

**IMPLEMENTASI METODE NAÏVE BAYES CLASSIFIER UNTUK
MENGKATEGORIKAN PESAN SMS BERBASIS SMS GATEWAY
STUDI KASUS SMK KOPERASI**

SKRIPSI



Oleh

Jefri Sugiarto

23 08 0377

Program Studi Sistem Informasi Fakultas Teknologi Informasi

Universitas Kristen Duta Wacana

2012

**IMPLEMENTASI METODE NAÏVE BAYES CLASSIFIER UNTUK
MENGKATEGORIKAN PESAN SMS BERBASIS SMS GATEWAY
STUDI KASUS SMK KOPERASI**

SKRIPSI



Ditujukan kepada Fakultas Teknik Sistem Informasi
Universitas Kristen Duta Wacana

Sebagai salah satu syarat dalam memperoleh gelar
Sarjana Komputer



Disusun Oleh :

Jefri Sugiarto

23 08 0377

**Program Studi Sistem Informasi
Universitas Kristen Duta Wacana**

2012

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa tugas akhir dengan judul:

IMPLEMENTASI METODE NAÏVE BAYES CLASSIFIER UNTUK MENKATEGORIKAN PESAN SMS BERBASIS SMS GATEWAY STUDI KASUS SMK KOPERASI

Yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan sarjana Program Studi Teknik Sistem Informasi, 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 bagain 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 yaitu pencabutan gelar kesarjanaan saya.



Yogyakarta, 10 Mei 2012

A handwritten signature in black ink, appearing to read 'Jefri Sugiarto', is written over the printed name.

(Jefri Sugiarto)

23 08 0377

HALAMAN PERSETUJUAN

Judul : Implementasi Metode Naïve Bayes Classifier untuk
mengategorikan pesan SMS berbasis SMS Gateway
Studi Kasus SMK Koperasi.

Nama : Jefri Sugiarto

Nim : 23080377

Mata Kuliah : Tugas Akhir

Kode : IN4036

Semester : Genap

Tahun Akademik : 2011/2012

Telah diperiksa dan disetujui
Di Yogyakarta,
Tanggal, 16 Mei 2012

Dosen Pembimbing I

(Erick Kurniawan, S.Kom., M.Kom.)

Dosen Pembimbing II

(Yetli Oslan, S.Kom, M.T.)

HALAMAN PENGESAHAN

SKRIPSI

**IMPLEMENTASI METODE NAÏVE BAYES CLASSIFIER UNTUK
MENGKATEGORIKAN PESAN SMS BERBASIS SMS GATEWAY
STUDI KASUS SMK KOPERASI**

Oleh : Jefri Sugiarto / 23080377

Dipertahankan di depan dewan Penguji Tugas Akhir/Skripsi

Program Studi Sistem Informasi Fakultas Teknik

Universitas Kristen Duta Wacana – Yogyakarta

Dan dinyatakan diterima untuk memenuhi salah satu

Syarat memperoleh gelar

Sarjana Komputer

Pada tanggal

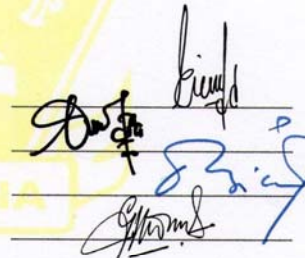
28 Mei 2012

Yogyakarta, 31 Mei 2012

Mengesahkan,

Dengan Penguji:

1. Erick Kurniawan, S.Kom., M.Kom
2. Yetli Oslan, S.Kom, M.T
3. Drs Wimmie Handiwidjojo., MIT
4. Drs.R.Gunawan Santosa., M.Si



Dekan

Ketua Program Studi



(Drs. Wimmie Handiwidjojo., MIT)

(Yetli Oslan, S.Kom, M.T)

UCAPAN TERIMA KASIH

Puji dan syukur kepada Tuhan Yesus Kristus karena kasih dan rahmat yang diberikan, sehingga penulis dapat menyelesaikan skripsi dengan judul Implementasi metode naïve bayes dalam mengkategorikan pesan SMS berbasis sms gateway studi kasus SMK Koperasi dengan baik dan tepat waktu.

Laporan hasil penelitian ini merupakan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer. Tujuan lainnya adalah melatih mahasiswa agar dapat menghasilkan sebuah karya yang dapat dipertanggungjawabkan secara ilmiah, dan diharapkan dapat berguna bagi penggunanya.

