

**Analisis Kepuasan Masyarakat Terhadap Pemindahan Bandara
Yogyakarta Dengan Metode *SentiStrength* Pada Media Sosial
*Twitter***

TUGAS AKHIR

**Diajukan untuk memenuhi salah satu persyaratan mencapai derajat sarjana
komputer**



Disusun Oleh :

Evan Dwi Librian

15 07 08435

**FAKULTAS TEKNOLOGI INDUSTRI
PROGRAM STUDI INFORMATIKA
UNIVERSITAS ATMA JAYA YOGYAKARTA
2020**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS KEPUASAN MASYARAKAT TERHADAP PEMINDAHAN BANDARA YOGYAKARTA
DENGAN METODE SENTISTRENGTH PADA MEDIA SOSIAL TWITTER

yang disusun oleh

EVAN DWI LIBRIAN

150708435

dinyatakan telah memenuhi syarat pada tanggal 07 Desember 2020

		Keterangan
Dosen Pembimbing 1	: Prof. Ir. Suyoto, MSc., PhD	Telah menyetujui
Dosen Pembimbing 2	: Prof. Ir. A. Djoko Budiyo, M.Eng., Ph.D.	Telah menyetujui
Tim Penguji		
Penguji 1	: Prof. Ir. Suyoto, MSc., PhD	Telah menyetujui
Penguji 2	: Yulius Harjoseputro, ST., MT.	Telah menyetujui
Penguji 3	: Dr. Andi Wahyu Rahardjo, BSEE., MSSE	Telah menyetujui

Yogyakarta, 07 Desember 2020
Universitas Atma Jaya Yogyakarta
Fakultas Teknologi Industri
Dekan

ttd

Dr. A. Teguh Siswanto, M.Sc

PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Evan Dwi Librian

NPM : 150708435

Program Studi : Informatika

Fakultas : Teknologi Industri

Judul Penelitian : Analisis Kepuasan Masyarakat Terhadap Pemindahan Bandara Yogyakarta dengan Metode *SentiStrength* pada Media Sosial *Twitter*

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 30 November 2020

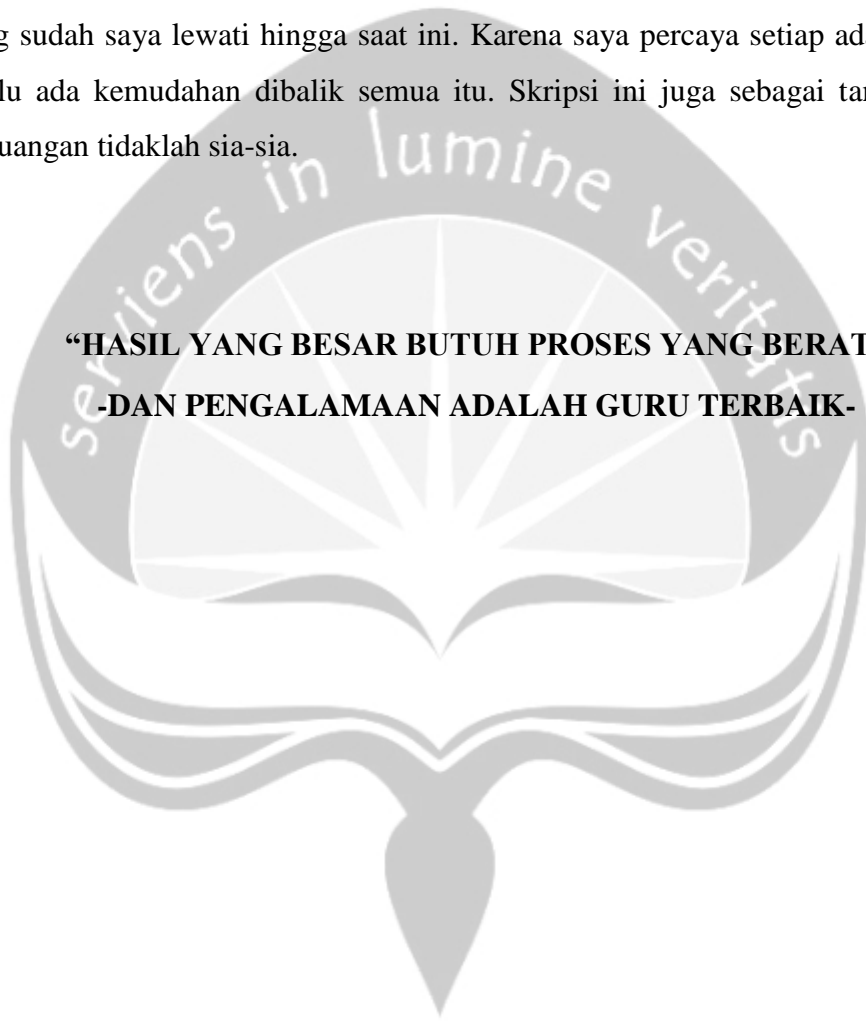
Evan Dwi Librian

150708435

HALAMAN PERSEMBAHAN

Tugas akhir ini saya persembahkan spesial kepada kedua orang tua saya yang sudah memperjuangkan dan memberikan yang terbaik untuk anak-anaknya termasuk saya, dan juga kepada teman-teman terdekat saya yang selalu mendukung saya. Terutama untuk kekasih saya, yang selalu menerima saya apa adanya dalam hal susah maupun senang dan selalu support apapun dalam usaha saya. Tidak henti-hentinya saya bersyukur kepada Tuhan atas segala pencapaian yang sudah saya lewati hingga saat ini. Karena saya percaya setiap ada kesulitan selalu ada kemudahan dibalik semua itu. Skripsi ini juga sebagai tanda bahwa perjuangan tidaklah sia-sia.

**“HASIL YANG BESAR BUTUH PROSES YANG BERAT”
-DAN PENGALAMAAN ADALAH GURU TERBAIK-**



KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa atas segala ramat dan karunianya, sehingga saya dapat menyelesaikan salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Universitas Atma Jaya Yogyakarta. Dalam menyelesaikan skripsi ini tentu banyak kesulitan dan hambatan. Pada kesempatan ini saya ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua yang selalu mendukung saya apapun jalan yang saya ambil.
2. Bapak Dr. A. Teguh Siswanto, M.Sc. selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
3. Bapak Prof. Ir. Suyoto, M.Sc., Ph.D. selaku dosen pembimbing I yang telah membimbing dengan baik dan selalu memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
4. Bapak Prof. Ir. A. Djoko Budiyanto, M.Eng., Ph.D. selaku dosen pembimbing II yang telah membimbing dan selalu memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir.
5. Ibu Stephanie Pamela Adithama, S.T., M.T. Selaku dosen pembimbing akademik penulis selama berkuliah di Universitas Atma Jaya Yogyakarta.
6. Seluruh Dosen, Staf pengajar, dan karyawan Fakultas Teknologi Industri UAJY yang telah membantu penulis selama menjalani kuliah hingga memungkinkan penyelesaian tulisan ini.
7. Seluruh teman-teman terdekat dan rekan-rekan bisnis saya yang tidak bisa saya sebut satu persatu namanya.

