

**ANALISIS SENTIMEN TWITTER HASIL PILKADA
DKI JAKARTA 2017 MENGGUNAKAN METODE
*SUPPORT VECTOR MACHINE***

TUGAS AKHIR

Diajukan untuk Memenuhi Sebagian Persyaratan
Mencapai Derajat Sarjana Teknik Informatika



Disusun oleh :
Sem Raka Widyamahendra
14 07 07688

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2018

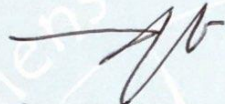
HALAMAN PENGESAHAN
TUGAS AKHIR BERJUDUL

**ANALISIS SENTIMEN TWITTER HASIL PILKADA DKI JAKARTA
2017 MENGGUNAKAN METODE SUPPORT VECTOR MACHINE**

Disusun oleh :
Sem Raka Widyamahendra
NIM : 14 07 07688

dinyatakan telah lengkap memenuhi syarat pada tanggal
April 2018

Dosen Pembimbing I,



(Ir. A. Djoko Budiyanto,
M.Eng., Ph.D)

Dosen Pembimbing II,



(Dra. Ernawati, M.T.)

Tim Penguji :

Dosen Penguji I,



(Ir. A. Djoko Budiyanto, M.Eng., Ph.D)

Dosen Penguji II,



(Thomas Adi Purnomo Sidhi,
S.T., M.T.)

Dosen Penguji III,



(Dr. Phil. Lukas S.
Ispandriarno)

Yogyakarta, April 2018
Universitas Atma Jaya Yogyakarta
Fakultas Teknologi Industri
Dekan,



(Dr. Drs. A. Teguh Siswantoro, M.Sc.)

KATA PENGANTAR

Puji dan Syukur kepada Tuhan Yang Maha Esa, atas berkat, rahmat, kemudahan, petunjuk dan kekuatan yang diberikan-Nya, sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul **"Analisis Sentimen Twitter Terhadap Hasil Pilkada DKI Jakarta 2017 Menggunakan Metode Support Vector Machine"**, merupakan prasyarat untuk mencapai gelar Sarjana Teknik Informatika.

Dalam penyusunan penulisan skripsi ini tidak terlepas dari pihak-pihak yang ikut membantu baik secara langsung maupun secara tidak langsung. Maka dalam kesempatan ini dengan penuh rasa hormat, penulis ingin menyampaikan rasa terima kasih kepada:

1. Tuhan Yesus Kristus yang selalu memberkati dan menyertai penulis dalam mengerjakan skripsi ini.
2. Universitas Atma Jaya Yogyakarta, Universitas yang telah mendidik dan membina penulis sehingga Tugas Akhir ini dapat diselesaikan.
3. Ir. A. Djoko Budiyanto SHR, M.Eng., Ph.D., selaku dosen pembimbing I yang selalu memberikan bimbingan yang baik dan memberikan ilmu-ilmu yang baru.
4. Dra. Ernawati, M.T. selaku dosen pembimbing II yang selalu membimbing dan mengarahkan dalam proses penulisan skripsi ini.
5. Dr. Phil. Lukas S. Ispandriarno yang selalu membimbing dan memberi ilmu baru tentang politik dalam pembentukan interpretasi skripsi ini.

6. Ayah, Ibu, Rut, dan dede Pua tercinta yang selalu memberikan dukungan dan doa setiap saat.
7. Leo, Andjar, Krisna, Dinan, Fibra, Resky, Icus, Rizky, Rista, Tyas, Anggrek, Agon, David, Gens, Audine dan semua teman-teman kampus yang ikut mendukung dalam proses pembuatan skripsi ini.
8. Gede, Nanto, Albert, Surya, Mas Albert, Jorghy, dan semua teman-teman dari kost yang selalu mendukung dalam pembuatan skripsi ini.
9. Ito, Alex, Maria, April, Jessica, Robert, Josef, Frisca, Shanti, Yumiko, dan semua teman-teman dari SS Pascasarjana yang selalu ada dalam suka dan duka dalam proses pembuatan skripsi ini.
10. Bule, Hardi, Faisal, Rudy, Aris, Rajtsan, Ando, Novi, Ica, dan semua teman-teman penulis yang ada di Bogor yang ikut mendukung dalam proses pembuatan skripsi ini.
11. Semua teman-teman dari Kelompok KKN 70 Klepu: Sonata, Fanny, Agri, Nina, Adit, Alan, Shinta, Winda, dan Windu yang ikut memberi dukungan moral dalam proses pembuatan skripsi ini.

Dengan segenap pengetahuan yang dimiliki, penulisan skripsi ini dibuat dengan harapan bisa bermanfaat dan menambah ilmu bagi penulis dan pembaca.

