

KLASIFIKASI TWEET MENGGUNAKAN ALGORITMA NAÏVE BAYES

kripsi



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

Disusun oleh

I WAYAN RYANDI SASKARA
22094757

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2014

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

KLASIFIKASI TWEET MENGGUNAKAN ALGORITMA NAÏVE 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, 17 Juni 2014



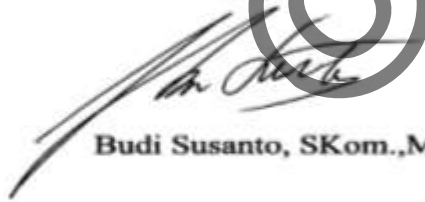
I WAYAN RYANDI SASKARA
22094757

HALAMAN PERSETUJUAN

Judul Skripsi : KLASIFIKASI TWEET MENGGUNAKAN
ALGORITMA NAÏVE BAYES
Nama Mahasiswa : I WAYAN RYANDI SASKARA
N I M : 22094757
Matakuliah : Skripsi (Tugas Akhir),
Kode : TIW276
Semester : Genap
Tahun Akademik : 2013/2014

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 28 Mei 2014

Dosen Pembimbing I



Budi Susanto, SKom.,M.T.

Dosen Pembimbing II



Antonius Rachmat C., SKom.,M.Cs

HALAMAN PENGESAHAN

KLASIFIKASI TWEET MENGGUNAKAN ALGORITMA NAÏVE BAYES

Oleh: I WAYAN RYANDI SASKARA / 22094757

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 11 Juni 2014

Yogyakarta, 17 Juni 2014

Mengesahkan,

Dewan Penguji:

1. Budi Susanto, S.Kom, MT.
2. Antonius Indranto G., S.Kom, M.Cs
3. Purnadi Basy Wuliyah, S.Si., MT.
4. Igratia Dhian E.K.R., S.Kom



Dekan



(Drs. Winnie Haridwidjaja, MT.)

Ketua Program Studi



(Nugroho Agus Haryono, M.Si)

UCAPAN TERIMAKASIH

Puji Tuhan Ida sang Hyang Widhi Wasa. Terimakasih atas penyertaan dan berkat-Mu sehingga penulis dapat menyelesaikan Tugas akhir yang berjudul Klasifikasi Tweet Menggunakan Algoritma *Naive Bayes* dengan baik.

Penulis mengucapkan terimakasih yang sebesar-besarnya kepada semua pihak yang telah membantu dan memberikan dukungan kepada penulis selama penyusunan Tugas Akhir ini, diantaranya:

1. Bapak **Budi Susanto, SKom.,M.T** selaku dosen pembimbing 1, yang telah memberikan ide, kritik, masukan dan saran dalam penulisan laporan dan pembuatan Tugas akhir ini.
2. Bapak **Antonius Rachmat C., SKom.,M.Cs** selaku dosen pembimbing 2, yang telah memberikan ide, kritik, masukan dan saran dalam penulisan laporan dan pembuatan Tugas akhir ini.
3. Kedua orang tua penulis. Papa I Made Rai Suaka dan mama Ni Wayan kianasih serta adik-adik penulis Ni Kadek reyka dan I Komang Aditya yang selalu memberikan dukungan, perhatian, kesabaran dan doa untuk penulis. Terimakasih atas kesempatan menuntut ilmu yang telah kalian berikan kepada penulis.
4. Keluarga Berkibi, Joseph Carlo, Guan, Jayadi, Prima Adi, Ni Made Ari Pratiwi, Yohanes Esthi, Daniel Mulia, Sherly Fanny Karlinda, Tinodan Yohana Gloria yang telah menjadi teman seperjuangan penulis yang selalu memberikan dukungan dan motivasi bagi penulis selama menempuh masa studi di Universitas Kristen Duta Wacana.
5. Gede Roy Sanjaya dan Ni Ketut Renik yang menjadi teman seperjuangan sekampung halaman yang selalu membantu penulis dalam berbagai hal.
6. I Wayan Pasek Ambara, yang telah menjadi teman penulis dalam setiap masalah yang terjadi dalam mengerjakan tugas akhir ini.
7. Keluarga KMHD UKDW, Putu Guna Satwam dan I Made Himawan Raharja yang telah menemani penulis dalam dunia perkuliahan selama 4 tahun ini dan memberikan semangat dan motivasi yang besar.