Yogyakarta, 30 November 2020

Evan Dwi Librian

150708435

ABSTRAK

Bandar udara atau biasa disebut dengan bandara merupakan sebuah fasilitas untuk pesawat terbang agar dapat mendarat dan lepas landas. Indonesia sendiri memiliki lebih dari 250 bandara, tidak sedikit dari masyarakat lebih memilih menggunakan alat transportasi udara karena waktu perjalanan yang relatif singkat dibandingkan dengan alat transportasi lainnya. Di Indonesia sendiri pembangunan bandara terus dilakukan, seperti yang baru-baru terjadi dikota Yogyakarta yaitu proyek pembangunan bandara New Yogyakarta International Airport (NYIA) di kecamatan Temon, Kabupaten Kulon Progo, Daerah Istimewa Yogyakarta. Penelitian ini dilakukan untuk mengetahui tingkat kepuasan masyarakat terhadap perpindahan bandara Yogyakarta di Kulonprogo, disini penulis menggunakan metode *Sentistrength* pada media sosial *Twitter* untuk membuat klasifikasi sentimen Analisis. Sentimen analisis ini juga bisa digunakan sebagai kesimpulan dalam mengukur suatu kepuasan masyarakat terhadap pemindahan bandara baru di Kulon Progo. Penggunaan metode *Sentistrength* ini dilakukan dengan mengkategorikan setiap sentimen publik kedalam beberapa kategori yaitu positif, negatif dan netral. Hasil penelitian menunjukkan bahwa nilai sentimen masyarakat terhadap perpindahan bandara Yogyakarta berbeda-beda dengan nilai sentimen netral lebih mendominasi sebanyak 63,82% dari 20.330 jumlah data. Kemudian untuk sentimen positif sebanyak 20,56% dan negatif sebanyak 15,62%.

Kata kunci: Analisis, Sentistrength, Python, Twitter, NYIA.

DAFTAR ISI

HALAMAN PERSEMBAHAN	i
KATA PENGANTAR	ii
ABSTRAK	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	vii
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	4
BAB II.....	5
TINJAUAN PUSTAKA	5
BAB III	8
LANDASAN TEORI.....	8
3.1. Data Mining.....	8
3.2. Analisis Sentimen.....	8
3.3. <i>Twitter</i>	9
3.4. <i>SentiStrength</i>	10
3.5. Phyton.....	12
BAB IV	14
METODOLOGI.....	14
4.1. Skema Metodologi Penelitian	14
4.2. Studi Literatur.....	15
4.3. <i>Data Collections</i>	16
4.4. <i>Preprocessing</i>	16
4.4.1. <i>Cleansing</i>	17
4.4.2. <i>Lowercase atau Case Folding</i>	17

4.4.3.	<i>Stopword Removal</i>	17
4.5.	Klasifikasi Analisis Sentimen Menggunakan <i>SentiStrength</i>	17
4.5.1.	Kamus Idiom	18
4.5.2.	Kamus <i>Emoticon</i>	20
4.5.3.	Kamus <i>Sentiment</i>	25
4.5.4.	Kamus <i>BoosterWords</i>	37
4.5.5.	Kata Negasi	37
4.5.6.	Kamus <i>Question Word</i>	38
4.6.	Penentuan Hasil Persentase Sentimen Publik	39
BAB V	40
HASIL DAN PEMBAHASAN	40
5.1.	Pengumpulan Data	40
5.2.	<i>Preprocessing</i>	42
5.2.1.	<i>Remove Duplicate</i>	43
5.2.2.	<i>Stopword</i>	45
5.3.	Klasifikasi Dengan Algoritma <i>SentiStrength</i>	48
5.4.	Analisis Sentimen	54
BAB VI	60
KESIMPULAN DAN SARAN	60
6.1.	Kesimpulan	60
6.2.	Saran	60
DAFTAR PUSTAKA	62
LAMPIRAN	64

DAFTAR GAMBAR

Gambar 4. 1 Skema Metodologi Penelitian	15
Gambar 4. 2 Kamus Idiom	18
Gambar 4. 3 Kamus Emoticon	21
Gambar 4. 4 Kamus Sentiment	25
Gambar 4. 5 Kamus BoosterWords	37
Gambar 4. 6 Kata Negasi	38
Gambar 4. 7 Kamus Question Words	38
Gambar 4. 8 Contoh Penghitungan Sentimen Negative.....	39
Gambar 5. 1 Query Pengumpulan Data Menggunakan Python.....	40
Gambar 5. 2 Proses Pengumpulan Data Menggunakan Python Yang Diaplikasi Kedalam CMD	41
Gambar 5. 3 Hasil Pengumpulan Data Dalam Bentuk Ms.Excel	42
Gambar 5. 4 Hasil Dari Proses Pemilihan Data Dalam Bentuk Ms.Excel.....	42
Gambar 5. 5 Banyaknya Hasil Perolehan Data Sebelum Dilakukan Remove Duplicate	44
Gambar 5. 6 Hasil Data Setelah Dilakukan Remove Duplicate Data	44
Gambar 5. 7 Mengganti Fromat Dari Ms.Excel Kedalam Txt.....	45
Gambar 5. 8 Proses Penargetan Data Yang Akan Dilakukan Stopwords	46
Gambar 5. 9 Proses Saat Memasukan Perintah mvn assembly:single	46
Gambar 5. 10 Proses Tahap Terakhir Stopwords	47
Gambar 5. 11 Data Setelah Dilakukan Proses Stopwords	47
Gambar 5. 12 Hasil Proses Case Folding Dalam Bentuk File Ms.Excel	48
Gambar 5. 13 Query Untuk Melakukan Analisis Sentimen.....	48
Gambar 5. 14 Hasil Klasifikasi Menggunakan SentiStrength Pada PyCharm.....	53

DAFTAR TABEL

Tabel 5. 1 Hasil Analisis Sentimen Menggunakan SentiStrength	54
Tabel 5. 2 Persentase Sentimen Masyarakat	58



BAB I

PENDAHULUAN

1.1 Latar Belakang

Bandar udara atau biasa disebut bandara merupakan sebuah fasilitas untuk pesawat terbang agar dapat mendarat dan lepas landas. Indonesia sendiri memiliki lebih dari 250 bandara, tidak sedikit dari masyarakat lebih memilih menggunakan alat transportasi udara karena waktu perjalanan yang relatif singkat dibandingkan dengan alat transportasi lainnya. Berdasarkan data yang dirilis oleh Airport Council International (ACI) terdapat daftar bandara tersibuk di dunia pada tahun 2019, dan bandara di Indonesia termasuk bandara tersibuk nomor 18 di dunia pada tahun 2019. Selain merilis daftar tersebut, Airport Council International (ACI) juga memberikan laporan bahwa angka penumpang global meningkat 6,6 persen pada tahun 2017 dengan total pergerakan pesawat tumbuh sekitar 2,4%. Airport Council International (ACI) juga memperkirakan pertumbuhan permintaan untuk layanan udara akan tumbuh sekitar 30% hingga tahun 2023.

