

**Analisis Sentimen Masyarakat
Tentang *Provider* Telekomunikasi Indonesia
Pada Twitter Menggunakan *Naive Bayes
Classifier***

TUGAS AKHIR

**Diajukan untuk Memenuhi Sebagian Persyaratan
Mencapai Derajat Sarjana Teknik Informatika**



Disusun Oleh:

Stefani Aprilia

14 07 07745

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS ATMA JAYA YOGYAKARTA

2018

HALAMAN PENGESAHAN
TUGAS AKHIR BERJUDUL
ANALISIS SENTIMEN MASYARAKAT
TENTANG PROVIDER TELEKOMUNIKASI INDONESIA
PADA TWITTER MENGGUNAKAN NAÏVE BAYES CLASSIFIER

Disusun Oleh:

Stefani Aprilia

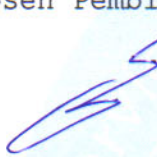
14 07 07745

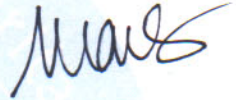
Dinyatakan telah memenuhi syarat

Pada tanggal: Mei 2018

Dosen Pembimbing I

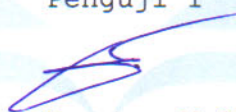
Dosen Pembimbing II


(Prof. Ir. Suyoto, M.Sc., Ph.D)



(Martinus Maslim, S.T., M.T.)

Tim Penguji

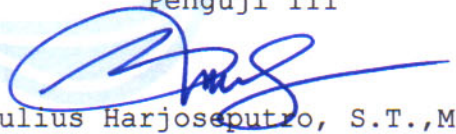
Penguji I


(Prof. Ir. Suyoto, M.Sc., Ph.D)

Penguji II


(Patricia Ardanari, S.Si., M.T.)
Yogyakarta,

Penguji III


(Yulius Harjoseputro, S.T., M.T.)
Mei 2018


PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS ATMA JAYA YOGYAKARTA



Dekan,


(Dr. A. Teguh Siswanto M.Sc.)

HALAMAN PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama: Stefani Aprilia

NPM: 140707745

Saya dengan ini menyatakan bahwa tugas akhir yang berjudul "Analisis Sentimen Masyarakat tentang *Provider Telekomunikasi Indonesia* Pada Twitter Menggunakan *Naïve Bayes Classifier*" merupakan hasil penelitian yang saya lakukan pada Tahun Akademik 2017/2018 yang merupakan penelitian asli dan tidak merupakan tindakan *plagiasi* dari karya manapun.

Jika dikemudian hari telah ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang sedang berlaku untuk mencabut gelar sarjana yang telah diberikan Universitas Atma Jaya Yogyakarta.

Demikian pernyataan ini dibuat sebagaimana mestinya dan dengan sebenar-benarnya.

Yogyakarta, 12 April 2018

Yang menyatakan

Stefani Aprilia

HALAMAN PERSEMBAHAN

**Tugas Akhir ini dipersembahkan kepada:
Tuhan Yesus Kristus, Keluarga, Teman dan
Sahabatku**



KATA PENGANTAR

Puji Syukur dan terima kasih penulis panjatkan kepada Tuhan Yesus Kristus atas segala berkat dan anugerah-Nya sehingga penulis dapat menyelesaikan pembuatan tugas akhir ini dengan baik dan tepat waktu. Tugas akhir ini bertujuan untu memenuhi salah satu syarat untuk mencapai gelar sarjana di Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta.

Tugas Akhir ini dapat diselesaikan dengan melibatkan banyak pihak yang membantu. Oleh sebab itu, penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yesus Kristus yang selalu melindungi dan memberkati penulis dalam menyelesaikan tugas akhir.
2. Bapak Prof. Ir. Suyoto, M.Sc., Ph.D selaku Dosen Pembimbing I yang telah membimbing, membantu dan memberikan motivasi dalam menyelesaikan tugas akhir ini.
3. Bapak Martinus Maslim, S.T., M.T. selaku Dosen Pembimbing II yang telah membimbing, membantu dan memberikan motivasi dalam menyelesaikan tugas akhir ini.
4. Segenap dosen dan karyawan Universitas Atma Jaya Yogyakarta, yang telah membimbing, mengajar selama penulis melakukan *studi* di Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta.

5. Keluarga tercinta, ibuk, bapak, mas tuga, dek grace, kakek dan seluruh bagian dari keluarga besar yang selalu memberikan motivasi, dukungan dan doa selama penulis menempuh pendidikan.
6. Rezky Owend Pratama yang selalu marah jika tidak mengerjakan skripsi, selalu jadi alarm kalau mau ketemu dosen, motivasi yang lucu dan semangat.
7. Seluruh teman-teman tercinta yang selalu memberikan dukungan, bantuan, dan menemani penulis dalam segala keadaan dan suasana. Terhusus. Teman-teman seperjuangan yang selalu direpotkan untuk menemani begadang anak kelas B 2014, Aya, Devi dan kak Anthy. Teman kost tercinta Keke, Hana, Eliz, Della, Tessa dan Ira.
8. Semua orang yang tidak dapat penulis sebutkan satu persatu yang telah membantu dalam segala hal dan memberikan semangat yang berarti.