8. Keluarga Galeri BEI UKDW, kak Ririn Safitri yang telah memberikan penulis kesempatan bekerja sebagai part timer di Galeri BEI UKDW dan mengenalkan dunia pasar modal, ibu presiden Evelen Besthari Pongsibidang yang menjadi inspirasi penulis dalam mengenal dunia organisasi. kakak volunteer abadi Laurentius Pitoyo yang menjadi inspirasi penulis dalam mengenal dunia usaha, nyonya Maria Binartia Sari Indralusiana yang selalu mengajak penulis dalam menimba ilmu diluar bidang akademis, master Thomas Alan setiawan yang memiliki semangat luar biasa dalam hal belajar, serta nyonya teliti Yulita Heni yang menjadi teman seperjuangan penulis dalam mengikuti lomba KNPM . Terimakasih Galeri BEI UKDW atas kesempatan dan pengalaman yang telah diberikan kepada penulis.
9. Keluarga KSPM UKDW, Kristian Broto, Andre, Erbia, Yolanda, Thania, Ferica, Richard dan teman-teman KSPM lainnya yang tidak penulis sebutkan satu persatu.
10. Pihak-pihak lain yang tidak dapat penulis sebutkan satu persatu secara langsung maupun tidak langsung yang telah mendukung penyelesaian Tugas Akhir ini.

Yogyakarta, 30 mei 2014

Penulis

I Wayan Ryandi Saskara

DAFTAR ISI

UCAPAN TERIMAKASIH	5
DAFTAR ISI.....	7
DAFTAR GAMBAR.....	10
DAFTAR TABEL	11
DAFTAR GRAFIK.....	12
DAFTAR LAMPIRAN.....	13
INTISARI	Error! Bookmark not defined.
BAB I.....	Error! Bookmark not defined.
PENDAHULUAN	Error! Bookmark not defined.
1.1 Latar Belakang Masalah	Error! Bookmark not defined.
1.2 Perumusan Masalah.....	Error! Bookmark not defined.
1.3 Batasan Masalah	Error! Bookmark not defined.
1.4 Tujuan Penelitian	Error! Bookmark not defined.
1.5 Metode Penelitian.....	Error! Bookmark not defined.
1.6 Sistematika Penulisan	Error! Bookmark not defined.
BAB II	Error! Bookmark not defined.
LANDASAN TEORI	Error! Bookmark not defined.
2.1 Tinjauan Pustaka.....	Error! Bookmark not defined.
2.2 Landasan Teori.....	Error! Bookmark not defined.
2.2.1 Twitter.....	Error! Bookmark not defined.
2.2.2 Twitter API.....	Error! Bookmark not defined.
2.2.3 Autentifikasi dan Authorisasi API.....	Error! Bookmark not defined.
2.2.4 Text Mining	Error! Bookmark not defined.
2.2.5 Naive Bayes Classifier.....	Error! Bookmark not defined.
2.2.6 Evaluasi Sistem.....	Error! Bookmark not defined.
BAB III.....	Error! Bookmark not defined.
ANALISIS DAN PERANCANGAN SISTEM.....	Error! Bookmark not defined.
3.1 Kebutuhan Fungsional dan Non Fungsional	Error! Bookmark not defined.
3.1.1 Kebutuhan Fungsional	Error! Bookmark not defined.