Dalam menyelesaikan program skripsi ini, tidak lepas dari bantuan serta dukungan dari berbagai pihak, baik secara langsung maupun tidak langsung. Untuk itu dengan segala kerendahan hati, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Tuhan Yesus Kristus yang telah menganugrahkan karunia-Nya, sehingga skripsi ini dapat terselesaikan.
2. Erick Kurniawan, S.Kom., M.Kom. selaku dosen pembimbing I yang telah memberikan bimbingan dengan sabar dan baik kepada penulis, ucapan terima kasih juga penulis sampaikan kepada
3. Yetli Oslan, S.Kom., M.T. selaku dosen pembimbing II yang telah memberikan bimbingan, petunjuk dan masukan yang diberikan selama pengerjaan skripsi ini.
4. Keluarga tercinta yang selalu memberikan dorongan serta semangat.
5. Teman-teman yang telah memberikan masukan, dukungan doa serta semangat.
6. Pihak lain yang tidak dapat penulis sebutkan satu per satu, sehingga skripsi ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa program dan laporan skripsi ini masih sangat jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari pembaca sekalian.

Akhir kata penulis meminta maaf jika dalam penulisan skripsi ini, penulis melakukan kesalahan ataupun hal-hal yang kurang baik. Dan semoga ini dapat berguna bagi kita semua.

Yogyakarta, 14 Mei 2012

Penulis

© UKDW

INTISARI

Implementasi Metode Naïve Bayes Classifier untuk mengkategorikan pesan SMS berbasis SMS Gateway studi Kasus SMK Koperasi

Mengidentifikasi isi pesan SMS berdasarkan kategori merupakan hal yang tidak mudah, apa lagi pesan tersebut tidak memiliki struktur yang jelas. Oleh sebab itu diperlukan teknik khusus untuk mengidentifikasi isi pesan tersebut, sehingga pesan mudah untuk dikategorikan berdasarkan ciri atau pola tertentu.

Untuk mengkategorikan pesan SMS tersebut, digunakan Metode Naïve Bayes yang merupakan metode klasifikasi dokumen dengan cara kerja menghitung probabilitas kata yang muncul terhadap data training. Proses yang dilakukan pertama oleh metode ini adalah mengolah data training dan menghitung probabilitas katanya di masing-masing kategori, kata-kata di setiap kategori yang telah ditentukan sebelumnya akan menjadi ciri-ciri dari masing-masing kategori. Untuk mempermudah dan mempercepat proses pengkategorian pesan, digunakan sistem informasi yang menggunakan metode Naïve Bayes Classifier dalam mengkategorikan pesan.

Dari hasil penelitian yang dilakukan dapat disimpulkan bahwa metode Naïve Bayes Classifier dapat memiliki akurasi atau ketepatan dalam mengkategorikan pesan sebesar 73%. Hal itu dipengaruhi oleh pemilihan fitur-fitur di masing-masing kategori dan jumlah pesan yang digunakan di masing-masing kategori. Sebagai saran untuk pengembangan penelitian, sebaiknya dilakukan pembobotan fitur di masing-masing kelas, hal ini akan membedakan mana fitur yang penting dan tidak begitu penting di masing-masing kategori.

DAFTAR ISI

HALAMAN JUDUL.....	
PERNYATAAN KEASLIAN SKRIPSI.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
UCAPAN TERIMA KASIH	iv
INTISARI	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
BAB 1 PENDAHULUAN	1
1.1 Latar belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Metode Pendekatan	3
1.6 Sistematika Penulisan	4
BAB 2 TINJAUAN PUSTAKA	5
2.1 Tinjauan Pustaka	5
2.2 Landasan Teori	9
BAB 3 ANALISIS DAN PERANCANGAN SISTEM	18
3.1 Data Yang digunakan dalam pembuatan sistem	18
3.2 Algoritma Pemrograman	19
3.3 Algoritma Metode Naïve Bayes Classifier	20
3.4 Rancangan Proses	24
3.5 Rancangan Database	33
3.6 Rancangan Masukan Sistem	47
3.7 Rancangan Keluaran Sistem	52

BAB 4 IMPLEMENTASI DAN ANALISIS SISTEM	54
4.1 Implementasi Sistem	54
4.2 Analisis Sistem	60
BAB 5 KESIMPULAN DAN SARAN	62
5.1 Kesimpulan	62
5.2 Saran	66