Yogyakarta, April 2018

Penulis

Sem Raka Widyamahendra

DAFTAR ISI

HALAMAN PENGESAHAN	Error! Bookmark not defined.
KATA PENGANTAR	ii
DAFTAR ISI	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	ix
INTISARI	xi
BAB I	1
PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian.....	4
1.5. Sistematika Penulisan Tugas Akhir.....	5
BAB II	7
TINJAUAN PUSTAKA DAN LANDASAN TEORI	7
2.1. Tinjauan Pustaka.....	7
2.2. Landasan Teori.....	15
a. Data Mining.....	15
b. Analisis Sentimen.....	15
c. Klasifikasi.....	16
d. Twitter.....	17
e. <i>Support Vector Machine</i>	17
f. Kernel pada <i>Support Vector Machine</i>	20
g. <i>MultiClass Support Vector Machine</i>	21
h. <i>Pre-Proccesing</i>	23
i. <i>SentiStrength</i>	23
j. <i>K-fold Cross Validation</i>	25
k. N-Gram.....	26
BAB III	27
METODOLOGI PENELITIAN	27
3.1. Studi Literatur.....	28
3.2. Pembentukan Dataset.....	28
a. Pengumpulan Data.....	28
b. <i>Pre-processing</i> Data.....	29
3.2. Pembentukan Model.....	31
a. Menentukan <i>Class Attribute</i>	31
b. Klasifikasi.....	32
3.3. Evaluasi Model.....	33
a. Menentukan Fitur.....	33

b.	Analisa Model.....	33
c.	Penggunaan Cross Validation.....	35
d.	Perbandingan Metode Support Vector Machine dengan Metode Klasifikasi Lain.....	36
3.4.	Interpretasi Data.....	36
a.	Mengidentifikasi Isu-Isu yang berpengaruh terhadap hasil Pilkada DKI Jakarta 2017.....	37
b.	Pengelompokan Sentimen dan Deteksi Emosi.....	37
BAB IV	38
HASIL DAN PEMBAHASAN	38
4.1.	Dataset.....	38
4.2.	Menentukan Class Attribute dengan SentiStrength	40
4.3.	Pengujian Dataset.....	42
c.	Hasil Pengujian dengan metode <i>Support Vector Machine</i>	43
d.	Hasil Pengujian dengan metode <i>Maximum Entropy</i>	49
c.	Perbandingan Hasil Pengujian.....	52
4.4.	Evaluasi Dataset.....	54
4.5.	Faktor yang mempengaruhi hasil Pilkada DKI Jakarta 2017.....	56
a.	Isu SARA.....	58
b.	Isu Radikalisme.....	59
c.	Isu Kampanye.....	61
d.	Isu Hoax.....	64
4.6.	Analisis Sentimen terhadap Isu yang mempengaruhi hasil Pilkada DKI Jakarta 2017.....	66
a.	Analisis Sentimen terhadap Isu SARA.....	66
b.	Analisis Sentimen terhadap Isu Radikalisme.....	69
c.	Analisis Sentimen terhadap Isu Kampanye.....	72
d.	Analisis Sentimen terhadap Isu hoax.....	75
4.7.	Analisa Hasil dari Opini Masyarakat.....	77
BAB V	79
KESIMPULAN DAN SARAN	79
5.1.	Kesimpulan.....	79
5.2.	Saran.....	80
DAFTAR PUSTAKA	81
LAMPIRAN	85

DAFTAR GAMBAR

Gambar 2.1 SVM berusaha untuk menemukan hyperplane terbaik yang memisahkan kedua kelas -1 dan +1.....	18
Gambar 2.2 Hyperplane terbentuk di antara class-1 dan +1.....	20
Gambar 3.1 Metodologi Penelitian.....	27
Gambar 3.2 Contoh hasil <i>grab data</i>	29
Gambar 3.3 Proses Steming.....	31
Gambar 3.4 <i>Confusion Matrix</i>	34
Gambar 3.5 Kurva ROC.....	35
Gambar 4.1 <i>Class Attribute</i> pada dataset.....	41
Gambar 4.2 Perbandingan performa kernel linear dan RBF.....	44
Gambar 4.3 Perbandingan AUC kernel linear dan RBF.....	45
Gambar 4.4 kurva ROC kelas positif dengan linear.....	46
Gambar 4.5 kurva ROC kelas netral dengan linear.....	47
Gambar 4.6 kurva ROC kelas negatif dengan linear.....	47
Gambar 4.7 kurva ROC kelas positif dengan kernel RBF..	48
Gambar 4.8 kurva ROC kelas netral dengan kernel RBF....	48
Gambar 4.9 kurva ROC kelas negatif dengan kernel RBF..	49
Gambar 4.10 kurva ROC kelas positif dengan <i>Maximum Entropy</i>	51
Gambar 4.11 kurva ROC kelas netral dengan <i>Maximum Entropy</i>	51
Gambar 4.12 kurva ROC kelas negatif dengan <i>Maximum Entropy</i>	52