3.1.2 Kebutuhan Non Fungsional.....	Error! Bookmark not defined.
3.2 Perancangan Sistem.....	Error! Bookmark not defined.
3.3 Diagram Use Case	Error! Bookmark not defined.
3.4 Rancangan Alur Kerja Sistem	Error! Bookmark not defined.
3.4.1 Perancangan proses Login	Error! Bookmark not defined.
3.4.2 Perancangan proses pengambilan tweet pada twitter.....	Error! Bookmark not defined.
3.4.3 Perancangan Proses Preprocessing Tweet Kotor.....	Error! Bookmark not defined.
3.4.4 Perancangan Proses Perhitungan Bobot TF-IDF	Error! Bookmark not defined.
3.4.5 Perancangan Proses <i>Naive Bayes Classifier</i>	Error! Bookmark not defined.
3.5 Kamus Data	Error! Bookmark not defined.
3.5.1 Tabel Tweet_kotor	Error! Bookmark not defined.
3.5.2 Tabel Stoplist.....	Error! Bookmark not defined.
3.5.3 Tabel tweet_bersih	Error! Bookmark not defined.
3.5.4 Tabel fs	Error! Bookmark not defined.
3.5.5 Tabel token_kategori_fs	Error! Bookmark not defined.
3.5.6 Tabel eval_sistem.....	Error! Bookmark not defined.
3.5.7 Tabel Klasifikasi	Error! Bookmark not defined.
3.5.8 Tabel Token_all_fs	Error! Bookmark not defined.
3.5.9 Tabel Naive.....	Error! Bookmark not defined.
3.6 Skema Diagram	Error! Bookmark not defined.
3.7 Rancangan Antarmuka Sistem.....	Error! Bookmark not defined.
3.8 Rancangan Pengujian Sistem.....	Error! Bookmark not defined.
BAB IV	Error! Bookmark not defined.
IMPLEMENTASI DAN ANALISIS SISTEM	Error! Bookmark not defined.
4.1 Menu Sistem dan Operasional.....	Error! Bookmark not defined.
4.1.1 Tampilan Menu Login	Error! Bookmark not defined.
4.1.2 Tampilan Menu Utama	Error! Bookmark not defined.
4.1.3 Tampilan Halaman Tweet Training	Error! Bookmark not defined.
4.1.4 Tampilan Halaman Generate Tweet Bersih	Error! Bookmark not defined.
4.1.5 Tampilan Halaman Lihat Bobot Tweet.....	Error! Bookmark not defined.
4.1.6 Tampilan Halaman Evaluasi Sistem.....	Error! Bookmark not defined.
4.2 Implementasi Sistem Klasifikasi	Error! Bookmark not defined.

4.2.1 Implementasi Proses Pengambilan Data Tweet	Error! Bookmark not defined.
4.2.2 Implementasi Proses Preprocessing Tweet Kotor	Error! Bookmark not defined.
4.2.3 Implementasi Proses Perhitungan TF-IDF.....	Error! Bookmark not defined.
4.2.4 Implementasi Proses Normalisasi Nilai TF-IDF.....	Error! Bookmark not defined.
4.2.5 Implementasi Proses Feature Selection Mode All	Error! Bookmark not defined.
4.2.6 Implementasi Proses Feature Selection Mode Kategori	Error! Bookmark not defined.
4.2.7 Implementasi Algoritma Naïve Bayes	Error! Bookmark not defined.
4.3 Analisis sistem.....	Error! Bookmark not defined.
4.3.1 Pengujian dan Evaluasi dengan <i>Feature Selection</i> 10% Mode Kategori	Error! Bookmark not defined.
4.3.2 Pengujian dan Evaluasi Dengan <i>Feature Selection</i> 20% Mode Kategori	Error! Bookmark not defined.
4.3.3 Pengujian dan Evaluasi dengan <i>Feature Selection</i> 40% Mode Kategori	Error! Bookmark not defined.
4.3.4 Pengujian dan Evaluasi dengan <i>Feature Selection</i> 10% Mode All	Error! Bookmark not defined.
4.3.5 Pengujian dan Evaluasi dengan <i>Feature Selection</i> 20% Mode All	Error! Bookmark not defined.
4.3.6 Pengujian dan Evaluasi Dengan <i>Feature Selection</i> 40% Mode All	Error! Bookmark not defined.
4.3.7 Perbandingan Precision Mode All.....	Error! Bookmark not defined.
4.3.8 Perbandingan Recall Mode All	Error! Bookmark not defined.
4.3.9 Perbandingan Precision Mode Kategori	Error! Bookmark not defined.
4.3.10 Perbandingan Recall Mode Kategori.....	Error! Bookmark not defined.
4.3.11 Perbandingan Akurasi Pengujian Mode Kategori dan All	Error! Bookmark not defined.
BAB V	Error! Bookmark not defined.
KESIMPULAN DAN SARAN	Error! Bookmark not defined.
5.1 Kesimpulan.....	Error! Bookmark not defined.
5.2 Saran	Error! Bookmark not defined.