Di Indonesia sendiri pembangunan bandara terus dilakukan, seperti yang baru-baru terjadi di kota Yogyakarta yaitu proyek pembangunan bandara New Yogyakarta International Airport (NYIA) di kecamatan Temon, Kabupaten Kulon Progo, Daerah Istimewa Yogyakarta. Proyek ini telah berlangsung sejak bulan Oktober tahun 2018. Yogyakarta sendiri sudah memiliki satu bandara yaitu bandara Adisutjipto, alasan dibangunnya bandara baru dikarenakan pada bandara Adisutjipto terjadi kelebihan kapasitas, yang sebenarnya hanya bisa menampung 1,8 juta jiwa penumpang pertahun, hanya saja dalam beberapa tahun kebelakang sudah terjadi kelebihan kapasitas. Seperti pada tahun 2017. Jumlah penumpang mencapai lebih dari 7,8 juta jiwa dan terus naik setiap tahunnya. Berdasarkan standar kementerian perhubungan, setiap satu penumpang seharusnya mempunyai titik ruang sebesar 8 meter persegi, tetapi pada kenyataan yang terjadi di dalam bandara Adisutjipto setiap satu penumpang hanya mempunyai titik ruang sebesar 1,2 meter persegi. Jauh dari batas standar yang diberikan oleh kementerian perhubungan. Selain kelebihan kapasitas ada beberapa alasan lain yang membuat Yogyakarta harus memiliki bandara baru yaitu sebagai berikut:

1. Kapasitas area parkir tempat pesawat dan landasan pacu sepanjang 2.200 meter tidak mampu menampung pesawat berbadan lebar. Apron hanya mampu menampung kurang lebih sebanyak 11 pesawat.
2. Bandara Adisutjipto sebenarnya merupakan Civil Enclave milik TNI Angkatan Udara yang sudah dibangun pada tahun 1938 dan dirancang untuk penerbangan militer kurang lebih selama 40 tahun. Kemudian bandara ini pun menjadi salah satu pangkalan utama TNI AU dan Pusdik Penerbang TNI AU.
3. Dilihat dari segi kendala alam dan lahan, bandara Adisutjipto tidak bisa untuk dikembangkan lagi.
4. Dan untuk memenuhi kebutuhan penumpang dan mendorong pertumbuhan daerah ataupun pertumbuhan nasional serta program pemerintah didalam bidang pariwisata demi mewujudkan 20 juta jiwa wisatawan mancanegara yang datang dan juga tentunya untuk DIY membutuhkan tempat bandara baru yang lebih representatif.

Pada proyek pembangunan infrastruktur bandara ini adapun beberapa masalah terkait pembebasan lahan terhadap warga yang menempati tanah di daerah proyek tersebut, tidak sedikit dari warga yang bertahan atau menolak untuk pindah dengan alasan rumah warisan ataupun rumah yang memiliki kenangan tersendiri. Pada kasus ini pembebasan tanah telah mencapai 98,15% dari total 587,3 hektare (ha). Meski berada di dekat pantai, New Yogyakarta International Airport (NYIA) Kulon Progo lokasi ini diklaim tahan gempa berkekuatan hingga 8,8 Skala Richter (SR), hanya saja akses bandara yang jauh dari pusat kota tersebut juga menjadi salah satu masalah bagi masyarakat Yogyakarta. Dan tidak sedikit pula dari masyarakat yang memberikan reaksi positif maupun negatif tentang pembangunan bandara baru ini, entah dari dalam kota maupun dari luar kota, yang kemudian dituangkan kedalam media social seperti Facebook, Instagram, *Twitter* dan sebagainya. Salah satu media sosial populer yang sering digunakan untuk menyampaikan pendapat, baik berupa diskusi maupun cerita yaitu *Twitter*. *Twitter* sendiri merupakan media sosial yang berkembang pesat saat ini karena pengguna dapat berinteraksi dengan pengguna lainnya dari komputer ataupun perangkat mobile, dimana pun dan kapanpun selagi tersedia jaringan

internet. Pengguna *Twitter* sendiri bisa terdiri dari berbagai macam kalangan yang para penggunanya dapat berinteraksi dengan teman, keluarga hingga rekan kerja. Pada penelitian kali ini penulis menggunakan metode *Sentistrength* untuk membuat klasifikasi sentimen Analisis. Penelitian di lakukan bertujuan untuk mengukur tingkat kepuasan masyarakat Yogyakarta terhadap proses pemindahan New Yogyakarta International Airport.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dijelaskan penulis di atas, sehingga muncul, rumusan masalah yaitu sebagai berikut:

- a. Bagaimana Sentimen publik terhadap rencana pemerintahan terkait pemindahan bandara New Yogyakarta International Airport?
- b. Bagaimana mengklasifikasikan sentimen publik berdasarkan data *tweets* pengguna *Twitter* menggunakan metode *Sentistrength*?
- c. Bagaimana pemerintah menanggapi respon masyarakat terkait masalah pembebasan lahan?
- d. Bagaimana mengklasifikasikan kepuasan masyarakat berdasarkan *tweet* pengguna *Twitter*?

1.3 Batasan Masalah

Berdasarkan rumusan masalah yang sudah disampaikan penulis diatas pada penelitian ini adalah sebagai berikut:

1. Analisis sentimen kepuasan yang dibuat diambil dari opini masyarakat terhadap pemindahan bandara Yogyakarta di Kulonprogo.
2. Analisis dilakukan dengan menggunakan metode *Sentistrength*.
3. Data sentimen kepuasan diambil dari media sosial *Twitter* dan hanya berupa *tweet* dalam Bahasa Indonesia.
4. Hasil keluaran analisis berupa sentimen positif, negatif dan netral.

1.4 Tujuan Penelitian

Berdasarkan batasan masalah yang sudah disampaikan penulis diatas pada penelitian ini adalah dengan menggunakan analisis sentimen, penulis ingin mengetahui tingkat kepuasan masyarakat Yogyakarta terhadap pemindahan bandara Yogyakarta di Kulonprogo.



