

**PENERAPAN ANALISIS SENTIMEN PADA PENGGUNA
TWITTER MENGGUNAKAN METODE *K-NEAREST
NEIGHBOR***

Skripsi
Untuk Memenuhi Sebagian Persyaratan
Mencapai Derajat Sarjana S-1
Program Studi Teknik Informatika



Disusun Oleh :

Akhmad Deviyanto

12651085

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA
PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2018



PENGESAHAN TUGAS AKHIR

Nomor : B-1019/Un.02/DST/PP.00.9/02/2018

Tugas Akhir dengan judul : Penerapan Analisis Sentimen pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor

yang dipersiapkan dan disusun oleh:

Nama : AKHMAD DEVIYANTO
Nomor Induk Mahasiswa : 12651085
Telah diujikan pada : Senin, 26 Februari 2018
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR

Ketua Sidang

Muhammad Didik Rohmad Wahyudi, S.T., MT.
NIP. 19760812 200901 1 015

Penguji I

Dr. Agung Fatwanto, S.Si., M.Kom.
NIP. 19770103 200501 1 003

Penguji II

Agus Mulyanto, S.Si., M.Kom.
NIP. 19710823 199903 1 003

Yogyakarta, 26 Februari 2018
UIN Sunan Kalijaga
Fakultas Sains dan Teknologi

DEKAN



Dr. Murtong, M.Si.
NIP. 19691212 200003 1 001



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Permohonan

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Akhmad Deviyanto

NIM : 12651085

Judul Skripsi : Penerapan Sentiment Analysis pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Teknik Informatika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 19 Februari 2018

Pembimbing

M. Didik R. Wahyudi, M.T.

NIP 19760812 200901 1 015

SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini:

Nama : Akhmad Deviyanto
NIM : 12651085
Program Studi : Teknik Informatika
Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi dengan judul **“Penerapan Sentiment Analysis pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor”** tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 19 Februari 2018

Yang Menyatakan



Akhmad Deviyanto
NIM. 12651085

KATA PENGANTAR

Alhamdulillah Robbil 'Alamin. Segala puji syukur ke hadirat Allah SWT yang telah memberikan pertolongan pada hambaNya yang lemah ini dalam setiap keadaan, termasuk dalam penelitian dan penulisan skripsi ini. Sholawat dan salam teruntuk Baginda Agung Nabi Muhammad SAW Sang *Insan Kamil* yang paling sempurna fisik dan akhlaknya yang sungguh kita rindukan perjumpaan dengan beliau di dunia terlebih di akhirat kelak.

Pelaksanaan penelitian dan penyusunan skripsi ini merupakan salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika di Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

Selanjutnya penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Ayahanda Rebi dan Ibu Warpinah serta dua adikku Fuad Hasyim, S.H. dan Fitri Lutfiani tercinta yang sejak lahir, sekarang, dan kelak akan senantiasa memberikan do'a, dukungan, dan pengorbanan serta menjadi sumber motivasi dan inspirasi.
2. Prof. Drs. Yudian Wahyudi, M.A., Ph.D. selaku rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Dr. Murtono, M.Si. selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Bapak Dr. Bambang Sugiantoro selaku Ketua Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta.

5. Bapak Agus Mulyanto, S.SI, M.Kom. selaku dosen pembimbing akademik yang telah senantiasa meluangkan waktu untuk memberikan arahan mengenai akademisi.
6. Bapak M. Didik R. Wahyudi, M.T. selaku dosen pembimbing yang telah banyak memberikan dukungan serta pengarahan demi kelancaran pelaksanaan penulisan skripsi.
7. Bapak dan Ibu Dosen Teknik Informatika UIN Sunan Kalijaga Yogyakarta yang telah memberikan ilmu dari awal perkuliahan.
8. Teman-teman Program Studi Teknik Informatika angkatan 2012 (KATAK) yang tidak saya sebutkan satu persatu namun tidak mengurangi rasa kagum dan bangga kepada kalian yang telah banyak memberikan bantuan dan memberikan warna dan kenangan selama masa kuliah.
9. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah memberikan dukungan dan semangat sehingga penulisan dapat menyelesaikan laporan skripsi ini.

Akhirnya, tiada gading yang tak retak. Penulis menyadari masih banyak kekurangan dan kelemahan dalam pelaksanaan dan penyusunan skripsi ini. Semoga penelitian ini dapat menjadi pengalaman yang berharga bagi penulis dalam mempersiapkan diri menghadapi persaingan di dunia kerja dan bermanfaat untuk masyarakat yang lebih luas.

Yogyakarta, 13 Februari 2018

Penulis

HALAMAN PERSEMBAHAN

PERSEMBAHAN

Aku persembahkan karya sederhana ini untuk:

- Ibu dan Bapakku, yang telah mendukungku, memberiku motivasi dan arahan dalam segala hal serta memberikan kasih sayang yang teramat besar dan semua perjuangan yang tak mungkin bisa ku balas dengan apapun.
- Adik-adikku tercinta yang selalu memberikan dukungan agar cepat selesai skripsi.
- Teman-teman TIF Mandiri 2012, *jazaakumullaah ahsana jazaa.*

MOTTO

“ الرَّاحِمُونَ يَرْحَمُهُمُ الرَّحْمَنُ ارْحَمُوا مَنْ فِي الْأَرْضِ يَرْحَمَكُم مَّنْ فِي السَّمَاءِ ”

“Orang-orang yang mengasihani akan dikasihani oleh Ar Rahman.
Berkasih sayanglah kepada siapa pun yang ada di bumi, niscaya Yang
ada di langit akan mengasihani kalian.”

