

**Klasifikasi Sentimen Terhadap Dataset Komentar Pemilu
Presiden 2014 Pada *Facebook* Menggunakan
*Support Vector Machine***

Tugas Akhir



Oleh :

JAYADI KURNIAWAN

71120146

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
YOGYAKARTA
2017

©UKDW

**Klasifikasi Sentimen Terhadap Dataset Komentar Pemilu Presiden 2014
Pada Facebook Menggunakan Support Vector Machine**

Tugas Akhir



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

Disusun oleh :

JAYADI KURNIAWAN

71120146

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

2017

©UKDW

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

KLASIFIKASI SENTIMEN TERHADAP DATASET KOMENTAR PEMILU PRESIDEN 2014 PADA FACEBOOK MENGGUNAKAN SUPPORT VECTOR MACHINE

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 keserjanaan 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 keserjanaan saya.

Yogyakarta, 12 Juni 2017



Jayadi Kurniawan
JAYADI KURNIAWAN

71120146

HALAMAN PERSETUJUAN

Judul Skripsi : KLASIFIKASI SENTIMEN TERHADAP DATASET
KOMENTAR PEMILU PRESIDEN 2014 PADA
FACEBOOK MENGGUNAKAN SUPPORT
VECTOR MACHINE

Nama Mahasiswa : JAYADI KURNIAWAN

N I M : 71120146

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Genap

Tahun Akademik : 2016/2017

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 8 Mei 2017

Dosen Pembimbing I



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

Dosen Pembimbing II



Yuan Lukito, S.Kom., M.Cs.

HALAMAN PENGESAHAN

KLASIFIKASI SENTIMEN TERHADAP DATASET KOMENTAR PEMILU PRESIDEN 2014 PADA FACEBOOK MENGGUNAKAN SUPPORT VECTOR MACHINE

Oleh: JAYADI KURNIAWAN / 71120146

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 26 Mei 2017


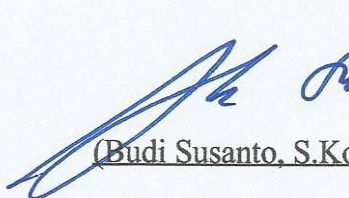
Yogyakarta, 12 Juni 2017
Mengesahkan,

Dewan Penguji:

1. Antonius Rachmat C., S.Kom., M.Cs.
2. Yuan Lukito, S.Kom., M.Cs.
3. Gloria Virginia, S.Kom., MAI, Ph.D.
4. Lukas Chrisantyo, S.Kom., M.Eng.

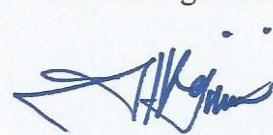


Dekan



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

Ketua Program Studi



(Gloria Virginia, Ph.D.)

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan penyertaan-Nya sehingga skripsi yang berjudul “Klasifikasi Sentimen Terhadap Dataset Komentar Pemilu Presiden 2014 Pada *Facebook* Menggunakan *Support Vector Machine*” ini dapat terselesaikan. Penulisan skripsi ini disusun sebagai karya tulis untuk memenuhi salah satu persyaratan guna memperoleh gelar Sarjana (S1) pada Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Duta Wacana.

Terselesainya penulisan skripsi ini tidak lepas dari campur tangan berbagai pihak, dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Allah Bapa, Tuhan Yesus Kristus dan Roh Kudus atas setiap penyertaan, berkat serta kasih yang tiada habisnya.
2. Bapak Antonius Rachmat C., S.Kom., M.Cs. selaku dosen pembimbing I atas arahan dan masukan selama pembuatan skripsi. Saya belajar banyak dari bapak.
3. Bapak Yuan Lukito, S.Kom, M.Cs. selaku dosen pembimbing II untuk setiap saran dan kritik selama pembuatan skripsi. Terima kasih untuk kesabaran bapak dalam membimbing.
4. Kedua orang tua terkasih, Haryadie Linang dan Rutse atas setiap doa, dukungan, kesabaran untuk mendengarkan keluh kesahku. Terima kasih sudah menjadi papah dan mamah yang luar biasa buat aku. Sehat terus!
5. *My beloved siblings* Cahaya Eka Novianti, Tri Novia Kalawa dan Satria Irpandi terima kasih untuk dukungan, candaan dan doa kalian. God bless!
6. Keluarga besar yang senantiasa mendukung dan mendoakan.
7. Dewi Setyaningsih S., S.Si. untuk motivasi, doa, perhatian dan dukungan yang begitu berarti buat aku. Kamu semangatku sejauh ini!
8. Diri sendiri yang telah berjuang melawan malas dan khawatir. *I did it!*
9. Teman-teman seperjuangan Teknik Informatika 2012 yang saling memberikan semangat dan dukungan satu sama lain.