BAB II TINJAUAN PUSTAKA

Twitter adalah salah satu media sosial yang banyak digunakan orang di dunia terutama di Indonesia. Media ini juga banyak digunakan oleh semua kalangan sebagai bentuk untuk mengungkapkan perasaan maupun pendapat seseorang. Seperti contohnya saat proses pemindahan bandara di Kulon Progo yang menjadi trending topik di *Twitter*. Tidak sedikit dari banyak pengguna *Twitter* yang ikut meramaikannya dalam arti mengkritik maupun mendukung pelaksanaan perpindahan bandara tersebut. Penulis melakukan analisis sentimen publik terkait pemindahan bandara Yogyakarta di kulonprogo.

Penelitian yang lain juga di lakukan oleh Habibi (2016) yang berjudul “Pemetaan Gaya Belajar Mahasiswa Dana Kecenderungan Emosi Pada *Twitter*”. Tujuan penelitian tersebut berkaitan dengan sentimen analisis pada *Twitter*, jika seseorang menyampaikan opini ataupun gagasan mereka di dalam media sosial *Twitter* dapat di pastikan *tweet* orang tersebut mempunyai sifat negatif, positif dan netral. Penulis menggunakan metode Backpropagation. Metode ini menggunakan Bahasa pemograman Java dan Mysql sebagai data basenya dan memakai Weka sebagai aplikasinya [1].

Terkait penelitian analisis sentimen pada media sosial *Twitter* sudah banyak di lakukan sebelumnya dengan menggunakan metode yang bermacam-macam. Seperti halnya pada penelitian yang dilakukan oleh Sianipar R dan Setiawan pada tahun (2015). Yang berjudul “Pendeteksian Kekuatan Sentimen Pada Teks *Tweet* Berbahasa Indonesia Menggunakan *Sentistrength*”. Tujuan peneliatan tersebut yaitu untuk menganalisis sentimen pada teks *tweet* yang berbahasa Indonesia, kemudian juga untuk melihat efek dari metode *Sentistregth* yang digunakan. Opini yang terdapat pada media sosial *Twitter* tersebut nantinya akan dibagi dalam 3 kategori yaitu, sentimen positif, negatif dan netral. Kemudian data set diambil dari enam akun fanbase yang berkaitan dengan kehidupan mahasiswanya, nantinya masing-masiing akun diambil 100 *tweet* secara acak sehingga jumlah data yang nanti digunakan adalah 600 *tweet*. Evaluasi menggunakan metode

Sentistrength yaitu adalah akurasi, uji hipotesis dan uji korelasi dengan metode t-test. Tetapi hasil nilai akurasi pada penelitian ini dianggap masih kurang memuaskan dikarenakan kurangnya terms pada leksikon yang sehingga tidak memenuhi syarat kebutuhan berbahasa Indonesia untuk mengklasifikasi *tweet* [2].

Penelitian yang lain juga dilakukan oleh Moein H (2016) yang berjudul “Penggunaan Media Sosial Untuk Mengukur Kepuasan Pelanggan Dalam Layanan Transit Publik”. Tujuan penelitian dilakukan yaitu untuk mensurvei kepuasan pelanggan (Csat), dengan cara mengumpulkan pendapat pelanggan dalam sistem angkutan umum. Studi ini menganalisis diskusi online di *Twitter* tentang pelayanan dalam bidang transportasi umum, dan untuk mengembangkan alat yang melengkapi survei CSat transit. Analisis ini mempunyai 3 kategori yaitu analisis, semantic, jejaring sosial dan sentimen. Hasilnya kemudian dibandingkan dengan survei CSat yang diterbitkan oleh agen transit Kanada, apakah media sosial dapat digunakan untuk menentukan perkiraan perubahan dalam kepuasan pelanggan [3].

Penelitian yang lain juga dilakukan oleh Rynaldi Tallamma (2019) yang berjudul “Analisis Sentimen Kepuasan Mahasiswa Terhadap UAJY Melalui Media *Twitter*”. Tujuan penelitian ini yaitu untuk menganalisis sentimen atau opini mahasiswa terhadap *service quality* pada Universitas Atma Jaya Yogyakarta. Penulis menggunakan metode algoritma *SentiStrength* untuk melakukan klasifikasi sentimen dengan menggunakan data yang berasal dari *Twitter* sebanyak 17.000 *tweet*. Kemudian dibagi dalam beberapa kategori yang terdiri dari daya tanggap, empati, garansi, kehandalan dan fisik (tangibles). Setiap dimensi nantinya akan dilakukan penilaian sentimennya menggunakan metode *Sentistrength*, kemudian diklasifikasikan kedalam tiga kategori yaitu positif, negatif dan netral. Hasil dari penelitian tersebut akan menunjukkan berapa banyaknya sentimen negatif pada dimensi fisik, kehandalan, daya tanggap, empati dan garansi. sehingga diperlukan peningkatan *service quality* pada setiap dimensi yang dinilai [4].

Tabel 2.1. merupakan tabel perbandingan Analisis Sentimen yang pernah dilakukan oleh beberapa peneliti beserta dengan metode-metode yang digunakan peneliti.

Tabel 2. 1 Perbandingan Analisis Sentimen

Pembanding	Habibi (2016)	Sianipar R dan Setiawan E B. (2015)	Moein Hosseini (2016)	Rynaldi Talama (2019)
Topik	Pemetaan Gaya Belajar Mahasiswa dan Kecendrungan Emosi pada <i>Twitter</i>	Pendeteksian Kekuatan Sentimen pada Teks <i>Tweet</i> Berbahasa Indonesia Menggunakan <i>SentiStrength</i>	Penggunaan Media Sosial untuk Mengukur Kepuasan Pelanggan dalam Layanan Transit Publik	Analisis Sentimen Kepuasan Mahasiswa terhadap UAJY melalui Media <i>Twitter</i>
Metode	<i>backpropagation</i>	<i>SentiStrength</i>	Sematic Analisis, <i>Sosial Network</i> dan Sentiment Analisis	<i>SentiStrength</i>
Database	<i>Mysql</i>	-	-	-
Platfrom	Dekstop	Dekstop	Dekstop	Dekstop

BAB III

LANDASAN TEORI

Pada bab ini penulis akan membahas tentang bagian serta metode-metode yang digunakan dan menjadi landasan dalam melakukan Analisis Kepuasan Masyarakat Terhadap Pemandangan Bandara Yogyakarta Dengan Metode *SentiStrength* Pada Media Sosial *Twitter*.