(Hadits Musalsal Bil Awwaliyyah)

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
SURAT PERSETUJUAN SKRIPSI / TUGAS AKHIR	iii
PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
HALAMAN PERSEMBAHAN	vii
MOTTO	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
INTISARI	xiii
ABSTRACT	xiv
BAB I 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 Manfaat Penelitian	3
1.6 Keaslian Penelitian	3
1.7 Sistematika penelitian	3
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	5
2.1 Tinjauan Pustaka	5
2.2 Landasan Teori	7
2.2.1 Analisis Sentimen	8
2.2.2 <i>Supervised Learning</i>	8
2.2.3 <i>Text Mining</i>	9
2.2.4 <i>Text Preprocessing</i>	9
2.2.5 Pembobotan Kata TF-IDF	11
2.2.6 Algoritma KNN (<i>K Nearest Neighbor</i>)	12
2.2.7 <i>Confusion Matrix</i>	14
BAB III METODE PENELITIAN	16
3.1 Metode Penelitian	16
3.2 Peralatan Penelitian	16
3.2.1 Perangkat Keras	16
3.2.2 Perangkat Lunak	16

3.3	Tahap - Tahap Penelitian	18
3.3.1	Pengumpulan Data	18
3.3.2	<i>Preprocessing Data</i>	22
3.3.3	Analisis Sentimen dengan <i>K-Nearest Neighbor</i>	23
BAB IV HASIL DAN PEMBAHASAN		25
4.1	Objek Penelitian	25
4.2	Proses Analisis Sentimen	25
4.2.1	<i>Preprocessing</i>	26
4.2.1.1	<i>Case Folding</i>	26
4.2.1.2	Tokenisasi	28
4.2.1.3	<i>Stopword Removal</i>	30
4.2.1.4	<i>Stemming</i>	32
4.2.2	Pembobotan Kata TF-IDF	34
4.2.3	Perhitungan <i>Cosine Similarity</i>	39
4.2.4	Penentuan Nilai Sentimen	44
4.3	Rancang Antar Muka Sistem.....	46
BAB V PENUTUP		54
5.1	Kesimpulan	54
5.2	Saran	54
DAFTAR PUSTAKA		55
LAMPIRAN		57

DAFTAR TABEL

Tabel 2.1 Daftar Penelitian Terdahulu	6
Tabel 2.2 <i>Confusion Matrix</i>	14
Tabel 4.1 Hasil Proses <i>Case Folding</i>	27
Tabel 4.2 Hasil Proses <i>Tokenisasi</i>	28
Tabel 4.3 Hasil Proses <i>Stopword Removal</i>	31
Tabel 4.4 hasil Proses <i>Stemming</i>	33
Tabel 4.5 Nilai TF-IDF Dokumen	37
Tabel 4.6 Nilai Sentimen Dokumen	39
Tabel 4.7 Hasil <i>Cosine Similarity</i> (Langkah 2)	41
Tabel 4.8 Hasil <i>Cosine Similarity</i> (Langkah 3)	42
Tabel 4.9 Hasil <i>Cosine Similarity</i> (Langkah 4)	44

DAFTAR GAMBAR

Gambar 2.1 Proses <i>Text Mining</i>	9
Gambar 2.2 Rumus TF-IDF	11
Gambar 3.1 Alur Pengambilan Data <i>Tweet</i>	19
Gambar 3.2 Halaman Pencarian Lanjutan Twitter	20
Gambar 3.3 Tampilan <i>Running Twitterscraper</i> pada <i>Command Promt</i>	21
Gambar 3.4 Tampilan File JSON pada MS Excel	21
Gambar 4.1 <i>Flowchart</i> Sistem Analisis Sentimen	25
Gambar 4.2 Halaman Awal	46
Gambar 4.3 Tampilan Menu Data Latih	47
Gambar 4.4 Tampilan Submenu Tambah Data Latih	47
Gambar 4.5 Submenu Edit pada Menu data Latih	48
Gambar 4.6 Tampilan Menu Data Uji	49
Gambar 4.7 Proses Analisis Sentimen Telah Selesai Dilakukan	49
Gambar 4.8 hasil Perhitungan sentimen untuk Satu Data Uji	50
Gambar 4.9 Halaman Menu Setting KNN	51
Gambar 4.10 Halaman Menu Akurasi	52
Gambar 4.11 Nilai Confusion Matrix Sistem	52
Gambar 4.12 Grafik Nilai <i>Confusion Matrix</i>	53

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PENERAPAN ANALISIS SENTIMEN PADA PENGGUNA TWITTER MENGGUNAKAN METODE *K-NEAREST NEIGHBOR*

Akhmad Deviyanto
12651085

INTISARI

Media sosial dewasa ini menjadi tempat untuk mencurahkan opini dari penggunanya terhadap berbagai hal salah satunya politik. Tema politik yang dalam beberapa waktu lalu menjadi *trending* di media sosial Twitter adalah tentang Pemilihan Gubernur (Pilgub) DKI Jakarta tahun 2017. Maka, pada penelitian ini akan dilakukan analisis sentimen para pengguna Twitter terhadap para kandidat gubernur di pilgub tersebut.