10. Terakhir kepada seluruh teman, kerabat yang tidak dapat penulis cantumkan satu per satu, terima kasih untuk setiap doa dan dukungan kalian. Tuhan memberkati.

Dalam penulisan tugas akhir ini penulis menyadari masih terdapat banyak kekurangan. Oleh sebab itu penulis sangat mengharapkan kritik maupun saran yang bersifat membangun agar penulisan skripsi pada masa yang akan datang dapat dilakukan lebih baik lagi. Terima kasih.

©UKDW

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa, karena atas rahmat dan penyertaanNya, sehingga tugas akhir ini dapat terselesaikan. Terima kasih yang sebesar-besarnya penulis ucapkan kepada setiap pihak untuk segala kritik dan saran kepada penulis dalam menyelesaikan tugas akhir ini. Penulis menyadari bahwa penulisan laporan tugas akhir ini masih terdapat kekurangan baik dari materi ataupun penyusunannya. Maka dari itu penulis mengapresiasi segala kritik dan saran yang tujuannya membangun. Semoga laporan tugas akhir ini dapat memberikan manfaat kepada kita sekalian.

Yogyakarta, 7 Mei 2017

INTISARI

Klasifikasi Sentimen Terhadap Dataset Komentar Pemilu Presiden 2014 Pada Facebook Menggunakan Support Vector Machine

Media sosial adalah sarana komunikasi yang sangat populer saat ini jika dilihat dari jumlah pemakainya. Salah satu media sosial yang banyak digunakan yaitu Facebook dengan berbagai tujuan penggunaan misalnya sebagai tempat diskusi, komunitas pecinta olahraga tertentu, online shop, bahkan dalam bidang politik digunakan dimanfaatkan untuk mendapatkan dukungan dari para publik.

Tujuan penelitian ini adalah untuk menghasilkan suatu sistem klasifikasi sentimen berbasis *web*. *Dataset* yang digunakan adalah komentar status Facebook pemilihan umum presiden Indonesia tahun 2014. Metode *Support Vector Machine* (SVM) digunakan untuk proses klasifikasi komentar positif atau negatif. Tahap *preprocessing* yang dilakukan adalah *case folding*, *stemming*, *stopwords removal* dan tokenisasi. Data komentar kebanyakan mengandung *special character*, angka dan *emoticon*. Pada penelitian ini *emoticon* dikonversi ke bentuk teks sesuai dengan makna masing-masing. Pembobotan dilakukan menggunakan *Term Frequency Document Inverse Frequency (TF IDF)*. Metode lain yang diterapkan yaitu *feature selection* untuk menentukan nilai *term* yang akan digunakan dengan tujuan dapat menghasilkan nilai akurasi yang baik.

Pengujian menggunakan tiga jenis *dataset* dengan komposisi berbeda yaitu komentar positif dan negatif tidak sama (870 dan 130), komentar positif dan negatif masing – masing berjumlah 130, dan menggunakan data latih dengan 2024 komentar positif serta 313 komentar negatif. Evaluasi sistem dilakukan menggunakan *confusion matrix* untuk mengukur akurasi sistem dalam melakukan klasifikasi. Rata – rata akurasi menggunakan *kernel linear* terhadap *dataset* uji yaitu 73,84%, 50,37% dan 74,99%. Sedangkan rata-rata akurasi menggunakan *kernel RBF* adalah 76,53%, 51,83% dan 80,81%.

Kata Kunci: Klasifikasi Sentimen, *Text Mining*, *Support Vector Machine*, *Text Preprocessing*, *Confusion Matrix*, *TF IDF*, *Feature Selection*