3.1. Data Mining

Text mining atau biasa disebut data mining adalah suatu proses untuk menemukan informasi dalam sebuah data yang besar (umumnya tidak terstruktur) dan kemudian secara otomatis akan mengidentifikasi pola tersembunyi dari hubungan setiap data tersebut [5]. *Text mining* dapat membantu pengguna untuk menganalisa dan mengelompokkan teks yang bersifat tidak terstruktur dalam jumlah besar termasuk data dari media *Twitter*. Hasil pengolahan data dalam data mining dapat dijadikan landasan dalam proses pengambilan keputusan. Penerapan data mining untuk mengambil keputusan juga dilakukan oleh beberapa peneliti. Seorang peneliti pada tahun 2017 menggunakan data mining untuk mengelompokkan penderita *Thalassaemia* untuk menentukan tindakan medis terhadap setiap kelompok penderita [6].

Selain itu, pada tahun 2014 seorang peneliti juga menggunakan data mining untuk mengklasifikasikan nasabah kredit sebuah bank untuk menentukan keputusan terkait pemberian kredit kepada nasabah. *Text mining* mengadopsi dan mengembangkan banyak teknik dari bidang lain, seperti *Data Mining*, *Information Retrieval*, Statistik dan Matematik, *Machine Learning*, *Linguistic*, *Natural Language Processing* dan *Visualization* untuk menganalisa sebuah data. Kegiatan riset untuk *text mining* antara lain ekstraksi dan penyimpanan teks, *preprocessing* akan konten teks, pengumpulan data statistik serta *indexing* dan Analisis Sentimen [7].

3.2. Analisis Sentimen

Sentimen atau opini adalah sesuatu hal yang berkaitan erat dengan kehidupan manusia karena memiliki pengaruh terhadap kebiasaan. Opini menjadi dasar dalam berbagai hal baik kehidupan pribadi, kebijakan publik, maupun

penilaian kualitas layanan sebuah organisasi. Analisis Sentimen atau biasa juga disebut dengan *opinion mining* adalah sebuah bidang studi yang mempelajari atau menganalisis tentang suatu opini, sentimen, evaluasi, *attitude*, serta emosi publik terhadap layanan, topik umum, produk, organisasi, isu-isu publik, dan lain sebagainya [8]. Sentimen analisis juga digunakan untuk mengklasifikasikan sentimen ke dalam kategori positif maupun negatif agar menjadi acuan untuk perbaikan suatu sistem, produk, layanan, dan sebagainya [9].

Selain itu di dalam dunia bisnis juga digunakan dalam menilai suatu layanan atau produk. Kemudian juga di dalam dunia politik untuk mengetahui pendapat publik terhadap seorang kandidat politik atau kebijakan yang dikeluarkan oleh pemerintah. Sentimen analisis terbagi menjadi 3 level, yang pertama yaitu level dokumen atau biasa disebut klasifikasi sentimen level dokumen dimana analisis ini dilakukan untuk seluruh dokumen dan kemudian diambil kesimpulan akhir untuk menentukan dokumen tersebut bernilai positif atau negatif. Analisis Sentimen level dokumen umumnya hanya digunakan untuk menilai suatu entitas, tidak untuk lebih dari satu entitas. Pada Level kedua yaitu level *sentence* dimana setiap *sentence* atau kalimat dinilai dengan kategori negatif, positif, dan netral (netral dalam arti dianggap tidak ada sentimen), dan level yang ketiga yaitu level entitas dan aspek dimana setiap aspek dalam sebuah kalimat dapat diklasifikasikan sentimennya [10].

3.3. Twitter

Media sosial menjadi salah satu media yang mudah digunakan untuk mengetahui sentimen publik karena media sosial menjadi tempat bagi masyarakat untuk menuangkan opini. Ketersediaan informasi yang banyak di media sosial dapat mempermudah penggalan dalam informasi terkait opini publik terhadap sebuah isu. *Twitter* merupakan salah satu media sosial yang memungkinkan penggunaannya untuk berbagi informasi dengan membaca, mengirim pesan blog (berbasis teks) dan mengirim video. Pesan yang di kirim dan dibaca pada media sosial *Twitter* biasa disebut *tweet*. Pada media sosial *Twitter* pencarian informasi dapat dipermudah menggunakan simbol # (*hashtag*). Simbol # (*hashtag*) biasanya digunakan untuk mengelompokkan suatu topik ataupun informasi yang dikirim pada media social *Twitter*. Simbol # dapat digunakan oleh pengguna *Twitter* untuk

mencari informasi maupun *tweet* terkait informasi yang spesifik. Penggunaan simbol hashtag diletakkan pada awal kata atau kalimat yang relevan untuk mengkategorikan *tweet* tersebut pada suatu topik tertentu agar mempermudah saat proses pencarian [10].

Media sosial *Twitter* juga banyak digunakan untuk suatu penelitian yang sejalan dengan luasnya penggunaan *Twitter* sebagai media diskusi oleh semua kalangan. Penggunaan media *Twitter* diseluruh dunia semakin mengalami peningkatan setiap harinya termasuk di Indonesia. Menurut catatan Kementerian Komunikasi dan Informatika RI, jumlah pengguna media social *Twitter* di Indonesia mencapai 19,5 juta orang dari 500 juta pengguna *Twitter* global dan Indonesia menempati posisi ke 5 di dunia. Hal ini memungkinkan penggunaan data dari media sosial *Twitter* untuk keperluan suatu penelitian [11].

3.4. *SentiStrength*

SentiStrength merupakan salah satu program classifier untuk Analisis Sentimen yang menggunakan pendekatan berbasis kamus atau biasa disebut dengan leksikon. Leksikon menggunakan pendekatan yang berfokus pada kekuatan sentimen dimana kekuatan sentimen akan menentukan emosi manusia (emosi positif dan negatif). Pada metode *SentiStrength*, sentimen akan diberikan skor/value dengan range value 1-5 dimana range 1 menunjukkan sentimen yang netral dan range 5 menunjukkan sentimen yang sangat positif atau sangat negatif. Penentuan kategori dari suatu kalimat dilakukan dengan melihat nilai positif tertinggi dan nilai negatif tertinggi dari setiap kata penyusunnya. Dari setiap kata nantinya akan diberi skor sesuai dengan skor dari kata tersebut dalam leksikon atau kamus sentimen [12]. Penentuan hasil akhir ditentukan oleh aturan sebagai berikut:

- Jika positif > negatif maka sentimen dikategorikan sebagai positif
- Jika negatif > positif maka sentimen dikategorikan sebagai negatif
- Jika positif = negatif, maka sentimen dikategorikan netral dimana sentimen netral dianggap tidak memiliki sentimen.