Peneliti menggunakan algoritma *K-Nearest Neighbor* dengan pembobotan kata TF-IDF dan fungsi *Cosine Similarity* untuk membangun sistem analisis sentimen. Nilai k dipakai adalah 3, 5, 7, 9, dan 15. Data Twitter yang diambil adalah sebanyak 2000 data dengan perincian 1500 sebagai data latih dan 500 sebagai data uji. Data diambil dari Twitter menggunakan salah satu *package* Python yaitu *Twitterscraper*. Data yang telah diperoleh kemudian akan melewati proses *preprocessing* terlebih dulu sebelum dilakukan analisis sentimen.

Metode *Confusion matrix* digunakan untuk mengukur kinerja algoritma. Dari metode ini akan diketahui nilai akurasi, presisi, dan *recall* dari algoritma KNN yang telah diterapkan pada sistem. Berdasarkan hasil pengujian sistem, diperoleh hasil akurasi terbesar adalah sebesar 67,2% ketika nilai $k = 5$, nilai presisi terbesar adalah 56,94% ketika $k = 5$, dan *recall* terbesar 78,24% ketika $k = 15$.

Kata kunci : analisis sentimen, *Twitterscraper*, *K-Nearest Neighbor*, *Cosine Similarity*, TF-IDF

IMPLEMENTATION OF SENTIMENT ANALYSIS FROM TWITTER USER USING K-NEAREST NEIGHBOR METHOD

Akhmad Deviyanto
12651085

ABSTRACT

Social media today is a place to express the opinions of its users about various things, including about politic. Political issue which sometime ago be a trend in social media Twitter is about the 2017 Election of Governor (Pilgub) of Special Capital Region Jakarta (DKI). So, this research will be explaining about sentiment analysis of Twitter users to the governor candidates in Pilgub.

Researcher uses the K-Nearest Neighbor algorithm with Cosine Similarity's distance calculation and TF-IDF term weighting in order to develop the sentiment analysis system. The k values used are 3, 5, 7, 9, and 15. The Twitter data used in this system is 2000 data, which is 1500 as training data and 500 as test data. The data taken from the Twitter using one of Python packages called Twitterscraper. The data that has been obtained then will go through the preprocessing phase first before the sentiment analysis is started.

Confusion matrix is a method to measure the performance of algorithm. With this method, can be obtained the accuracy, precision, and recall percentages of KNN algorithm that has been implemented on the system. Based on the test results, the highest accuracy is 67,2% when $k = 5$, the highest precision is 56,94% with $k = 5$, and the highest recall is 78,24% when $k = 15$.

Keywords : sentiment analysis, Twitterscraper, K-Nearest Neighbor, Cosine Similarity, TF-IDF

BAB I

PENDAHULUAN

1.1 Latar Belakang

Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) telah melakukan survei pada tahun 2016. Ada sekitar 132,7 juta pengguna internet di Indonesia (naik secara signifikan dari tahun 2014 sebanyak 88 juta pengguna). Dari jumlah tersebut, sebanyak 97,4% (129,2 juta) adalah pengguna yang menggunakan internet untuk mengakses media sosial. Lima media sosial dengan pengguna terbanyak adalah Facebook, Instagram, Youtube, Google Plus, dan Twitter.

Social media Twitter memiliki 7,2 juta pengguna di Indonesia. Setiap harinya, paling tidak ada 4,1 juta *tweet* yang berasal dari Indonesia (CNN Indonesia, 2016). Jumlah yang cukup besar tersebut merupakan cuitan para penggunanya tentang banyak hal: pendidikan, hiburan, pekerjaan, dan termasuk juga politik.

Salah satu isu politik yang menjadi *trending topic* di Twitter pada tahun 2017 adalah tentang pemilihan gubernur Jakarta (Pilkada DKI). Terdapat tiga pasangan calon pada Pilkada DKI 2017. Ketiga paslon tersebut adalah Agus Harimurti Yudhoyono - Sylviana Murni, Basuki Tjahaja Purnama - Djarot Saiful Hidayat, Anies Rasyid Baswedan - Sandiaga Salahuddin Uno.

Sentimen Analisis atau *opinion mining* adalah jenis *natural language* yaitu pengolahan kata untuk melacak *mood* masyarakat tentang produk atau topik

tertentu. Analisis sentimen, disebut *opinion mining*. (G.Vinodhini, M.Chandrasekaran 2012).

Penulis pada penelitian ini akan melakukan analisis sentimen para pengguna Twitter terhadap ketiga pasangan kandidat pada Pilkada DKI 2017. Dengan input berupa data *tweet* dalam Bahasa Indonesia, akan dilakukan klasifikasi dengan algoritma KNN (K-Nearest Neighbor) untuk menentukan apakah *tweet* tersebut bersentimen positif atau negatif.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, dapat dirumuskan permasalahan yang akan diselesaikan dalam penelitian ini adalah bagaimana mengamati dan menganalisis opini pengguna Twitter mengenai Pilkada DKI 2017 dan kemudian mengklasifikasikan sentimen data *tweets* tersebut.