DAFTAR ISI

PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN	v
UCAPAN TERIMA KASIH.....	vi
KATA PENGANTAR.....	viii
INTISARI.....	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
BAB I.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Sistem	3
1.4. Tujuan Penelitian	3
1.5. Metode Penelitian	3
1.6. Sistematika Penulisan.....	4
BAB II.....	5
2.1. Tinjauan Pustaka	5
2.2. Landasan Teori	7
2.2.1. <i>Text Mining</i>	8
2.2.2. Analisa sentimen (<i>Sentiment Analysis</i>)	8
2.2.4. <i>Term Frequency – Inverse Document Frequency (TF-IDF)</i>	10
2.2.5. <i>Feature Selection</i>	11
2.2.6. Pemanfaatan Emoticon.....	12
2.2.7. <i>Support Vector Machine</i>	12
2.2.8. Evaluasi Sistem	16
BAB III	18
3.1. Analisis Kebutuhan Sistem	18
3.1.1. Kebutuhan Perangkat Lunak (<i>Software</i>)	18
3.1.2. Kebutuhan Perangkat Keras (<i>Hardware</i>).....	18
3.2. Perancangan Sistem.....	18
3.2.1. Alur Kerja Sistem	19
3.2.2. <i>Flowchart Preprocessing</i>	19
3.2.2.1. <i>Flowchart Case Folding</i>	20
3.2.2.2. <i>Flowchart Stemming</i>	21
3.2.2.3. <i>Flowchart Stopwords Removal</i>	21
3.2.2.4. <i>Flowchart Tokenisasi</i>	22

3.2.3.	<i>Flowchart Perhitungan TF IDF</i>	23
3.3.1	Kamus Data	26
3.3.1.1.	Tabel <i>users</i>	26
3.3.1.2.	Tabel <i>stopwords</i>	27
3.3.1.3.	Tabel <i>emoji</i>	27
3.3.1.4.	Tabel <i>dataset_latih</i>	27
3.3.1.5.	Tabel <i>token_latih</i>	28
3.3.1.6.	Tabel <i>df_idf_latih</i>	28
3.3.1.7.	Tabel <i>hasil_preprocessing_latih</i>	29
3.3.1.8.	Tabel <i>selected_feature</i>	29
3.3.1.9.	Tabel <i>data_latih_ready</i>	30
3.3.1.10.	Tabel <i>model_latih</i>	30
3.3.1.11.	Tabel <i>dataset_uji</i>	31
3.3.1.12.	Tabel <i>token_uji</i>	31
3.3.1.13.	Tabel <i>df_idf_uji</i>	32
3.3.1.14.	Tabel <i>hasil_uji</i>	32
3.4.1	Halaman <i>home</i>	33
3.4.1.	Halaman Pelatihan SVM.....	33
3.4.2.	Halaman Pengujian SVM.....	34
3.5.1.	Pengujian <i>Feature Selection</i>	36
3.5.2.	Pengujian Menggunakan Data Tidak Seimbang.....	36
3.5.3.	Pengujian Menggunakan Data Seimbang.....	36
3.5.4.	Pengujian Menggunakan Data Latih sebagai Data Uji	37
BAB IV	38
4.1.	Implementasi Antarmuka Sistem	38
4.1.1.	Halaman <i>Home</i>	38
4.1.2.	Halaman Pelatihan.....	38
4.1.3.	Halaman Pengujian	40
4.2.	Implementasi Sistem	41
4.2.1.	Implementasi <i>Preprocessing</i>	41
4.2.2.	Implementasi <i>Support Vector Machine</i>	45
4.3.	Hasil Pengujian	47
4.3.1.	Hasil Pengujian menggunakan Data Tidak Seimbang	47
4.3.2.	Hasil Pengujian menggunakan Data Seimbang.....	50
4.3.3.	Hasil Pengujian menggunakan Data Latih sebagai Data Uji.....	53
4.4.	Analisis Hasil Pengujian	56
BAB V	59
5.1.	Kesimpulan	59
5.2.	Saran	59
DAFTAR PUSTAKA	61
LAMPIRAN B	72
LAMPIRAN C	74