DAFTAR GAMBAR

GAMBAR	KETERANGAN	HALAMAN
Gambar 2.1	Algoritma Naive Bayes	11
Gambar 3.1	Gambaran Umum Sistem	18
Gambar 3.2	Use Case Diagram	19
Gambar 3.3	Flowchart Proses Autentifikasi Oauth	21
Gambar 3.4	Flowchart Pengambilan Tweet	22
Gambar 3.5	Flowchart Preprocessing Tweet Kotor	24
Gambar 3.6	Flowchart proses perhitungan TF-IDF	27
Gambar 3.7	Flowchart Proses Pembentukan Feature	28
Gambar 3.8	Flowchart proses normalisasi TF-IDF	29
Gambar 3.9	Flowchart Proses Naive Bayes Classifier	31
Gambar 3.10	Skema Diagram Sistem	38
Gambar 3.11	Perancangan Halaman Utama Sistem	39
Gambar 3.12	Perancangan Halaman Pengambilan Tweet	39
Gambar 3.13	Perancangan Halaman Data <i>Stoplist</i>	40
Gambar 3.14	Perancangan Halaman Svaluasi Sistem	40
Gambar 4.1	Tampilan Halaman Authorize App	43
Gambar 4.2	Tampilan Halaman Utama Sistem	44
Gambar 4.3	Tampilan Halaman Tweet Training	45
Gambar 4.4	Tampilan Halaman Tweet Bersih	46
Gambar 4.5	Tampilan Halaman Lihat Bobot Tweet	47
Gambar 4.6	Halaman Evaluasi Sistem	49

DAFTAR TABEL

TABEL	KETERANGAN	HALAMAN
Tabel 2.1	Tabel Contoh Tweet	6
Tabel 2.2	Tabel Confussion Matrix	14
Tabel 2.3	Tabel Notasi Confussion Matrix 3 Kelas	15
Tabel 3.1	Tabel Tweet_Kotor	32
Tabel 3.2	Tabel Stoplist	32
Tabel 3.3	Tabel Tweet_Bersih	33
Tabel 3.4	Tabel fs	33
Tabel 3.5	Tabel token_kategori_fs	34
Tabel 3.6	Tabel Eval_Sistem	34
Tabel 3.7	Tabel klasifikasi	39
Tabel 3.8	Tabel token_all_fs	37
Tabel 3.9	Tabel naive	37
Tabel 3.10	Tabel confusion matrix untuk klasifikasi 4 kelas	41
Tabel 4.1	Tabel Contoh Nilai Bobot Token	48
Tabel 4.2	Tabel Feature 10% Mode Kategori	57
Tabel 4.3	Tabel Feature 20% Mode Kategori	59
Tabel 4.4	Tabel Feature 40% Mode Kategori	61
Tabel 4.5	Tabel Feature 10% Mode <i>All</i>	63
Tabel 4.6	Tabel Feature 20% Mode <i>All</i>	66
Tabel 4.7	Tabel Feature 40% Mode <i>All</i>	68
Tabel 4.8	Tabel Perbandingan Nilai <i>Precision</i> Mode <i>All</i>	70
Tabel 4.9	Tabel Perbandingan <i>Recall</i> Mode <i>All</i>	71
Tabel 4.10	Tabel Perbandingan <i>Precision</i> Mode Kategori	72
Tabel 4.11	Tabel Perbandingan <i>Recall</i> Mode Kategori	73
Tabel 4.12	Tabel Perbandingan Akurasi Mode Kategori	74

DAFTAR GRAFIK

Grafik	Keterangan	Halaman
Grafik 4.1	Perbandingan Nilai Precision Pada Mode <i>All</i>	71
Grafik 4.2	Perbandingan Nilai <i>Recall</i> Pada Mode <i>All</i>	72
Grafik 4.3	Perbandingan Nilai Precision Pada Mode Kategori	73
Grafik 4.4	Perbandingan Nilai <i>Recall</i> Pada Mode Kategori	74
Grafik 4.5	Perbandignan Nilai akurasi	75

©UKDW

DAFTAR LAMPIRAN

LAMPIRAN	KETERANGAN	HALAMAN
Lampiran A	Lampiran tweet testing	A-1
Lampiran B	Lampiran Source Code	B-1

©UKDW

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Jejaring sosial merupakan salah satu media penyebaran konten berita yang sangat cepat. salah satu sosial media tersebut adalah *Twitter*. *Twitter* merupakan situs jejaring sosial yang memiliki tingkat kecepatan *update* informasi yang cukup tinggi. Sekarang hampir semua portal berita mempunyai akun *Twitter*, yang digunakan sebagai media ‘perpanjangan tangan’ penyebaran informasi. *Tweet* merupakan cara *user* dari pengguna akun twitter berbagi pesan atau informasi di situs jejaring sosial ini. *Twitter* mengirimkan rata-rata 58 juta *tweet* setiap harinya, dan 9100 *tweet* per detik. (<http://www.statisticbrain.com/Twitter-statistics/>, 2013).