Gambar 4.13 Hasil perbandingan klasifikasi metode <i>Support Vector Machine</i> dan <i>Maximum Entropy</i>	53
Gambar 4.14 Grafik perbandingan k-fold cross validation	55
Gambar 4.15 Kumpulan Opini Masyarakat terhadap hasil Pilkada DKI Jakarta 2017.....	56
Gambar 4.16 Opini Masyarakat.....	57
Gambar 4.17 kelas sentimen isu SARA.....	67
Gambar 4.18 Polaritas emosi isu SARA.....	67
Gambar 4.19 kelas sentimen isu radikalisme.....	70
Gambar 4.20 Polaritas emosi isu radikalisme.....	70
Gambar 4.21 kelas sentimen isu kampanye.....	73
Gambar 4.22 Polaritas emosi isu kampanye.....	73
Gambar 4.23 kelas sentimen isu hoax.....	75
Gambar 4.24 Polaritas emosi isu hoax.....	76

DAFTAR TABEL

Tabel 2.1 Algoritma Support Vector Machine untuk klasifikasi data.....	13
Tabel 2.2 Perbandingan algoritma Support Vector Machine dengan metode lainnya.....	14
Tabel 2.3 Kernel pada Support Vector Machine.....	21
Tabel 2.4 Contoh hasil SentiStrength.....	24
Tabel 2.5 Aturan dalam SentiStrength.....	25
Tabel 3.1 Batasan Pengambilan data Twitter.....	28
Tabel 3.2 Rangkaian proses data cleansing.....	30
Tabel 3.3 Contoh Kamus Senti-Strength.....	32
Tabel 3.4 <i>k-fold cross validation</i>	36
Tabel 4.1 Contoh tweet yang dikumpulkan.....	38
Tabel 4.2 Proses <i>preprocessing</i>	39
Tabel 4.3 Contoh hasil pembobotan SentiStrenght.....	40
Tabel 4.4 Contoh hasil dari <i>Class Attribute</i>	42
Tabel 4.5 Hasil pengujian dengan metode <i>Support Vector Machine</i> menggunakan kernel linear.....	43
Tabel 4.6 Hasil pengujian dengan metode <i>Support Vector Machine</i> menggunakan kernel RBF.....	43
Tabel 4.7 Hasil AUC kernel linear dan RBF.....	45
Tabel 4.8 Hasil pengujian dengan metode <i>Maximum Entropy</i>	49
Tabel 4. 9 Nilai AUC dari metode <i>Maximum Entropy</i>	50
Tabel 4.10 Hasil <i>k-fold cross validation</i>	54
Tabel 4.11 Contoh <i>tweet</i> isu SARA.....	58

Tabel 4.12 Contoh <i>tweet</i> isu radikalisme.....	60
Tabel 4.13 Contoh <i>tweet</i> kampanye.....	62
Tabel 4.14 Contoh <i>tweet</i> hoax.....	65
Tabel 4.15 Contoh <i>tweet</i> dengan emosi marah dan sentimen negatif terhadap isu SARA.....	68
Tabel 4.16 Contoh <i>tweet</i> dengan emosi marah dan sentimen negatif terhadap isu Radikalisme.....	71
Tabel 4.17 Contoh <i>tweet</i> dengan emosi terkejut dan sentimen positif terhadap isu Kampanye.....	74
Tabel 4.18 Contoh <i>tweet</i> dengan emosi marah dan sentimen negatif terhadap isu Hoax.....	77

ANALISIS SENTIMEN TWITTER HASIL PILKADA DKI JAKARTA 2017 MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

INTISARI

Sem Raka Widyamahendra (140707688)

Microblogging merupakan tren alat komunikasi yang sangat populer. Twitter merupakan salah satu platform situs web microblogging yang populer. Jutaan pengguna berbagi tweets setiap harinya. Oleh karena itu Twitter adalah sumber data yang kaya untuk analisis sentimen. Penelitian ini bertujuan untuk mengumpulkan opini publik terhadap hasil Pilkada DKI Jakarta 2017 yang disampaikan di Twitter.

Data yang dikumpulkan berjumlah 12670 tweet. Data tersebut dibentuk menjadi model klasifikasi dengan melakukan preprocessing dan pembobotan menggunakan SentiStrength untuk mengelompokkan ke dalam tiga kelas sentimen, yaitu: positif, netral, dan negatif. Metode Support Vector Machine digunakan untuk melakukan klasifikasi atas model yang sudah dibentuk. Hasil dari klasifikasi diinterpretasi untuk menemukan opini yang berhubungan dengan hasil Pilkada dan mengolahnya menjadi faktor-faktor yang berpengaruh terhadap hasil Pilkada.

Hasil pengujian terhadap model klasifikasi yang dibangun memperlihatkan bahwa metode Support Vector Machine dengan kernel RBF mampu menghasilkan akurasi sebesar 89%, recall sebesar 79,3%, presisi sebesar 84%, dan F-Measure sebesar 81%. Hasil dari interpretasi data ditemukan empat isu yang berpengaruh terhadap hasil Pilkada, faktor tersebut adalah isu SARA, isu radikalisme, isu kampanye, dan isu hoax.

Kata Kunci: *analisis sentimen, pilkada dki jakarta 2017, klasifikasi, support vector machine, twitter*