Penggunaan *SentiStrength* sangat luas di berbagai belahan negara dan digunakan untuk menganalisis berbagai jenis bahasa, namun leksikon yang tersedia pada *SentiStrength* menggunakan bahasa Inggris. *SentiStrength* juga memiliki suatu kelebihan dimana bahasa dalam leksikonnya dapat diterjemahkan ke berbagai jenis bahasa termasuk dalam bahasa Indonesia. Hal ini yang memungkinkan *SentiStrength* dapat digunakan untuk menganalisis Sentimen dalam berbagai jenis bahasa. *SentiStrength* juga merupakan metode Analisis Sentimen yang bersifat fleksibel dimana leksikon dapat diatur ataupun diubah sesuai dengan kebutuhan bahasa yang akan dianalisis oleh pengguna [12]. Kamus yang tersedia di dalam program untuk Analisis Sentimen dalam bahasa Indonesia dikembangkan oleh Azhari N dan Wahid pada tahun 2016 dan terdiri dari kamus sentimen dalam bahasa Indonesia, kamus negasi, kamus *emoticon*, kata tanya, kamus kata penguat, dan kamus idiom [13]. Penjelasan terkait setiap kamus tersebut dijabarkan sebagai berikut:

a. Kamus sentimen

Kamus sentimen merupakan kamus yang terdiri dari kata dasar yang sudah diberi skor atau nilai sentimen mulai dari -1 (tidak memiliki sentimen negatif) sampai -5 (sentimen negatif yang sangat kuat) dan 1 (tidak memiliki sentimen negatif) sampai 5 (memiliki sentimen positif yang sangat kuat). Kamus ini juga menjadi dasar penilaian sentimen karena digunakan untuk menilai atau memberi bobot dari setiap kata yang terkandung dalam suatu kalimat.

b. Kamus *Emoticon*

Kamus *emoticon* adalah kamus yang berisi *emoticon* atau gabungan dari simbol-simbol yang sudah diberi skor atau nilai sentimen berdasarkan makna dari *emoticon* itu sendiri. *emoticon* juga diberi nilai karena gabungan simbol-simbol tersebut biasanya digunakan untuk menggambarkan suatu ekspresi manusia ketika mengungkapkan sebuah kalimat dan dapat menegaskan emosi yang terkandung dalam data yang berbentuk teks.

c. Kamus Idiom

Kamus idiom adalah kamus yang berisi sebuah rangkaian kata yang memiliki makna yang menyatu dan tidak dapat ditafsirkan secara terpisah. Dalam sentimen analisis kamus idiom diberi nilai atau skor tersendiri sehingga tidak dapat ditafsirkan per kata.

d. Kamus Kata Tanya

Kamus kata tanya merupakan kamus yang berisi daftar kata tanya seperti apakah, akankah, benarkah, bagaimana, dan lain sebagainya. Kata tanya dalam sentimen analisis dapat mengubah suatu makna dari kalimat tersebut, jika sebuah kalimat diawali dengan kata tanya maka sentimennya dikategorikan netral atau tidak mengandung sentimen.

e. Kamus Kata Negasi

Kamus kata negasi merupakan kamus yang berisi sebuah kata yang dapat mengubah makna dari suatu kalimat. Misalnya kata “cantik” memiliki makna yang positif, namun akan berbeda jika ditambah dengan kata negasi “tidak” menjadi “tidak cantik” maka maknanya akan menjadi negatif.

f. Kamus *Boosterwords*

Kamus *Boosterwords* adalah sebuah kata yang dapat meningkatkan atau menambah makna dari kata. Penambahan *boosterword* pada sebuah kata dapat meningkatkan makna kata tersebut sehingga diberi bobot atau skor lebih tinggi dibandingkan jika tanpa *boosterwords*. Misalnya kata “cantik” memiliki skor bobot 4, namun jika ditambahkan *boosterword* seperti “sangat” menjadi “sangat cantik” maka skor bobotnya akan menjadi 5.

3.5. Phyton

Python merupakan salah satu bahasa pemrograman yang dapat dijalankan pada berbagai sistem operasi baik di Windows, MacOS maupun di Linux dengan cara melakukan konfigurasi terlebih dahulu. Python juga merupakan sebuah bahasa pemrograman yang dikategorikan sebagai bahasa pemrograman tingkat tinggi dari bahasa pemrograman lainnya.

Dan dikarenakan code yang ditulis kemudian akan di *compile* menjadi byte code, dan lalu dieksekusi, sehingga membuat Python tepat digunakan untuk bahasa pemograman scripting, aplikasi web dan juga untuk bahasa pemograman lainnya [14].



BAB VI

KESIMPULAN DAN SARAN

6.1. Kesimpulan

Berdasarkan penelitian Analisis Kepuasan Masyarakat Terhadap Perpindahan Bandara Yogyakarta dengan metode *Sentistrength* dapat digunakan untuk menganalisis sentimen dengan melakukan klasifikasi, dalam hal ini di bagi menjadi 3 kategori yaitu, negatif, positif dan netral. Pada penelitian ini hasil analisis kepuasan masyarakat terhadap perpindahan bandara Yogyakarta umumnya netral. Dengan hasil sentimen netral sebanyak 63,82% dari 20.330 jumlah data. Kemudian untuk sentimen positif sebanyak 20,56% dan yang terakhir negatif dengan hasil sentimen 15,62%.

Data yang diperoleh menunjukkan bahwa masyarakat lebih dominan untuk tidak ikut campur atau mempermasalahkan terkait adanya perpindahan bandara baru. Namun ada juga warga yang turut mendukung pemindahan bandara baru Yogyakarta dikarenakan masyarakat berpikir terkait kemajuan ekonomi disekitar bandara baru. Ada juga warga yang menolak pemindahan bandara baru dikarenakan beberapa faktor diantaranya adalah adanya PHK beberapa karyawan, warga yang tidak ingin pindah tempat, maupun ganti rugi yang diberikan pihak bandara tidak sesuai dengan harapan warga. Dari data yang diperoleh juga diketahui bahwa pemerintah melakukan upaya untuk mengatasi permasalahan terkait beberapa warga yang menolak pindah yaitu dengan cara membeli tanah warga dengan harga yang lebih tinggi.

6.2. Saran

Berikut merupakan saran dari penulis untuk pembaca sebagai berikut :

- a. Dari hasil penelitian ini, perlu dilakukan perbaikan dan relevansi isi kamus pada program *SentiStrength* untuk pengembangan penelitian selanjutnya.
- b. Perlu dilakukan evaluasi dan pelayanan terhadap beberapa fasilitas Transportasi umum menuju bandara baru Yogyakarta di Kulonprogo, karena dari data ditemukan bahwa masih banyak penumpang yang masih menggunakan kendaraan pribadi dibandingkan dengan Transportasi umum. Sehingga tidak menimbulkan kemacetan pada jalur menuju

bandara.