© UKDW

DAFTAR TABEL

Tabel 2.1	Tarif SMS beberapa Operator Selular 2012	8
Tabel 2.2	Hasil Eksperimen Yudi Wibisono pada klasifikasi dokumen berbahasa Indonesia	15
Tabel 2.3	Hasil Eksperimen Fatimah Wulandini dan Anto Satriyo Nugroho	15
Tabel 3.1	Daftar Pesan Kategori dan Fitur Kemunculanya	21
Tabel 3.2	Nilai Probabilitas Data Training Tiap Kata Setiap Kategori 1 ...	22
Tabel 3.3	(Sambungan)	22
Tabel 3.4	(Sambungan)	22
Tabel 3.5	Tabel Entitas Pesan	41
Tabel 3.6	Tabel Entitas Tok_pesan	41
Tabel 3.7	Tabel Entitas Pesan Kategori	42
Tabel 3.8	Tabel Entitas Kategori	42
Tabel 3.9	Tabel Entitas Training	42
Tabel 3.10	Tabel Entitas Token_training	43
Tabel 3.11	Tabel Entitas Prob_training	43
Tabel 3.12	Tabel Entitas Hitung_prob	44
Tabel 3.13	Tabel Entitas Kata Singkatan	44
Tabel 3.14	Tabel Entitas Kata Stopwords	45
Tabel 3.15	Tabel Entitas Admin	45
Tabel 3.16	Tabel Entitas Kata Unik	45
Tabel 4.1	Hasil Perhitungan akurasi pengkategorian terhadap sampel data dan jumlah dokumen	60

DAFTAR GAMBAR

Gambar 2.1	Tahapan dalam Text Mining	5
Gambar 2.2	Ilustrasi Topologi SMS Gateway	7
Gambar 2.3	Cara Kerja Program PHP menggunakan web server.....	11
Gambar 2.4	Simbol Proses	16
Gambar 2.5	Simbol Masukan/ Keluaran	16
Gambar 2.6	Simbol Keputusan	16
Gambar 2.7	Simbol Garis Alur	17
Gambar 2.8	Simbol Penghubung antar halaman	17
Gambar 2.9	Simbol penghubung suatu halaman	17
Gambar 2.10	Gambar simbol titik terminal	17
Gambar 3.1	Topologi SMS Gateway	19
Gambar 3.2	Diagram Konteks	27
Gambar 3.3	Data Flow Diagram Level 0	28
Gambar 3.4	DFD Level 1 Proses 1.....	28
Gambar 3.5	DFD Level 2 Proses 2	29
Gambar 3.6	DFD Level 2 Proses 3	30
Gambar 3.7	Flowchart Proses Data Training	31
Gambar 3.8	Flowchart Proses Mengkategorikan Pesan	32
Gambar 3.9	Proses MDL 1	33
Gambar 3.10	Proses MDL 2.....	35
Gambar 3.11	Proses MDL 3	37
Gambar 3.12	Proses MDL 4.....	38
Gambar 3.13	Proses MDL 6	40
Gambar 3.14	Rancangan relasi antar tabel Sistem Informasi Gateway	46
Gambar 3.15	Tampilan Rancangan Awal Login Administrator	47
Gambar 3.16	Rancangan Awal Website	47
Gambar 3.17	Rancangan tampilan Tulis Pesan	48

Gambar 3.18	Rancangan tampilan Pesan Masuk	49
Gambar 3.19	Rancangan tampilan Daftar kotak keluar	49
Gambar 3.20	Rancangan tampilan Kata Singkatan	50
Gambar 3.21	Rancangan tampilan Menu Stopwords	50
Gambar 3.22	Rancangan tampilan Kata Kategori	51
Gambar 3.23	Tampilan Rancangan Data Training	51
Gambar 3.24	Rancangan tampilan Laporan Pesan Masuk	52
Gambar 3.25	Rancangan tampilan Laporan Kotak Keluar	52
Gambar 3.26	Rancangan tampilan Laporan Probabilitas Training	53
Gambar 4.1	Perintah Koneksi Ke Database	54
Gambar 4.2	Tampilan Program Untuk Menghilangkan Simbol	54
Gambar 4.3	Perintah Menghilangkan Stopwords	55
Gambar 4.4	Tampilan Program Menghitung Probabilitas	55
Gambar 4.5	Tampilan Daftar Perhitungan Probabilitas Kata dan Kategori....	56
Gambar 4.5	Tampilan Daftar Perhitungan Probabilitas Kata dan Kategori....	56
Gambar 4.6	Langkah Awal Proses Membuat Grafik	57
Gambar 4.7	Langkah Akhir Proses Membuat Grafik	57
Gambar 4.7	Langkah Akhir Proses Membuat Grafik	57
Gambar 4.8	Tampilan Diagram Hasil Perkalian Probabilitas	58
Gambar 4.9	Tampilan Hasil Prediksi Akhir	58
Gambar 4.10	Tampilan Hasil Parsing Data Pesan Masuk Sebelum Menjadi Data training	59
Gambar 4.11	Hasil akhir pengolahan data training	59

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sekolah merupakan salah satu tempat pendidikan yang umum atau dikenal di semua masyarakat. Proses pendidikan yang dilakukan di sekolah merupakan kegiatan utama yang dilakukan di sekolah. Aliran keuangan yang lancar, sarana dan prasarana yang baik akan sangat mendukung proses belajar mengajar yang diadakan di sekolah.