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Data yang digunakan untuk penelitian ini adalah data *tweets* dalam Bahasa Indonesia.
2. Data *tweets* yang akan dinilai sentimennya adalah *tweet* yang mengandung salah satu *keyword* berikut : ahy, ahok, atau anies.
3. Jumlah data *tweets* yang digunakan adalah 2000 data yang telah difilter yang diambil dari tanggal 01 Januari 2017 sampai dengan 31 Januari 2017
4. Metode yang digunakan untuk pengklasifikasian dalam penelitian ini adalah metode K-Nearest Neighbor (KNN)

5. Hasil keluaran pengklasifikasian sentimen pada penelitian ini dengan menggunakan metode KNN berupa sentimen positif dan negatif.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengimplementasikan algoritma KNN (*K - Nearest Neighbor*) dalam analisis sentimen pengguna Twitter dengan topik Pilkada DKI 2017 dan kemudian mengetahui nilai akurasi, presisi, dan *recallnya*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik kepada penulis maupun pembaca tentang gambaran sentimen para pengguna Twitter terhadap para kandidat Pilkada DKI 2017.

1.6 Keaslian Penelitian

Penelitian terkait analisis sentimen Twitter hingga saat ini telah banyak dilakukan. Namun, berdasarkan referensi dan tinjauan pustaka, penelitian yang diajukan sebagai Tugas Akhir S1 pada program studi Teknik Informatika Fakultas Sains Dan Teknologi UIN Sunan Kalijaga mengenai analisis sentimen Twitter dengan menggunakan algoritma *K-Nearest Neighbor* belum pernah dilakukan.

1.7 Sistematika Penulisan

Laporan penelitian tugas akhir ini disusun secara sistematis dibagi dalam beberapa bab. Penyusunan laporan tugas akhir ini dibagi dalam lima bab dengan rincian sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini dijelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, keaslian penelitian, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Tinjauan pustaka dan landasan teori berisi tentang penelitian terdahulu dan teori-teori dasar yang digunakan penulis untuk membangun sistem.

BAB III METODE PENELITIAN

Metode penelitian berisi tentang penjelasan mengenai langkah-langkah yang harus dilakukan untuk mencapai tujuan dan kesimpulan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini dijelaskan tentang hasil analisis penelitian dan pembahasan mengenai penerapan metode yang digunakan.

BAB V PENUTUP

Bab terakhir ini berisikan kesimpulan dari hasil analisis serta saran untuk perbaikan penelitian selanjutnya.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, diperoleh kesimpulan bahwa penelitian tentang analisis sentimen pengguna Twitter terhadap topik Pilkada DKI 2017 dengan menggunakan metode *K Nearest Neighbor* telah berhasil dilakukan. Hasil akurasi terbesar adalah 67,2% dengan nilai $k=5$. Sedangkan nilai presisi tertinggi sebesar 56,94% saat $k=5$ dan *recall* terbesar 78,24 % ketika $k=15$.

5.2 Saran

Penelitian ini sudah barang tentu tidak terlepas dari kekurangan dan kelemahan. Oleh karena itu, peneliti menyarankan adanya perbaikan-perbaikan :

1. Penelitian selanjutnya diharapkan menggunakan dua atau lebih metode atau algoritma untuk meningkatkan keakuratan sistem.
2. Sistem belum dapat mengidentifikasi slang word / kata-kata yang disingkat yang berpengaruh pada keakuratan perhitungan sehingga diperlukan adanya fitur untuk mengidentifikasi masalah tersebut.
3. Sistem yang dibangun masih terpisah antara modul untuk mengambil data dari Twitter dan modul untuk perhitungan sentimen analisis. Pada penelitian selanjutnya diharapkan dapat dibuat sistem yang lebih terintegrasi.

DAFTAR PUSTAKA

- Agarwal, A. X. 2011. *Sentiment Analysis of Twitter Data*. 1-9.
- Bukhari, Varian Habbie. 2015. *Sentiment Analysis Menggunakan K-Nearest Neighbor dengan Perbandingan Fungsi Jarak (Studi Kasus : Twitter Indosat dan Telkomsel)*. Skripsi S1 Universitas Widyatama Bandung.
- Ganisaputra, Y., & Tan, R. 2013. *Pembuatan Aplikasi Datamining Facebook dan Twitter dengan Naïve Bayes Classifier*. Jurnal Informatika, Vol. 9, No. 2, 173188.
- Ginanjari, A. M., & Lubis, R. 2012. *Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit*. Jurnal Komputer dan Informatika (KOMPUTA) Edisi I, Volume I, 53-57.
- Hidayatullah, Ahmad Fathan. 2014. *Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Data Twitter Menggunakan Naive Bayes Classifier*. MCs Thesis, Yogyakarta: Universitas Gajah Mada.
- Ilyas, Husni. 2017. *Cosine Similarity Antar Dokumen: Sebuah Contoh*. <https://komputasi.files.wordpress.com/2011/01/cosine-similarity-antar-dokumen-sebuah-contoh.pdf>.
- Indriyono, B. V., Utami, E., & Sunyoto, A. 2015. *Pemanfaatan Algoritma Porter Stemmer Untuk Bahasa Indonesia Dalam Proses Klasifikasi Jenis Buku*. Jurnal Buana Informatika, Volume 6, Nomor 4, 301310.
- Liu, B. (2012). *Sentiment Analysis and Subjectivity. Synthesis Lectures on Human Language Technologies*. USA: editor: Graeme Hirst Morgan & Claypool Publishers.
- Mandala, R. 2014. *Bahan Kuliah Sistem Temu Balik Informasi. Institut teknologi Informasi*.
- Marpaung, Fransiska Humida. 2017. *Analisis Sentimen Opini Masyarakat Indonesia Mengenai Calon Gubernur Dki Jakarta 2017 Menggunakan Naïve Bayes Classifier*. Skripsi S1 Universitas Atma Jaya Yogyakarta.
- Powers, David M W (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation". *Journal of Machine Learning Technologies*. 2(1): 37–63
- Rosdiansyah, Defri. 2014. *Analisis Sentimen Twitter Menggunakan Metode K-Nearest Neighbor dan Pendekatan Lexicon*. Skripsi S1 Universitas Islam Negeri Sultan Syarif Kasim Pekanbaru.