DAFTAR ISI

HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
ABSTRAK	xi
BAB 1	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	5
1.5. Metode Penelitian	5
1.6. Sistem Penulisan Tugas Akhir	9
BAB 2	10
TINJAUAN PUSTAKA	10
2.1. Tinjauan Pustaka	10

BAB 3	19
LANDASAN TEORI	19
3.1. Analisis Sentimen	20
3.2. Jejaring Sosial Twitter	21
3.3. Metode <i>Naïve Bayes Classifier</i>	22
3.4. Data Mining	24
BAB 4	28
METODOLOGI PENELITIAN	28
4.1. Pengumpulan Data	28
4.1.1. Kamus Kata Positif	29
4.1.2. Kamus Kata Negatif	29
4.1.3. Data Latih dan Data Uji	30
4.1.4. Data Asli	31
4.2. Pembuatan Model Analisis Sentimen	34
4.2.1. <i>Naïve Bayes Classifier</i>	34
4.2.1.1. Tokenisasi	36
4.2.1.2. <i>Stopword</i>	39
4.2.1.3. <i>Stemming</i>	40
4.2.1.4. Labeling dalam Excel	41
4.2.1.5. Penerapan <i>Naïve Bayes Classifier</i>	42
4.2.2. WEKA	46
4.3. Menguji dan Membandingkan Hasil Sentimen	46
4.4. Analisis Sentimen Data Asli	48
BAB 5	50
HASIL PENELITIAN DAN PEMBAHASAN	50
5.1. Metodologi Penelitian	50
BAB 6	57

KESIMPULAN DAN SARAN	57
6.1. Kesimpulan	57
6.2. Saran	58

DAFTAR GAMBAR

Gambar 4.1. Fungsi Pengambilan Data	31
Gambar 4.2. Fungsi Tokenisasi	36
Gambar 4.3. Fungsi <i>Stopword</i>	38
Gambar 4.4. Fungsi Stemming	39
Gambar 5.1. Contoh Data Twitter	51
Gambar 5.2. Kamus Kata Positif	52
Gambar 5.3. Labeling dengan Excel	53
Gambar 5.4. Jumlah Data	53
Gambar 5.5. Data Training 25%	55
Gambar 5.6. Data Training 50%	56
Gambar 5.7. Data Training 60%	56
Gambar 5.8. Data Training 75%	57
Gambar 5.9. Data Training 80%	57

DAFTAR TABEL

TABEL 2.1. Perbandingan dengan Penelitian Sebelumnya..	16
TABEL 4.1. Contoh Data Tabel Latih.....	42
TABEL 4.2. Contoh Tabel Uji.....	42
TABEL 4.3. Ilustrasi <i>Confusion Matrix</i>	46
TABEL 5.1. Tabel Hasil Pengujian dengan Weka.....	54

**Analisis Sentimen Masyarakat
Tentang *Provider* Telekomunikasi Indonesia
Pada Twitter Menggunakan *Naïve Bayes Classifier***

Disusun Oleh:

Stefani Aprilia/140707745

Abstrak

Perkembangan media sosial menjadikan media sosial sebagai salah satu media pengambilan data yang besar. Salah satu pemanfaatan dari data media sosial adalah untuk mengetahui pendapat atau sentimen pengguna media sosial terhadap suatu topik dan masalah. Topik penting yang dibicarakan oleh pengguna jejaring sosial adalah *provider* telekomunikasi. Masyarakat dapat mengungkapkan pendapat, pernyataan, maupun tanggapan tentang *provider* telekomunikasi melalui internet terkhusus jejaring sosial. Pendapat masyarakat melalui Twitter disampaikan dengan cara *posting tweets*. *Postingan tweets* tersebut dapat dilakukan analisis sentimen terhadap topik *provider* telekomunikasi.

Analisis sentimen adalah salah satu cabang penelitian dari *Text Mining* yang berguna untuk mengklasifikasikan dokumen teks berupa opini berdasarkan sentimen. Pendapat tentang *provider* telekomunikasi dikumpulkan dan dijadikan sebagai data untuk dianalisis dan diamati. Salah satu pembelajaran mesin untuk analisis sentimen adalah *Naïve Bayes Classifier* (NBC). Metode *Naïve Bayes Classifier* (NBC) terkenal dengan metode yang sederhana tetapi memiliki keakuratan yang tinggi dalam kasus analisis sentimen berupa dokumen dibandingkan SVM.

Hasil analisis sentimen dengan metode *Naïve Bayes Classifier* (NBC) akan mendapatkan informasi dan pengetahuan yang dapat dijadikan pedoman dan masukan kepada pihak-pihak terkait untuk dapat meningkatkan layanan dan mempertahankan

kepercayaan masyarakat terhadap *provider* telekomunikasi di Indonesia terutama Telkomsel.

Kata Kunci: Analisis Sentimen, Twitter, *Provider* Telekomunikasi, *Naïve Bayes Classifier* (NBC), Data Mining