Di setiap kegiatan yang dilakukan di sekolah tersebut memerlukan program kerja yang baik, serta evaluasi menyeluruh di setiap kegiatan yang dilakukan. Kegiatan di sekolah yang tidak sedikit serta unit-unit bagian penunjang yang banyak, membutuhkan informasi yang tidak sedikit untuk melakukan evaluasi secara menyeluruh.

Proses pengumpulan informasi untuk melakukan evaluasi secara menyeluruh, dapat dilakukan dengan melakukan survei lapangan secara langsung, atau dapat melalui tanggapan-tanggapan secara langsung dari siswa-siswi, dalam hal ini sebagai objek dalam proses pembelajaran.

Pada penelitian yang dilakukan melalui tugas akhir ini, memproses tanggapan-tanggapan dari para siswa-siswi, guru, serta karyawan sekolah melalui pesan yang dikirimkan menggunakan SMS (*Short Message Service*). Pesan yang dikirimkan tersebut kemudian di proses menggunakan algoritma metode naïve bayes classifier untuk dikategorikan sesuai dengan isi pesan SMS yang dikirimkan.

Proses mengkategorikan pesan tersebut, akan sangat memudahkan pihak sekolah untuk mengumpulkan informasi yang digunakan untuk evaluasi, selain proses pengumpulan informasi yang cepat, proses mengkategorikan dilakukan secara otomatis.

1.2 Rumusan Masalah

Dari permasalahan diatas dapat dirumuskan beberapa masalah antara lain sebagai berikut:

1. Mampukah sistem yang akan dibuat mengkategorikan pesan yang tidak terstruktur berbentuk sms menggunakan metode naïve bayes?
2. Apakah metode naïve bayes mampu mengkategorikan pesan dengan tepat untuk kasus sms masuk di SMK Koperasi ?

1.3 Batasan Masalah

Dalam pembuatan sistem ini digunakan beberapa batasan-batasan antara lain :

1. Pengaduan suara hanya dilakukan melalui pesan SMS.
2. Kategori pesan disesuaikan dengan topik-topik atau bahan yang sering di evaluasi di SMK Koperasi.
3. Sistem informasi yang dibuat menggunakan bahasa pemrograman PHP & Database MySQL sebagai tempat untuk menyimpan serta mengelola data.
4. Pemrosesan pesan untuk dikategorikan menggunakan algoritma metode Naïve bayes classifier.
5. Pesan yang tidak sesuai dengan data training, akan dianggap sebagai pesan SPAM.



1.4 Tujuan Penelitian

Adapun tujuan penulisan skripsi ini adalah sebagai berikut :

1. Memahami bagaimana membuat program berbasis SMS Gateway.
2. Memahami bagaimana teknik dalam mengkategorikan sebuah pesan yang dikirimkan menggunakan metode Naïve Bayes Classifier.
3. Mengerti bagaimana menganalisa sebuah pesan apakah termasuk sebuah spam atau tidak.
4. Memahami penggunaan teknik *text mining* dalam melakukan proses parsing data.
5. Sebagai tugas akhir untuk memenuhi syarat kelulusan program Sarjana.

1.5 Metode Pendekatan

Metode pendekatan yang dilakukan dalam penelitian ini untuk mengumpulkan data dilakukan sebagai berikut :

- a. Studi Pustaka
Mempelajari buku-buku literatur yang berkaitan dengan permasalahan yang akan diteliti dengan tujuan mendapatkan sumber-sumber pemahaman yang dapat membantu proses dalam penyelesaian permasalahan dalam penelitian.
- b. Pengumpulan Melalui Media Internet
Mengumpulkan artikel-artikel, jurnal atau beberapa karya ilmiah yang mendukung tentang penelitian yang dibahas atau pernah membahas permasalahan tersebut dengan menggunakan metode lain.
- c. Perancangan
Setelah data berhasil dikumpulkan, penulis merancang suatu sistem yang nantinya dapat membantu proses penyelesaian dari penelitian. Sistem tersebut akan dibuat menggunakan bahasa pemrograman PHP dan bantuan aplikasi Gammu untuk mengakses file SMS Gateway yang dibuat. Untuk pemecahan permasalahan diatas digunakan metode Naïve bayes Classifier yang nantinya digunakan sebagai teknik penyelesaian dari penelitian yang penulis bahas.