- Saridewi, Ayu Indah. 2015. *Perancangan Dan Implementasi Sistem Peminatan Siswa Sma Dengan Algoritma C4.5 Pada Smk Harapan Denpasar*. Skripsi S1 Universitas Udayana Bali.
- Sasti, Laksminta. 2017. *Analisis Sentimen Komentar Pada Sistem Penilaian Kinerja Instruktur Training ICT (Information and Communication Technology) UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier*. Skripsi S1 UIN Sunan Kalijaga Yogyakarta.
- Shelby.M.I., Warih,M., Adiwijaya. 2013. *Opinion Mining Pada Twitter Menggunakan Klasifikasi Sentimen pada Hastag berbasis Graf*.
- Taspinar, Ahmet dan Lasse Schuirmann. 2017. *Twitterscraper 0.2.7 : Python Package Index*. <https://pypi.python.org/pypi/twitterscraper/0.2.7>.



LAMPIRAN

Daftar Stopword

a	agar	aku	antara	atas
aa	agus	akui	antaranya	atau
ab	ah	akulah	ap	ataukah
abg	ahh	akun	apa	ataupun
abu	ahok	al	apaan	ati
ac	ahy	ala	apabila	ati2
acc	ai	alam	apakah	ato
acr	aint	alat	apalagi	atuh
ad	air	ali	apatah	awa
ada	aj	alie	ape	awak
adab	aja	all	api	awal
adalah	ajaa	am	apps	awalnya
adanya	ajah	ama	aq	awk
adapun	ajak	amat	arah	ayah
adat	ajal	amatlah	are	ayam
ade	aje	ampe	arek	ayat
adek	ak	an	arti	b
adik	akal	anak	artinya	ba
adl	akan	anas	as	bab
adlh	akankah	and	asa	babe
adn	akhir	anda	asal	baca
ae	akhiri	andalah	asalkan	back
af	akhirnya	ane	asap	badja
agak	akn	anies	asik	bae
agaknya	aksi	aniessandi	ask	bagai
agan	akta	antar	at	bagaikan

bagaimana	bas	benar	berkeingi	betul
a	basuki	benarkah	nan	betulkah
bagaimana	baswedan	benarlah	berkenaan	bg
akah	batu	ber	berlainan	bgmn
bagaimana	bau	berada	berlalu	bgt
apun	bawa	berakhir	berlangsung	bgtu
bagi	bawah	berakhirlah	berlebihan	bhw
bagian	bayi	berakhirnya	bermacam	bhwa
bah	bbrp	berapa	bermacam-	bi
bahkan	bdg	berapakah	macam	biasa
bahu	be	berapakah	bermaksud	biasanya
bahwa	bea	berapalah	bermula	bila
bahwasanya	beb	berapapun	bersama	bilakah
baik	beberapa	berarti	bersama-	bill
bail	begini	berawal	sama	bin
baju	beginian	berbagai	bersiap	bir
bak	beginikah	berdatangan	bersiap-	biru
bakal	beginilah	berg	siap	bisa
bakalan	begitu	beri	bertanya	bisakah
bakn	begitukah	berikan	bertanya-	bla
bala	begitulah	berikut	tanya	blg
bale	begitupun	berikutnya	berturut	bljr
bali	beib	a	turut-	bmw
balik	bekerja	berjumlah	turut	bngt
bang	belakang	berkali-	bertutur	bnyk
bank	belakanga	kali	berujar	bobo
banyak	n	berkata	berupa	bola
bapa	beli	berkehendak	besar	boleh
bapak	belum	best	best	bolehkah
baru	belumah			

bolehlah	buta	city	dam	depa
bowo	bvas	ckck	damn	depan
boxy	bwat	ckp	dan	der
bpak	bwt	clbk	dana	desa
bpk	bye	cm	dapat	desi
bray	byk	cmon	dari	dg
brb	c	co	daripada	dgn
bro	c1	coil	data	dh
brp	ca	come	datang	di
bru	cabe	como	daud	dia
bs	cak	con	daun	diakhiri
bsa	call	cont	daur	diakhirinya
bsk	cam	cp	day	a
bth	can	cr	daya	dialah
btp	cape	cu	dc	diantara
btw	cara	cukup	ddl	diantaranya
bu	caranya	cukupkah	de	a
buat	cari	cukuplah	dear	diberi
buka	cb	cul	deh	diberikan
bukan	cc	cuma	dek	diberikan
bukankah	cctv	cup	dekat	nya
bukanlah	cdma	cupu	del	dibuat
bukannya	cek	cwe	demi	dibuatnya
buku	ces	d	demikian	didapat
bulan	cewe	da	demikianl	didatang
bung	ci	dah	ah	an
buru	cie	dahulu	den	digunakan
bus	ciee	dak	deng	diibaratka
but	cih	dalam	dengan	n
			deni	diibaratka
				nya
				diingat