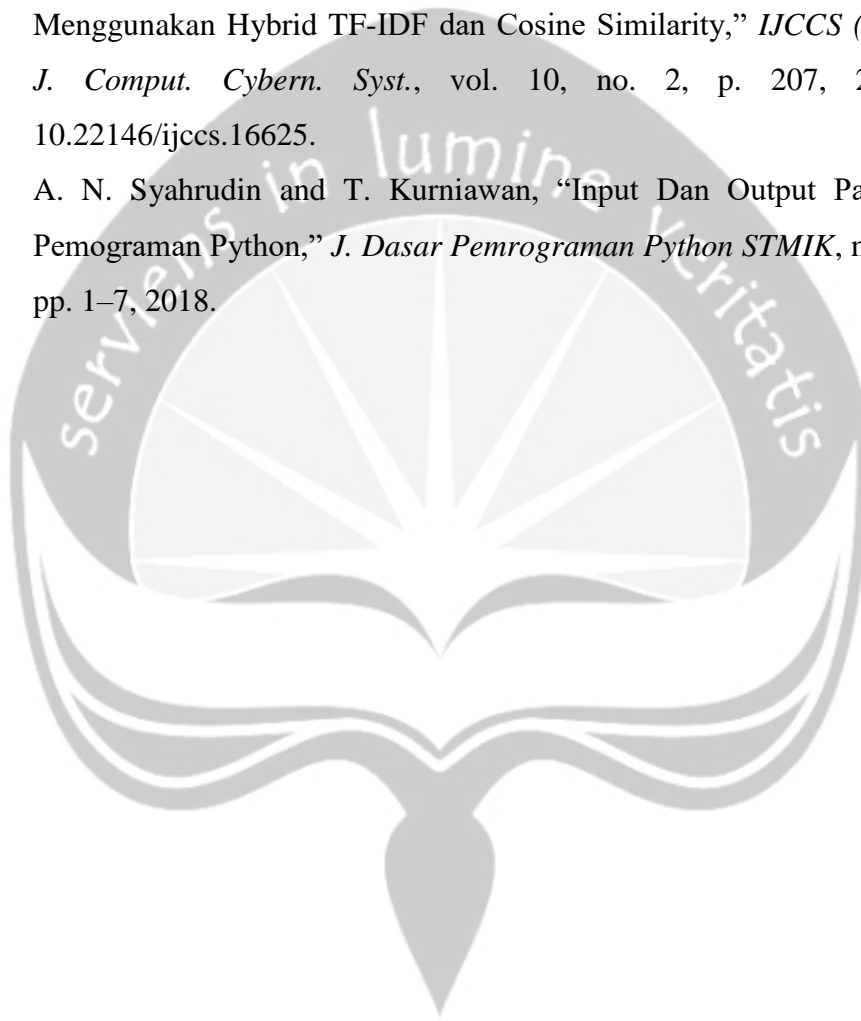
- c. Perlu dilakukan pengembangan lebih lanjut pada metode *SentiStrength* sehingga data yang dihasilkan lebih akurat.



DAFTAR PUSTAKA

- [1] R. Habibi, “Pemetaan Gaya Belajar Mahasiswa Dan Kecenderungan Emosi Pada Twitter,” pp. 1–39, 2016.
- [2] R. Sianipar and E. B. Setiawan, “Pendeteksian Kekuatan Sentimen Pada Teks Tweet Berbahasa Indonesia Menggunakan SentiStrength,” *e-Proceeding Eng. Vol.2, No.3 Desember 2015*, vol. 2, no. 3, pp. 7922–7928, 2015.
- [3] M. Hosseini, “Use of Social Media to Measure Customer Satisfaction in Public Transit Services Use of Social Media to Measure Customer Satisfaction with Public Transit Service,” pp. 6–30, 2016.
- [4] R. TALLAMMA, “Analisis Kepuasan Mahasiswa Terhadap UAJY Pada Media Sosial Twitter,” *J. Inf.*, vol. 10, pp. 1–16, 2009.
- [5] R. Feldman and J. Sanger, “The Text Mining Handbook,” *J. Texture Stud.*, vol. 15, no. 4, pp. 417–423, 1984, doi: 10.1111/j.1745-4603.1984.tb00397.x.
- [6] H. Sulastrri and A. I. Gufroni, “Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 299–305, 2017, doi: 10.25077/teknosi.v3i2.2017.299-305.
- [7] I. H. Witten, K. J. Don, M. Dewsnip, and V. Tablan, “Text mining in a digital library,” *Int. J. Digit. Libr.*, vol. 4, no. 1, pp. 56–59, 2004, doi: 10.1007/s00799-003-0066-4.
- [8] B. Liu, “Sentiment analysis: Mining opinions, sentiments, and emotions,” *Sentim. Anal. Min. Opin. Sentim. Emot.*, no. May, pp. 1–367, 2015, doi: 10.1017/CBO9781139084789.
- [9] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1750–1757, 2017.
- [10] F. Anugratami, M. Christin, and B. Primadani, “Pengaruh Penggunaan Media Sosial Twitter Terhadap Motivasi Followers Pada Akun @Merryriana the Effect of Social Media Twitter Usage on Followers’

- Motivation in @Merryriana Account,” vol. 2, no. 2, pp. 2256–2261, 2014.
- [11] R. Anisa and R. Rachmaniar, “Media Informasi dan Komunikasi Pemerintah Provinsi Jawa Barat,” *J. Lontar*, vol. 7, no. 1, pp. 36–40, 2019.
- [12] M. Thelwall, “The Heart and Soul of the Web? Sentiment Strength Detection in the Social Web with SentiStrength,” vol. 5, pp. 119–134, 2017, doi: 10.1007/978-3-319-43639-5_7.
- [13] D. H. Wahid and A. SN, “Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 10, no. 2, p. 207, 2016, doi: 10.22146/ijccs.16625.
- [14] A. N. Syahrudin and T. Kurniawan, “Input Dan Output Pada Bahasa Pemograman Python,” *J. Dasar Pemrograman Python STMIK*, no. January, pp. 1–7, 2018.