1.6 Sistematika Penulisan

Laporan Penelitian ini akan terbagi dalam beberapa bab dengan sistematika penulisan sebagai berikut :

BAB I : Pendahuluan yang membahas mengenai latar belakang masalah, bagaimana masalah itu kemudian dikaji ulang dalam rumusan masalah, mendefinisikan batasan-batasan permasalahan, menggali manfaat dari penelitian ini serta bagaimana metode yang digunakan dalam mengumpulkan data-data yang akan digunakan dalam penelitian.

BAB II : Pada bab ini terdiri dari Tinjauan pustaka dan Landasan teori yang membahas mengenai beberapa teori yang didapatkan dari sumber-sumber yang berhubungan dengan penelitian Tugas Akhir ini. Selain itu, pada bab ini juga dijelaskan bagaimana konsep-konsep yang diperlukan untuk memecahkan permasalahan yang dibahas pada penelitian Tugas Akhir ini.

BAB III : Pada bab ini akan dibahas bagaimana sumber-sumber teori pada bab sebelumnya yang telah dibahas digunakan dalam sistem yang akan dibuat, sehingga terbentuk rancangan-rancangan, simulasi penyelesaian kasus yang diuraikan menggunakan metode yang dipilih.

BAB IV : Pada bab ini dijelaskan mengenai implementasi dari hasil penelitian yaitu sebuah sistem yang bekerja menggunakan teknik metode yang diambil, pada bab ini berisi mengenai gambar-gambar foto yang menjelaskan bagaimana sebuah sistem memecahkan kasus yang diambil.

BAB V : Pada bab ini berisi mengenai kesimpulan dari hasil penelitian yang dibuat berhubungan analisis yang dilakukan, serta saran-saran untuk kegiatan riset ke depan

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari penelitian yang dilakukan diperoleh kesimpulan sebagai berikut:

1. Metode naïve bayes mampu mengkategorikan pesan SMS yang dikirimkan oleh pengguna dengan nilai rata-rata prosentase sebesar 73%.
2. Pemilihan fitur yang digunakan sebagai data training, sangat penting untuk menentukan ciri-ciri dari masing-masing kelas.
3. Selisih token atau fitur sebuah kategori terhadap kategori lainya juga menentukan hasil perkalian probabilitas, kategori dengan token yang terlalu rendah dari kategori lain, berdampak hasil prediksi akan lebih memihak kategori dengan jumlah token terkecil. Sebaliknya jumlah token yang terlalu tinggi membuat hasil probabilitas akan berkurang, sehingga jumlah token di masing-masing kategori sebaiknya memiliki selisih yang tidak terlalu jauh.

5.2 Saran

Untuk pengembangan penelitian sebaiknya metode naïve bayes ditambahkan dengan metode lain untuk menghitung pembobotan sebuah kata terhadap sebuah kelas, sehingga proses mengkategorikan akan lebih akurat.

DAFTAR PUSTAKA

- ^[1]Ahmad Solichin.(2010).*MySQL 5 dari pemula hingga mahir*.Jakarta.
- ^[2]Edison Tarigan,Daud.(2012).*Membangun SMS Gateway Berbasis Web Dengan CodeIgniter*.Yogyakarta:Lolomedia.
- ^[3]Feldman,R & Sanger,J. (2007).*The Text Mining Handbook Advanced Approaches in Analyzing Unstructured*. Cambridge University Press:New York
- ^[4]Hakim,Lukmanul.(2009).*Trik rahasia master PHP terbongkar lagi*.Sleman:Lokomedia.
- ^[5]Han,J & Kamber, M.(2006).*Data Mining Concepts and Techniques*.Morgan Kaufmann Publishers:San Fransisco.
- ^[6]Indranandita, A.(2008).*Sistem Klasifikasi dan Pencarian Jurnal dengan Menggunakan Metode Naïve Bayes dan Vector Space Model*. Jurnal Informatika Vol.no 2
- ^[8]Ni Wayan Sumartini Saraswati.(2011).*Text Mining dengan Metode Naïve Bayes Classifier Dan support vector machine untuk sentiment analysis*.
- ^[9]Saputra,Agus&Agustin,Feni.(2012).*62 Trik dan Plugin Terbaik JQuery*.Jakarta:PT.Elex Media Komputindo.
- ^[10]Wahidin.(2010).*Aplikasi SMS dengan PHP untuk Orang Awam*. Palembang:Maxikom.