diingatkan	dimulai	disini	dkt	dur
diinginka	h	disinilah	dl	dvd
n	dimulainya	ditambahk	diam	dwi
dijawab	dimungki	an	dlh	dy
dijelaskan	nkan	ditandask	dll	e
dijelaskan	dini	anya	dIm	eden
nya	dipastikan	ditanya i	dlu	eeh
dikarenak	diperbuat	ditanyaka	dm	efek
an	diperbuat	n	dmi	eh
dikatakan	nya	ditegaska	dn	ehh
dikatakan	diperguna	n	dng	ein
nya	kan	ditujukan	do	eks
dikerjakan	diperkirak	an	dodo	el
diketahui	an	ditunjuk	doi	ela
diketahui	diperlihat	kan	dok	ele
nya	kan	ditunjukk	dong	elo
dikira	diperluka	n	donk	elu
dilakukan	n	ditunjukk	dont	elus
dilalui	diperluka	annya	dp	emg
dilihat	nnya	ditunjuk n	dpet	empat
dimaksud	dipersoalk	ya	dpt	en
dimaksud	an	dituturkan	dr	end
kan	dipertanya	kan	dri	eneg
dimaksud	kan	nya	dtg	enggak
kannya	dipunyai	diucapkan	du	enggakny
dimaksud	diri	nya	dua	a
nya	dirinya	diungkapk	duh	entah
diminta	disampaik	an	duit	entahlah
dimintai	an	an	dulu	ep
dimisalka	disebut	dj	dki	epl
n	disebutka	djarot		
dimulai	n	dka		
	disebutka	dki		
	nnya			

era	from	gt	hehe	hy
es	frum	gtu	hell	i
esok	fun	gua	hendak	ia
est	ga	gue	hendaklah	ialah
esta	gaib	guna	hendakny a	ibarat
eta	game	gunakan	her	ibaratkan
etti	gara	guru	here	ibaratnya
eu	gaul	gus	hey	ibnu
euy	gaya	guys	hi	ibu
ex	gd	gw	hiar	ibu2
exo	ge	h	hii	icha
fa	gea	haa	hingga	icw
face	gede	haha	hm	idea
fake	gem	hai	hmm	idl
fare	ges	hak	hmmm	idol
fav	get	hal	hny	ieu
fb	geus	hale	hnya	if
fc	gini	hampir	ho	ih
film	gitu	hand	home	ihsg
fire	gk	hanya	how	ii
fjb	gmn	hanyalah	hoy	iin
for	go	hape	hp	ijo
fort	goal	hari	hr	ika
foto	god	harus	hrs	ikan
fouj	gol	haruslah	ht	iki
fpjp	good	harusnya	htm	iku
fps	grgr	have	htt	ikut
fr	grup	hawa	hub	il
free	gst	hay	huk	im

imho	iyo	jelaslah	juta	karenanya
in	izi	jelasnya	jwb	karo
ind	j	jeng	k	kasus
indo	ja	jg	ka	kat
info	jadi	jiga	kaca	kata
ingat	jadilah	jika	kae	katakan
ingat- ingat	jadinya	jikalau	kah	katakanla h
ingin	jaga	jin	kai	katanya
inginkah	jak	jiwa	kak	kate
inginkan	jam	jk	kaka	kau
ini	jan	jkw	kaki	kaum
inikah	jangan	jl	kala	ke
inilah	jangan janganlah	jl jln	kala kalaulah	keadaan
int	jari	join	kalaupun	kebetulan
io	jasa	jok	kali	kecil
ir	jauh	joko	kalian	kedua
irma	jawa	jose	kalo	keduanya
is	jawab	jroc	kam	keep
isi	jawaban	ju	kami	keinginan
iso	jawabnya	jual	kamilah	keju
ist	bjj	jude	kamu	kek
it	jct	juga	kamulah	kelamaan
its	jd	jumlah	kan	kelihatan
itu	jdi	jumlahnya	kang	kelihatann ya
itukah	je	juni	kapan	kelima
itulah	jede	juru	kapankah	keluar
iu	jelas	just	kapanpun	kembali
iya	jelaskan	justru	karena	kemudian
iyaa				

kemungki nan	kita	kota	laki	lima
kemungki nannya	kitalah	kpd	laku	lini
kena	kite	kpps	lalu	link
kenapa	kitu	krm	lama	list
kepada	kjp	krna	lamanya	live
kepadany a	kjs	kt	lanjut	lo
kesampaia n	kk	ktk	lanjutnya	loe
keseluruh an	kl	ku	lari	loh
keseluruh annya	klan	kudu	las	lol
kete	klik	kui	last	los
keterlalua n	klo	kuis	laut	love
ketika	klub	kuku	lbh	low
kg	km	kull	ldii	lps
khususnya	km09	kurang	ldr	lu
ki	kmk	kurt	le	luar
kian	kmrn	kw	lebih	luas
kick	kmu	kyak	leg	lucy
kim	kn	kyk	leh	lupa
kini	knal	l	lek	m
kinilah	know	la	leo	ma
kira	knp	laa	les	maam
kira-kira	ko	laga	lewat	maap
kiranya	kobe	lagi	lg	macam
kiri	koe	lagian	lgi	madu
kit	koit	lagu	lho	maem
	kok	lah	lht	maen
	koki	lain	lig	mah
	kopi	lainnya	liga	maha
	kot	lak	like	mahu

