Sehingga semakin tinggi *feature selection*, semakin tinggi akurasinya.
3. Data ISEAR kurang berpengaruh dalam penelitian ini saat menentukan hasil klasifikasi dengan dokumen non formal seperti status media sosial, hal ini didapat dari hasil evaluasi menggunakan data ISEAR nilai akurasi tertinggi hanya sebesar 50 %.

5.2 Saran

Saran yang diberikan untuk perbaikan dari sistem ini adalah sebagai berikut :

1. Membuat kurva ROC untuk mengevaluasi lebih lagi sistem klasifikasi dalam penelitian ini.

2. Diperlukan penambahan kata dalam kamus tidak baku, berupa singkatan dan kata yang sering dipakai dalam bahasa dalam status media sosial.
3. Menambahkan *stemming* dalam tahap *preprocessing* sehingga pencocokan frekuensi kemunculan kata dapat baik.
4. Melakukan klasifikasi dengan menerapkan tahap pembobotan terlebih dahulu tiap kelas, sehingga hasil perhitungan probabilitas menjadi lebih baik.

©UKDW

DAFTAR PUSTAKA

- Agustina, P. A. (2014). Klasifikasi Trending Topic Twitter dengan Penerapan Metode Naïve Bayes. *Jurnal UMRAH* .
- Danisman, T., & Alpkocak, A. (2008). Feeler: Emotion Classification of Text Using Vector Space Model. *AISB Convention* , 1.
- Destuardi, I., & Sumpeno, S. (2009). Klasifikasi Emosi untuk Teks Bahasa Indonesia Menggunakan Metode Naive. *Institut Teknologi Sepuluh* .
- Limpo, E. U., Wibowo, A., & Lim, R. (2014). *Pembuatan Jaringan Sosial Peneliti Berbasis Facebook Yang Memanfaatkan Situs Sitasi Artikel Ilmiah Di Pusat Penelitian Universitas Kristen Petra*. Surabaya: Universitas Kristen Petra.
- Manning, C. D. (2008). *Introduction To Information Retrieval*. Cambridge: Cambridge University Press.
- Nilamsari, Maharani, W., & Sulisty, M. D. (2013). Klasifikasi Emosi pada Opini Teks Bahasa Indonesia Menggunakan Vector Space Model.
- Samodra, J., Sumpeno, S., & Hariadi, M. (2009). Klasifikasi Dokumen Teks Berbahasa Indonesia dengan Menggunakan Naïve Bayes. *Seminar Nasional Electrical, Informatics, and It's Educations* .
- Shimodaira, H. (2014). Text Classification using Naive Bayes. *Informatics 2B* .
- Sindang, E. (2013). Manfaat Media Sosial Dalam Ranah Pendidikan dan Pelatihan. *Pusdiklat KNPk* .
- Utami, D. (2010). *Karakteristik Penggunaan Bahasa pada Status Facebook*. Surakarta: Universitas Sebelas Maret.