DAFTAR GAMBAR

Gambar 2.1. Himpunan <i>Hyperplane</i>	13
Gambar 3.1. <i>Flowchart</i> Sistem Secara Umum	19
Gambar 3.2. <i>Flowchart</i> Preprocessing.....	20
Gambar 3.3. <i>Flowchart</i> Case Folding.	20
Gambar 3.4. <i>Flowchart</i> Stemming.	21
Gambar 3.5. <i>Flowchart</i> Stopwords Removal	22
Gambar 3.6. <i>Flowchart</i> Tokenisasi.....	22
Gambar 3.7. <i>Flowchart</i> Perhitungan <i>TF</i>	23
Gambar 3.8. <i>Flowchart</i> Perhitungan <i>DF</i> , <i>IDF</i> dan <i>TF.IDF</i>	24
Gambar 3.9. Skema <i>Database</i>	25
Gambar 3.10. Rancangan Halaman <i>Home</i>	33
Gambar 3.11. Rancangan Halaman Pelatihan <i>SVM</i>	33
Gambar 3.12. Rancangan Halaman Nilai <i>TF.IDF</i>	34
Gambar 3.13. Rancangan Halaman Pengujian.....	34
Gambar 3.14. Rancangan Halaman Pengujian Manual.	35
Gambar 3.15. Komposisi Data Uji dan Data Latih.....	35
Gambar 4.1. Halaman <i>Home</i>	38
Gambar 4.2. Halaman Pelatihan <i>SVM</i>	39
Gambar 4.3. Menu untuk Menentukan <i>Feature Selection</i>	39
Gambar 4.4. Halaman Pengujian <i>SVM</i>	40
Gambar 4.5. Halaman Hasil Pengujian <i>SVM</i>	40
Gambar 4.6. Contoh Data pada Tabel <i>dataset_latih</i>	42
Gambar 4.7. Contoh Data pada Tabel <i>token_latih</i>	42
Gambar 4.8. Contoh Data pada Tabel <i>selected_feature</i>	43
Gambar 4.9. Data Model <i>SVM</i>	46
Gambar 4.10. Perbandingan Rata-rata Hasil Pengujian	58

DAFTAR TABEL

Tabel 2.1 Hasil Penelitian Wibisono (2005).	5
Tabel 2.2. Hasil Penelitian Wulandini dan Nugroho (2009).	6
Tabel 2.3. Ringkasan Penelitian Sentimen Analisis.	7
Tabel 2.4. Contoh Hasil Tokenisasi.	9
Tabel 2.5. Contoh Hasil <i>Case Folding</i>	9
Tabel 2.6. Contoh Hasil <i>Stemming</i>	9
Tabel 2.7. Contoh Hasil <i>Stopwords Removal</i>	10
Tabel 2.8. Contoh Hasil Perhitungan <i>TF.IDF</i>	11
Tabel 2.9. <i>Confusion Matrix</i>	17
Tabel 3.1. Tabel users.	26
Tabel 3.2. Tabel stopwords.	27
Tabel 3.3. Tabel emoji.....	27
Tabel 3.4. Tabel dataset_latih.....	28
Tabel 3.5. Tabel token_latih.....	28
Tabel 3.6. Tabel df_idf_latih	29
Tabel 3.7. Tabel hasil_preprocessing_latih.....	29
Tabel 3.8. Tabel selected_feature.....	30
Tabel 3.9. Tabel data_latih_ready	30
Tabel 3.10. Tabel model_latih.....	31
Tabel 3.11. Tabel komentar.	31
Tabel 3.12. Tabel token_uji.....	31
Tabel 3.13. Tabel df_idf_uji	32
Tabel 3.14. Tabel hasil_uji.....	32
Tabel 3.15. Jumlah Komentar Positif dan Negatif.....	36
Tabel 4.1. Komposisi Data Uji	41
Tabel 4.2. Hasil Pengujian <i>Kernel Linear C = 1</i>	47
Tabel 4.3. Hasil Pengujian <i>Kernel Linear C = 1000</i>	48

Tabel 4.4. Hasil Pengujian <i>Kernel Linear</i> $C = 10000000$	48
Tabel 4.5. Hasil Pengujian <i>Kernel RBF</i> $C = 1$	49
Tabel 4.6. Hasil Pengujian <i>Kernel RBF</i> $C = 1000$	49
Tabel 4.7. Hasil Pengujian <i>Kernel RBF</i> $C = 10000000$	50
Tabel 4.8. Hasil Pengujian <i>Kernel Linear</i> $C = 1$	50
Tabel 4.9. Hasil Pengujian <i>Kernel Linear</i> $C = 1000$	51
Tabel 4.10. Hasil Pengujian <i>Kernel Linear</i> $C = 10000000$	51
Tabel 4.11. Hasil Pengujian <i>Kernel RBF</i> $C = 1$	52
Tabel 4.12. Hasil Pengujian <i>Kernel RBF</i> $C = 1000$	52
Tabel 4.13. Hasil Pengujian <i>Kernel RBF</i> $C = 10000000$	53
Tabel 4.14. Hasil Pengujian <i>Kernel Linear</i> $C = 1$	53
Tabel 4.15. Hasil Pengujian <i>Kernel Linear</i> $C = 1000$	54
Tabel 4.16. Hasil Pengujian <i>Kernel Linear</i> $C = 10000000$	54
Tabel 4.17. Hasil Pengujian <i>Kernel RBF</i> $C = 1$	55
Tabel 4.18. Hasil Pengujian <i>Kernel RBF</i> $C = 1000$	55
Tabel 4.19. Hasil Pengujian <i>Kernel RBF</i> $C = 10000000$	56
Tabel 4.19. Rata-rata Akurasi <i>Kernel Linear</i>	566
Tabel 4.20 .Rata-rata Akurasi <i>Kernel RBF</i>	57