mail	masing-	memihak	menanya	mengingat
main	masing	meminta	menanyai	mengingat
mak	mata	memintak	menanyak	kan
maka	mati	an	an	mengingi
makanya	mau	memisalk	mendapat	nkan
make	maupun	an	mendapat	mengira
makin	maut	memperb	kan	mengucap
mako	maya	uat	mendatan	kan
malah	mb	memperg	g	mengucap
malahan	mba	unakan	mendatan	kannya
mall	mbak	memperki	gi	mengungk
malu	mbok	rakan	mendatan	apkan
mama	mcm	memperli	gkan	menjadi
mame	me	hatkan	menegask	menjawab
mampu	mei	mempersi	an	menjelask
mampuka	mel	apkan	mengakhi	an
h	melainkan	alkan	ri	menuju
man	melakuka	memperta	mengapa	menunjuk
mana	n	nyakan	mengatak	menunjuk
manakala	melalui	mempunya	an	i
manalagi	melihat	ai	mengatak	menunjuk
mari	melihatny	memulai	annya	kan
mas	a	memung	mengenai	menunjuk
masa	memang	inkan	mengerjak	nya
masalah	memastik	menaiki	an	menurut
masalahn	an	menamba	mengetah	menuturk
ya	memberi	hkan	ui	an
masalahn	memberik	menandas	mengguna	menyamp
masih	an	kan	kan	aikan
masihkah	memberik	menanti	menghend	menyangk
masing	an	menanti-	aki	ut
	membuat	nanti	mengibara	menyatak
	memerluk	nanti	tkan	an
	an	menantika	mengibara	menyebut
		n	tkannya	kan

menyelur uh	mood	naik	nk	ojo
menyiapk an	most	nak	nkri	okay
merasa	mou	nama	nmpk	oke
mereka	move	namun	nnt	ol
merekalah	mrk	nan	nnti	old
merupaka n	ms	nanti	no	oleh
meski	msh	nantinya	noah	olehnya
meskipun	msia	nap	noh	olh
meyakini	msih	nasi	non	olok
meyakink an	mski	nay	not	om
min	mt	nda	now	omdo
minta	mu	ne	nsa	on
mirip	muda	neh	ntar	one
misal	muka	nek	ntr	only
misalkan	mula	neko	nu	oong
misalnya	mulai	neng	nuh	oot
misi	mulailah	new	ny	opo
miss	mulanya	news	nya	or
mjd	muluk	next	nyaris	ora
mlm	mulu	nge	nyatanya	org
mmg	mun	ngp	nyo	org2
mnrt	mung	ni	o	orng
mo	mungkin	niat	obat	otak
mod	mungkin ah	nie	of	otot
moga	mw	nih	ofc	otw
moi	my	nihh	off	out
mom	n	nike	oh	p
	na	nimo	ohh	pa
	nah	njuk	ojek	pada

padahal	perlukah	podr	que	rs
padanya	perlunya	poin	quem	rsud
padu	pernah	pon	r	rt
pagi	persoalan	pool	r80	rupa
pain	pertama	por	ra	rupanya
pair	pertama- tama	pos	rabb	ruu
pak		post	race	s
pali	pertanyaa n	ppbn	raga	sa
paling	pertanyak an	ppp	raib	saal
pan		prak	raih	saat
panjang	php	pro	raja	saatnya
pantas	pi	prof	ran	sabo
papa	pia	psg	rasa	sad
para	pic	psm	rasanya	saik
pas	pihak	pt	rata	saja
pasti	pihaknya	puan	ratu	sajalah
pastilah	pin	pukul	raya	saje
paus	piur	pula	rcti	saling
pd	piye	pun	read	sama
pda	pk	punya	real	sama- sama
pe	pkb	pup	rela	sambil
pee	pkok	pura	reus	sampai
peka	pks	pw	ri	sampai- sampai
penting	plg	q	ribu	sampaika n
pentingny a	plis	qc	rip	sana
per	plus	qs	rmh	sandiaga
percuma	pm	qsan	roma	sang
pere	pnuh	qta	rp	
perlu	po	qu	rp3	