Pada situs portal berita, berita atau informasi digolongkan berdasarkan topik atau kategori tertentu, seperti kategori olahraga, politik dan lain-lain, sehingga dapat memudahkan pembaca dalam memilih konten berita yang akan di baca. Namun pada *Twitter*, akun portal berita tidak mengkategorikan berita yang di posting, sehingga hal ini menyulitkan *follower* dari portal berita tersebut untuk menelusuri berita dari akun portal berita tersebut untuk mendapatkan berita berdasarkan kategori yang di inginkan.

Oleh karena itu, akan dirancang sebuah sistem klasifikasi *tweet* yang dapat mengklasifikasikan kategori berita yang di posting oleh akun portal berita tersebut dengan menggunakan algoritma *Naïve bayes Classifier* kedalam kategori yang telah ditentukan. Adapun kategori yang akan di gunakan sebagai klasifikasi berita adalah olahraga, keungan, politik, dan teknologi. Sumber data yang digunakan sebagai dasar

pengujian sistem berasal dari akun *Twitter* portal berita Indonesia seperti @detikcom, @tribunnews, @Okezone dan tempodotco.

1.2 Perumusan Masalah

Rumusan masalah yang akan dibahas dalam penelitian ini adalah seberapa akurat metode *Naïve Bayes* melakukan klasifikasi kumpulan *tweet* yang diposting oleh @detikcom, @tribunnews, @Okezone, dan @tempodotco kedalam kategori olahraga, keuangan, politik, dan teknologi .

1.3 Batasan Masalah

Ruang lingkup masalah dalam penelitian ini dibatasi oleh hal-hal sebagai berikut :

Batasan data :

- Data *tweet* yang digunakan dalam penelitian ini adalah *tweet* yang disampaikan dalam bahasa Indonesia yang diposting oleh akun @detikcom, @tribunnews, @okezone, dan @tempodotco.
- *Tweet* akan di kategorikan kedalam 4 kategori olahraga, keuangan, politik, dan teknologi
- Bentuk keluaran sistem adalah berupa *tweet* yang telah memiliki label kategori berdasarkan hasil klasifikasi.
- *Stoplist* menggunakan *database* stopword yang di ambil dari <http://lecturer.ukdw.ac.id/budsus/pdf/stopwordID.csv>. *User* dapat menambahkan daftar stopword.

Batasan Sistem :

- Sistem melakukan pembobotan *tweet* training dengan metode TF-IDF, pengklasifikasian menggunakan *Naive Bayes Classifier* serta evaluasi sistem dengan menggunakan metode *holdout* dengan dengan pengukuran nilai *precision*, *recall* dan akurasi hasil klasifikasi pengujian .

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk :

1. Untuk membangun suatu aplikasi yang dapat melakukan filtering berita yang disampaikan melalui *tweet* dalam bahasa Indonesia oleh akun portal berita di twitter.
2. Untuk mengetahui akurasi klasifikasi *Naïve Bayes* ketika di aplikasikan pada klasifikasi text dengan jumlah token penyusun yang sedikit.
3. Mempermudah *follower* dari akun portal berita dalam otomatisasi pengkategorian dari *tweet*.

1.5 Metode Penelitian

Berikut uraian metode yang di gunakan dalam penelitian ini adalah :

1. Studi Literatur, melakukan berbagai pengumpulan bahan referensi, seperti Jurnal Penelitian, buku-buku teori dan sumber-sumber informasi lain yang di peroleh dari internet.
2. Analisa sistem, melakukan analisis kebutuhan dalam pengembangan sistem
3. Perancangan sistem, meliputi pengumpulan data *training*, proses dan perancangan antarmuka aplikasi.
4. Implementasi program & *testing* , merupakan proses penulisan kode program dan pembangunan basis data yang disesuaikan dengan desain proses, data , dan antarmuka yang telah di buat pada proses perancangan sistem.
5. Analisis hasil klasifikasi dan evaluasi, merupakan proses menganalisis keektifan program yang dibuat dalam melakukan klasifikasi terhadap data uji.

1.6 Sistematika Penulisan

Sistem penulisan laporan tugas akhir ini dibagi menjadi berapa bab, yaitu sebagai berikut :

Bab I, Pendahuluan, yang dapat memberikan gambaran umum tentang penelitian. pendahuluan berisi latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, metodologi penelitian, dan sistematika penulisan.

Bab II, Landasan teori, yang terdiri dari dua bagian utama, yaitu tinjauan pustaka dan landasan teori. Tinjauan Pustaka menguraikan berbagai teori yang diperoleh dari jurnal ataupun karya ilmiah yang berkaitan dengan topik penelitian. sedangkan landasan teori memuat mengenai penjelasan terhadap konsep dan teori-teori yang diperlukan dalam memecahkan masalah dari penelitian.