INTISARI

Klasifikasi Sentimen Terhadap Dataset Komentar Pemilu Presiden 2014 Pada Facebook Menggunakan Support Vector Machine

Media sosial adalah sarana komunikasi yang sangat populer saat ini jika dilihat dari jumlah pemakainya. Salah satu media sosial yang banyak digunakan yaitu Facebook dengan berbagai tujuan penggunaan misalnya sebagai tempat diskusi, komunitas pecinta olahraga tertentu, online shop, bahkan dalam bidang politik digunakan dimanfaatkan untuk mendapatkan dukungan dari para publik.

Tujuan penelitian ini adalah untuk menghasilkan suatu sistem klasifikasi sentimen berbasis *web*. *Dataset* yang digunakan adalah komentar status Facebook pemilihan umum presiden Indonesia tahun 2014. Metode *Support Vector Machine (SVM)* digunakan untuk proses klasifikasi komentar positif atau negatif. Tahap *preprocessing* yang dilakukan adalah *case folding*, *stemming*, *stopwords removal* dan tokenisasi. Data komentar kebanyakan mengandung *special character*, angka dan *emoticon*. Pada penelitian ini *emoticon* dikonversi ke bentuk teks sesuai dengan makna masing-masing. Pembobotan dilakukan menggunakan *Term Frequency Document Inverse Frequency (TF IDF)*. Metode lain yang diterapkan yaitu *feature selection* untuk menentukan nilai *term* yang akan digunakan dengan tujuan dapat menghasilkan nilai akurasi yang baik.

Pengujian menggunakan tiga jenis *dataset* dengan komposisi berbeda yaitu komentar positif dan negatif tidak sama (870 dan 130), komentar positif dan negatif masing – masing berjumlah 130, dan menggunakan data latih dengan 2024 komentar positif serta 313 komentar negatif. Evaluasi sistem dilakukan menggunakan *confusion matrix* untuk mengukur akurasi sistem dalam melakukan klasifikasi. Rata – rata akurasi menggunakan *kernel linear* terhadap *dataset* uji yaitu 73,84%, 50,37% dan 74,99%. Sedangkan rata-rata akurasi menggunakan *kernel RBF* adalah 76,53%, 51,83% dan 80,81%.

Kata Kunci: Klasifikasi Sentimen, *Text Mining*, *Support Vector Machine*, *Text Preprocessing*, *Confusion Matrix*, *TF IDF*, *Feature Selection*

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi saat ini sangat mempengaruhi kehidupan manusia. Dalam proses tersebut ditandai dengan kemunculan produk dari teknologi itu sendiri, salah satunya adalah media sosial yang digunakan untuk saling berkomunikasi (*communication*), selain itu juga memungkinkan pengguna untuk berbagi *file* (*sharing*), mengirimkan pesan instan (*instant messenger*), bahkan membentuk suatu kelompok sosial. Salah satu media sosial yang banyak digunakan adalah *Facebook*. Menurut hasil survey dari We Are Social masyarakat di Indonesia lebih banyak menggunakan *Facebook* dibandingkan dengan media sosial lainnya seperti *Twitter*, *Instagram*, *LinkedIn* dan media sosial lainnya (Lukman, 2015).

Secara umum *Facebook* adalah media sosial yang memudahkan pengguna untuk saling berkomunikasi dengan cepat meskipun dibatasi oleh letak geografis (Pardidi, 2013). Pengguna *Facebook* di Indonesia sangat beragam mulai dari kalangan pelajar, pelaku bisnis, bahkan para tokoh politik pun menggunakannya. Dalam dunia pendidikan, Pardidi (2013) melalui penelitiannya menyatakan bahwa para pelajar cenderung menggunakan *Facebook* untuk saling berkomunikasi dan berinteraksi dengan teman-temannya yang juga menggunakan *Facebook*. Dalam dunia bisnis, *Facebook* dimanfaatkan sebagai sarana untuk promosi barang dagangan secara *online* (Gemilang, 2011). Selain itu, sosial media juga berpengaruh terhadap dunia politik. Sebagai contoh pada pemilu presiden 2014 lalu, kedua pasang calon presiden dan calon wakil presiden menggunakan *Facebook* sebagai sarana untuk menerapkan strategi *political marketing* untuk mendapatkan dukungan dari para calon pemilih. Hal ini dilakukan mengingat kampanye politik di Indonesia membutuhkan biaya yang sangat besar, nilai tersebut dapat mencapai 9,6 triliun rupiah. Penggunaan *Facebook* dalam *political marketing* dianggap ampuh untuk menjangkau para calon pemilih dengan dana yang tidak besar (Juditha, 2015). Sanur (2014) beranggapan bahwa media sosial