sangat	sebaik- baiknya	seenaknya	sela	semua
sangatlah	sebaiknya	segala	selain	semuanya
sapa	sebaliknya	segalanya	selaku	semula
sapi	sebaliknya	segera	selalu	sen
sara	sebanyak	seharusny a	selama	sendiri
satu	sebegini	sehingga	selama- lamanya	sendirian
saw	sebegitu	seingat	selamanya	sendirinya
say	sebelum	sejak	selanjutny a	seng
saya	sebelumn ya	sejauh	self	seni
sayalah	sebenarny a	sejenak	seluruh	seolah
sbb	seberapa	sejumlah	seluruhny a	seolah- olah
sbg	sebesar	sek	sem	seorang
scr	sebetulny a	sekadar	semacam	sepa
sctv	sebisanya	sekadarny a	semakin	sepanjang
sd	sebuah	sekali	semampu	sepantasn ya
sdg	sebut	sekali-kali	semampu nya	sepantasn yalah
sdh	sebutlah	sekalian	semasa	seperluny a
sdm	sebutnya	sekalipun	semasih	seperti
sdr	secara	sekarang	semata	sepertinya
se	secukupn ya	sekecil	semata- mata	sepihak
sebab	sedang	seketika	semaunya	ser
sebabnya	sedangkan	sekiranya	sementara	seri
sebagai	sedemikia n	sekitar	semisal	sering
sebagaima na	sedikit	sekitarnya	semisaln ya	seringnya
sebagainy a	sedikitnya	sekurang- kurangnya	sempat	serta
sebagian	see	sekurangn ya	semu	seru
sebaik				serupa

sesaat	show	sms	suku	tapi
sesama	si	smua	supaya	tara
sesampai	sia	smw	susu	tau
sesegea	sia2	snsd	swt	tba
sesekali	siap	so	sy	td
seseorang	siapa	soal	sya	tdk
sesi	siapakah	soalnya	sylvi	tdur
sesuatu	siapapun	sob	t	te
sesuatunya	sih	sol	ta	team
sesudah	sik	solo	tadi	tegas
sesudahnya	sin	son	tadinya	tegasnya
setelah	sing	soo	tah	teh
setempat	sini	sop	tahu	tek
setengah	sinilah	sore	tahun	telah
seterusnya	sip	spam	taik	tema
setiap	sipu	sprt	tak	tempat
setiba	sisa	spt	tali	temu
setibanya	sisi	sr	tambah	tena
setidak-tidaknya	siti	ssk	tambahnya	tengah
setidakny	situ	stlh	a	tentang
a	sj	sto	tampak	tentu
setinggi	skg	su	tampaknya	tentulah
seusai	skrg	suatu	a	tentunya
sewaktu	sktr	sud	tandas	tepat
sf	sky	sudah	tandasnya	tepi
sgla	sku	sudahkah	tani	terakhir
sgt	sm	sudahlah	tanpa	terasa
sgtu	sma	suk	tanya	terbanyak
	smp	suke	tanyakan	terdahulu
			tanyanya	

terdapat	th	toh	tvS	und
terdiri	that	tok	twit	ungkap
terhadap	thd	tol	u	ungkapny a
terhadapn ya	thdp this	tong tppi	uang ubah	ungu
teringat	thn	tq	ucap	uni
teringat- ingat	thx	tree	ucapnya	uno
terjadi	ti	tren	ucl	utk
terjadilah	tiap	trs	ucup	untuk
terjadinya	tiba	truk	ud	up
terkira	tiba-tiba	trus	uda	us
terlalu	tidak	tsb	udah	us\$
terlebih	tidakkah	ttg	udh	usah
terlihat	tidaklah	ttp	ugb	usai
termasuk	tiga	tts	ujan	usi
ternyata	tim	tu	ujar	usia
tersampai kan	time	tua	ujarnya	usul
tersebut	tin	tuan	ukm	usut
tersebutla h	tinggi tipe	tubi tugu	ulah ulin	utd utk
tertentu	tips	tuh	uma	uts
tertuju	tkp	tuk	umar	uu
terus	tl	tunjuk	umno	v
terutama	tlah	turi	umt	var
tetap	tlh	turut	umum	vay
tetapi	tni	tutur	umumnya	ve
teu	tnt	tuturnya	umur	via
tgk	to	tv	un	vin
tgl	toel	tv3	una	vit
				vote

voth	wani	wis	yaa	ygh
vou	was	with	yaaa	yng
vs	we	wk	yah	yo
w	weh	wks	yaitu	yoga
wa	wes	wkwk	yak	yok
waduh	wgf	wni	yakin	yoke
wae	wgl	wong	yakni	you
wah	what	x	yang	your
wahai	who	xd	ye	yu
waktu	why	xi	yeah	
waktunya	wib	xx	yes	
walau	will	y	yey	
walaupun	win	ya	yg	



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR RIWAYAT HIDUP

Data Pribadi

Nama : Akhmad Deviyanto
Tempat, Tanggal Lahir : Brebes, 10 Januari 1992
Jenis Kelamin : Laki - Laki
Agama : Islam
Status : Belum Menikah
Kewarganegaraan : Indonesia
Alamat : Jl. Banjarmelati Rt 07/08 Dukuh
Banjarmelati Desa Bangsri Kec. Bulakamba Kab. Brebes
Nomor HP : 0857 2572 6600
Email : akhmad.deviyanto@ gmail.com



Pendidikan Formal :

2012 – 2018 UIN Sunan Kalijaga Yogyakarta Teknik Informatika
2006 – 2009 MAN Rejoso Peterongan Jombang IPA
2003 – 2006 MTs N Model Babakan Lebaksiu -
Tegal
1997 – 2003 SDN Bangsri 3 Bulakamba Brebes -

Pendidikan Non Formal :

2006 – 2009 Pondok Pesantren Darul Ulum Rejoso Peterongan Jombang
2003 – 2006 Pondok Pesantren Ma'hadut Tholabah Babakan Lebaksiu
Tegal