Bab III, Analisis dan perancangan sistem, yang berisi spesifikasi kebutuhan sistem yang akan dibangun, arsitektur sistem, use-case diagram, tahapan dalam perancangan dan pembangunan sistem , aliran data dan antarmuka input output sistem.

Bab IV, Implementasi dan analisis sistem, membahas tentang implementasi perancangan sistem yang dibuat pada bab III yang disertai analisis dari sistem yang dibuat terhadap beberapa set data test.

Bab V, Kesimpulan dan saran, berisi mengenai kesimpulan dari hasil penelitian serta memberikan masukan berupa saran yang dapat digunakan sebagai acuan dalam pengembangan sistem yang telah dibuat agar menjadi lebih baik lagi.

©UKDW

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan implementasi sistem, maka diperoleh kesimpulan sebagai berikut :

- Sistem klasifikasi *Naive Bayes* cukup baik untuk digunakan dalam klasifikasi tweet dengan nilai akurasi tertinggi pada mode kategori sebesar 0.89 dengan feature selection 40% .
- Untuk *feature* selection mode all, nilai akurasi tertinggi berada pada *feature* selection 20% dengan nilai akurasi 0.776.
- Pada mode all, *classifier* cenderung mengklasifikasikan tweet sebagai kategori olahraga. Hal ini terlihat dari nilai *precision* kategori olahraga paling rendah jika dibandingkan dengan kategori lain.
- Akurasi yang diperoleh dengan menggunakan mode kategori lebih baik daripada dengan menggunakan mode all.

5.2 Saran

- Sistem Dapat dikembangkan dengan menambahkan isi *tweet* dengan mengambil konten berita yang direferensikan dari link tiap *tweet*. Sehingga didapatkan lebih banyak variasi *token* untuk *tweet* tersebut.
- Sistem dapat dikembangkan untuk perangkat mobile sehingga dapat memudahkan *user* dalam mengkategorikan *tweet* kedalam kategori yang diinginkan.

- Sistem dapat ditambahkan kategori lainnya selain kategori politik, keuangan, teknologi dan olahraga, seperti kategori netral yang digunakan untuk *tweet* yang tidak termasuk kedalam 4 kategori tersebut.

©UKDW

DAFTAR PUSTAKA

- Aliandu, Paulina (2012). *Analisis Sentimen Tweet Berbahasa Indonesia Di Twitter*. Yogyakarta: Universitas Gadjah Mada.
- Anugroho, Preatsetyo. 2010. *Klasifikasi Email Spam Dengan Metode Naïve Bayes Classifier Menggunakan Java Programming*, Surabaya. Politeknik Elektronika Negeri Surabaya.
- Feldman, R. & Sanger, J. (2007). *The Text Mining Handbook*. New York: Cambridge University Press.
- Indranandita, A., Susanto, B., & Rachmat, Antonius C. (2008). *SISTEM KLASIFIKASI DAN PENCARIAN JURNAL DENGAN MENGGUNAKAN METODE NAÏVE BAYES DAN VECTOR SPACE MODEL*. Diakses tanggal 3 Maret 2014 dari <http://ti.ukdw.ac.id/ojs/index.php/informatika/article/view/48/19>.
- Intan, R. & Defeng, A. Hard: Subject-Based Search Engine Menggunakan Tf-idf dan Jaccard's Coefficient. Diakses tanggal 2 Februari 2014 dari <http://puslit.petra.ac.id/files/published/journals/IND/IND060801/IND06080106.pdf>.
- Kandaga, T . 2012. *Sentiment Classification menggunakan Machine Learning: Metode Naïve-Bayes dan Support Vector Machines (Studi kasus: movie reviews imdb.com)*. Di akses tanggal 12 Agustus 2013 dari <http://repository.maranatha.edu/2573/1/Sentiment%20Classification%20menggunakan%20Machine%20Learning.pdf>.
- Kusumo, G.D. (2013). *Klasifikasi Sentimen Pembaca Berdasar Komentar Terhadap Berita Menggunakan Metode naive bayes Classifier*. (Undergraduate thesis, Duta Wacana Christian University, 2013).
- Manning, C.D., Raghavan, P., & Schütze, H. (2008). *An Introduction to Information Retrieval*. New York : Cambridge University Press.
- Sriram, Bharath. (2010). *Short Text Classification In Twitter To Improve Information Filtering*. Ohio: The Ohio State University.