mampu menjadi sarana dialog antara partai politik dan pendukungnya karena dapat menjaga keakurasian informasi.

Facebook memiliki beberapa fitur salah satunya adalah *wall*. Melalui fitur ini para pengguna dapat membuat sebuah pesan status yang kemudian *diposting* dan dapat dilihat oleh pengguna lainnya (Kusumaningtyas, 2010). Fitur *comment wall-to-wall* digunakan untuk memberikan komentar pada suatu status yang sebelumnya telah *diposting* pada *wall*. Kedua pasang calon presiden pun memanfaatkan fitur-fitur tersebut dalam menjalankan *political marketing*. Juditha (2015) menyampaikan bahwa pada masa kampanye terbuka pada 13 Juni – 4 Juli 2014, akun *Facebook* milik calon presiden Prabowo Subianto telah *memposting* 47 status, sedangkan akun milik Jokowi Widodo telah membuat *posting* sebanyak 46 status. Berdasarkan status-status yang telah *diposting* tersebut, para calon pemilih dapat berkomentar untuk menanggapi apa yang telah disampaikan masing-masing calon presiden. Sari (2015) dalam penelitiannya pada bidang sosial mengatakan bahwa komentar para calon pemilih tentunya sangat beragam, mulai dari komentar yang bersifat positif (setuju, mendukung) bahkan ada yang memberikan komentar negatif (berkomentar kasar, *bullying*). Penulis melakukan penelitian klasifikasi sentimen atau yang juga dikenal sebagai *opinion mining* untuk melakukan klasifikasi terhadap *dataset* komentar pemilu presiden 2014 (Rachmat C. dan Lukito, 2015) pada *Facebook* menggunakan *Support Vector Machine (SVM)*. Penelitian ini dilakukan karena *dataset* akan diklasifikasikan kedalam kategori positif atau negatif serta untuk melihat seberapa baik penggunaan *SVM* dalam melakukan klasifikasi.

1.2. Rumusan Masalah

1. Bagaimana sistem dapat menerapkan SVM untuk mengklasifikasikan sentimen terhadap *dataset* komentar pemilu presiden 2014 ?
2. Bagaimana keakuratan SVM dalam mengklasifikasi sentimen dataset komentar pemilu presiden 2014 jika dilihat dari segi *precision*, *recall* dan *f-measure* ?

1.3. Batasan Sistem

1. Dataset yang digunakan adalah dataset komentar pada *Facebook* terhadap pemilu presiden 2014 (Rachmat C. dan Lukito, 2015).
2. Metode yang digunakan adalah *SVM* dan tidak dibandingkan dengan metode lain.
3. Menggunakan *TF-IDF* untuk proses pembobotan token dalam suatu dokumen (Ghag dan Shah, 2014).
4. *Dataset* yang digunakan adalah bahasa Indonesia dan menggunakan *library Sastrawi stemmer* versi 1.0.2 (<https://github.com/sastrawi/sastrawi>).
5. Klasifikasi menggunakan *library SVM PHP-ML* versi 0.4.1 (<https://github.com/php-ai/php-ml>).

1.4. Tujuan Penelitian

Penelitian ini bertujuan untuk menghasilkan suatu sistem yang mampu mengklasifikasikan sentimen dengan baik terhadap dataset komentar pemilu presiden 2014 menggunakan *SVM*.

1.5. Metode Penelitian

Metode penelitian yang digunakan adalah sebagai berikut :

1. Studi Pustaka
Melalui studi pustaka penulis mencari dan mempelajari beberapa literatur yang memuat teori dan pembahasan terkait dengan metode yang digunakan untuk klasifikasi sentimen.
2. Konsultasi
Penulis melakukan konsultasi terhadap kedua dosen pembimbing untuk memperoleh saran terhadap sistem yang dibangun.
3. Pengumpulan Data
Data yang digunakan merupakan data komentar pada *Facebook* yang telah diberikan label positif atau negatif dan merupakan hasil penelitian Rachmat C. dan Lukito (2015).

4. Perancangan Sistem

Tahap ini merupakan perancangan *database* dan antarmuka sistem yang akan dibangun.

5. Pembuatan Sistem

Pada tahap ini akan dibangun suatu sistem sesuai dengan hasil perancangan *database* dan antarmuka.

6. Implementasi dan Pengujian Sistem

Pada bagian ini akan dilakukan pengujian terhadap sistem yang telah dibangun dengan menghasilkan suatu *output* berupa hasil klasifikasi.

7. Analisis Hasil dan Evaluasi Sistem

Pada tahap ini dilakukan analisis terhadap hasil yang diperoleh dari proses klasifikasi. Evaluasi sistem dilakukan untuk mengetahui seberapa akurat metode yang digunakan dalam melakukan klasifikasi.

1.6. Sistematika Penulisan

BAB I, Pendahuluan, bagian ini berisi latar belakang masalah, rumusan masalah, batasan sistem, tujuan penelitian, metodologi penelitian, sistematika penulisan. BAB II, Tinjauan Pustaka, bab ini dicantumkan beberapa penelitian yang pernah dilakukan serta landasan teori yang berkaitan dengan klasifikasi sentimen, seperti metode *SVM*, *text mining*, evaluasi sistem. BAB III, Analisis dan Perancangan Sistem, memaparkan tentang kebutuhan sistem, perancangan sistem, desain antarmuka serta perancangan *database*. BAB IV, Implementasi dan Analisis Sistem, berisikan pembahasan hasil perancangan sistem serta pemaparan mengenai analisa sistem. BAB V, Penutup, memuat kesimpulan dari hasil penelitian dan saran untuk mengembangkan sistem menjadi lebih baik.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan implementasi dan analisis hasil pengujian sistem, maka penulis membuat kesimpulan sebagai berikut :

1. Metode *Support Vector Machine* (SVM) dapat diterapkan pada sistem untuk melakukan klasifikasi sentimen menggunakan *dataset* komentar pemilu presiden 2014.
2. Berdasarkan hasil penelitian yang telah dibahas pada bab 4 mengindikasikan bahwa sistem belum mampu menentukan *hyperplane* yang memiliki *margin* terbaik untuk memisahkan kedua kelas.
3. Performa *kernel linear* dengan nilai $C = 1$ dan $C = 1000$ memberikan akurasi yang lebih baik pada pengujian komposisi data seimbang.
4. *Kernel RBF* dengan nilai $C = 1000$ mampu memberikan akurasi tertinggi sebesar 93,84% pada pengujian menggunakan data latih sebagai data uji dengan *feature selection* 10%.
5. Penggunaan nilai $C = 10000000$ mampu memberikan akurasi tertinggi pada pengujian dengan komposisi data seimbang yaitu 61,57% dengan *feature selection* 40%.

5.2. Saran

Berikut adalah saran-saran yang dapat digunakan untuk penelitian mengenai klasifikasi sentimen selanjutnya :

1. Menerapkan pemanfaatan *emoticon* yang lebih kompleks seperti pada penelitian Hogenboom et al (2013) untuk meningkatkan akurasi klasifikasi.
2. Memvisualisasikan data *vector* untuk mempermudah dalam menentukan jenis *kernel* dan *hyperplane* yang akan digunakan.

3. Menggunakan nilai C yang bervariasi untuk menentukan *margin* yang dapat membentuk model SVM dengan baik sehingga akurasi yang dihasilkan juga baik.

©UKDW

DAFTAR PUSTAKA

- Barber, I. 2010. *Bayesian Opinion Mining*. [Online]. Tersedia di : <http://phpir.com/bayesian-opinion-mining> [diunduh: 28 Feb 2016].
- Cortes, C. & Vapnik, V. 1995. *Support-Vector Networks. Machine Learning*, 20. Tersedia di: <http://www.springerlink.com/content/k238jx04hm87j80g/> [Diunduh: 25 Feb 2016].
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). *SMOTE: Synthetic minority over-sampling technique*. *Journal of Artificial Intelligence Research*, 16, 321–357. <http://doi.org/10.1613/jair.953>
- Gemilang, Rr. Dian A. 2011. *Peran Facebook sebagai Media Komunikasi Bisnis Online*. Universitas Pembangunan Nasional “Veteran”, Ilmu Komunikasi. Surabaya : UPN “Veteran”.
- Ghag, K., Shah, K. . 2014. *SentiTFIDF – Sentiment Classification using Relative Term Frequency Inverse Document Frequency*. *IJACSA*, 5, 38.
- Gries, D., & Schneider, F. B. (1997). *An introduction to Kolmogorov complexity and its applications. (Second edition)*. *Computers & Mathematics with Applications* (Vol. 34). [http://doi.org/10.1016/S0898-1221\(97\)90213-3](http://doi.org/10.1016/S0898-1221(97)90213-3)
- Hogenboom, A., Bal, D., Frasinca, F., Bal, M., de Jong, F., & Kaymak, U. (2013). Exploiting emoticons in sentiment analysis. *Proceedings of SAC’13, the 28th Annual ACM Symposium on Applied Computing*, 703–710. <http://doi.org/10.1145/2480362.2480498>
- Jiawei, H., Kamber, M., Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. San Francisco, CA, itd: Morgan Kaufmann. <http://doi.org/10.1016/B978-0-12-381479-1.00001-0>
- Juditha, Christian. 2015. *Political Marketing dan Media Sosial*. Balai Besar Pengkajian dan Pengembangan Komunikasi dan Informatika. Makassar.
- Kusumaningtyas, Ratih D. 2010. *Peran Media Sosial Online (Facebook) Sebagai Saluran Self Disclosure Remaja Putri di Surabaya*. Universitas Pembangunan Nasional “Veteran”, Ilmu Komunikasi. Surabaya : UPN “Veteran”.

- Lukman, E. 2015. *The latest numbers on web, mobile, and social media in Indonesia (INFOGRAPHIC)*. Diambil dari <https://www.techinasia.com/indonesia-web-mobile-data-start-2015>. [Diunduh: 20 Feb 2016].
- Manning, Christopher D., Raghavan, P., Schütze H. 2009. *An Introduction to Information Retrieval*.
- Novantirani, Anita, Mira K, Veronikha E. 2015. *Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine*. Universitas Telkom, Bandung.
- Pardidi, N. (2013). Motif Penggunaan Facebook Di Kalangan Siswa, 1(3), 479–491.
- Patil, M. G., Galande, M. V., Kekan, V., & Dange, M. K. (2014). Sentiment Analysis Using Support Vector Machine, 2607–2612. <http://doi.org/10.1109/I4CT.2014.6914200>
- Rachmat C, A., & Lukito, Y. (2015). Implementasi Crowdsourced Labelling Berbasis Web.
- Rachmat C, A., & Lukito, Y. (2016). Klasifikasi Sentimen Komentar Politik dari Facebook Page Menggunakan Naïve Bayes.
- Rajkumar, N. (2013). *Information Retrieval Using Second Order Co-occurrence PMI*, 9, 79–88.
- Sanur, Debora L. 2014. *Efektivitas Kampanye Pemilu Legislatif 2014*. http://berkas.dpr.go.id/pengkajian/files/info_singkat/Info%20Singkat-VI-7-I-P3DI-April-2014-43.pdf, diakses tanggal 10 February 2016.
- Saraswati, N. W. (2011). *Text Mining dengan Metode Naive Bayes Classifier dan Support Vector Machines untuk Sentiment Analysis*. Universitas UDAYANA, Teknik Elektro. Denpasar: Universitas UDAYANA.
- Sari, Ats Tsaqafi Ludiro. 2015. *Bullying Terhadap Calon Presiden pada Pemilihan Umum Presiden Tahun 2014*. Universitas Muhammadiyah Surakarta, Pendidikan Pancasila dan Kewarganegaraan. Surakarta : UMS.
- Verma, V. K., Ranjan, M., & Mishra, P. (2015). *Text mining and information professionals: Role, issues and challenges*, 133–137. <http://doi.org/10.1109/ETTLIS.2015.7048186>

Wibisono, Y. (2005). *Klasifikasi Berita Berbahasa Indonesia menggunakan Naïve Bayes Classifier*. In Seminar Nasional Matematika. Diakses pada tanggal 24 Feb 2016 (Vol. 27).

Wulandini, F., & Nugroho, A. S. (2009). *Text Classification Using Support Vector Machine for Webmining Based Spatio Temporal Analysis of the Spread of Tropical Diseases*. Proc. International Conference on Rural Information and Communication Technology.

©UKDW