LAMPIRAN

Lampiran 1. Query NetBeans

```
package com.uttesh.exude.stemming;

import com.uttesh.exude.ExudeData;
import com.uttesh.exude.exception.InvalidDataException;
import java.io.*;
import java.sql.Connection;
import java.sql.PreparedStatement;
import java.sql.ResultSet;
import java.sql.Statement;
import dao.myDao;
import static dao.myDao.c;
import static dao.myDao.path;
import static dao.myDao.url;
import java.sql.DriverManager;
import java.util.StringTokenizer;
import jdk.nashorn.internal.parser.TokenStream;

/**
 * Stemmer, implementing the Porter Stemming Algorithm
 *
 * The Stemmer class transforms a word into its root form. The input
 * word can be provided a character at time (by calling add()), or at once
 * by calling one of the various stem(something) methods.
 */

public class Stemmer
{
    public static Connection c;
    public static final String driver = "sun.jdbc.odbc.JdbcOdbcDriver";
    public static myDao mydao = new myDao();
    public static String url = "jdbc:ucanaccess://";
    public static final String path = "E:" + File.separator + "Kuliner.mdb";
    private char[] b;
    private int i, /* offset into b */
               i_end, /* offset to end of stemmed word */
               j, k;
    private static final int INC = 50;
    /* unit of size whereby b is increased */

    public Stemmer()
    {
        b = new char[INC];
        i = 0;
        i_end = 0;
    }
}
```

```

public void add(char ch)
{
    if (i == b.length)
    {
        char[] new_b = new char[i+INC];
        for (int c = 0; c < i; c++) new_b[c] = b[c];
        b = new_b;
    }
    b[i++] = ch;
}

```

```

/** Adds wLen characters to the word being stemmed contained in a portion
 * of a char[] array. This is like repeated calls of add(char ch), but
 * faster.
 */

```

```

public void add(char[] w, int wLen)
{
    if (i+wLen >= b.length)
    {
        char[] new_b = new char[i+wLen+INC];
        for (int c = 0; c < i; c++) new_b[c] = b[c];
        b = new_b;
    }
    for (int c = 0; c < wLen; c++) b[i++] = w[c];
}

```

```

/**
 * After a word has been stemmed, it can be retrieved by toString(),
 * or a reference to the internal buffer can be retrieved by getResultBuffer
 * and getResultLength (which is generally more efficient.)
 */

```

```

public String toString() { return new String(b,0,i_end); }

```

```

/**
 * Returns the length of the word resulting from the stemming process.
 */

```

```

public int getResultLength() { return i_end; }

```

```

/**
 * Returns a reference to a character buffer containing the results of
 * the stemming process. You also need to consult getResultLength()
 * to determine the length of the result.
 */

```



```

public char[] getResultBuffer() { return b; }

/* cons(i) is true <=> b[i] is a consonant. */

private final boolean cons(int i)
{
    switch (b[i])
    {
        case 'a': case 'e': case 'i': case 'o': case 'u': return false;
        case 'y': return (i==0) ? true : !cons(i-1);
        default: return true;
    }
}

/* m() measures the number of consonant sequences between 0 and j. if c is
a consonant sequence and v a vowel sequence, and <..> indicates arbitrary
presence,

    <c><v>      gives 0
    <c>vc<v>    gives 1
    <c>vcvc<v>  gives 2
    <c>vcvcvc<v> gives 3
    ....
*/

private final int m()
{
    int n = 0;
    int i = 0;
    while(true)
    {
        if (i > j) return n;
        if (! cons(i)) break; i++;
    }
    i++;
    while(true)
    {
        while(true)
        {
            if (i > j) return n;
            if (cons(i)) break;
            i++;
        }
        i++;
        n++;
        while(true)
        {
            if (i > j) return n;
            if (! cons(i)) break;
            i++;
        }
        i++;
    }
}

```



```

    }
}

/* vowelinstem() is true <=> 0,...j contains a vowel */

private final boolean vowelinstem()
{
    int i; for (i = 0; i <= j; i++) if (! cons(i)) return true;
    return false;
}

/* doublec(j) is true <=> j, (j-1) contain a double consonant. */

private final boolean doublec(int j)
{
    if (j < 1) return false;
    if (b[j] != b[j-1]) return false;
    return cons(j);
}

/* cvc(i) is true <=> i-2,i-1,i has the form consonant - vowel - consonant
and also if the second c is not w,x or y. this is used when trying to
restore an e at the end of a short word. e.g.

    cav(e), lov(e), hop(e), crim(e), but
    snow, box, tray.

*/

private final boolean cvc(int i)
{
    if (i < 2 || !cons(i) || cons(i-1) || !cons(i-2)) return false;
    {
        int ch = b[i];
        if (ch == 'w' || ch == 'x' || ch == 'y') return false;
    }
    return true;
}

private final boolean ends(String s)
{
    int l = s.length();
    int o = k-1+l;
    if (o < 0) return false;
    for (int i = 0; i < l; i++) if (b[o+i] != s.charAt(i)) return false;
    j = k-1;
    return true;
}
}

```



```

private final void setto(String s)
{ int l = s.length();
  int o = j+1;
  for (int i = 0; i < l; i++) b[o+i] = s.charAt(i);
  k = j+1;
}

/* r(s) is used further down. */

private final void r(String s) { if (m() > 0) setto(s); }

/* step1() gets rid of plurals and -ed or -ing. e.g.

   caresses -> caress
   ponies   -> poni
   ties     -> ti
   caress   -> caress
   cats     -> cat

   feed     -> feed
   agreed   -> agree
   disabled -> disable

   matting  -> mat
   mating   -> mate
   meeting  -> meet
   milling  -> mill
   messing  -> mess

   meetings -> meet

*/

private final void step1()
{ if (b[k] == 's')
  { if (ends("sses")) k -= 2; else
    if (ends("ies")) setto("i"); else
    if (b[k-1] != 's') k--;
  }
  if (ends("eed")) { if (m() > 0) k--; } else
  if ((ends("ed") || ends("ing")) && vowelinstem())
  { k = j;
    if (ends("at")) setto("ate"); else
    if (ends("bl")) setto("ble"); else
    if (ends("iz")) setto("ize"); else
    if (doublec(k))

```



```

        { k--;
          { int ch = b[k];
            if (ch == 'l' || ch == 's' || ch == 'z') k++;
          }
        }
        else if (m() == 1 && cvc(k)) setto("e");
    }
}

/* step2() turns terminal y to i when there is another vowel in the stem. */
private final void step2() { if (ends("y") && vowelinstem()) b[k] = 'i'; }

/* step3() maps double suffixes to single ones. so -ization (= -ize plus
-ation) maps to -ize etc. note that the string before the suffix must give
m() > 0. */
private final void step3() { if (k == 0) return; /* For Bug 1 */ switch (b[k-1])
{
    case 'a': if (ends("ational")) { r("ate"); break; }
              if (ends("tional")) { r("tion"); break; }
              break;
    case 'c': if (ends("enci")) { r("ence"); break; }
              if (ends("anci")) { r("ance"); break; }
              break;
    case 'e': if (ends("izer")) { r("ize"); break; }
              break;
    case 'l': if (ends("bli")) { r("ble"); break; }
              if (ends("alli")) { r("al"); break; }
              if (ends("entli")) { r("ent"); break; }
              if (ends("eli")) { r("e"); break; }
              if (ends("ousli")) { r("ous"); break; }
              break;
    case 'o': if (ends("ization")) { r("ize"); break; }
              if (ends("ation")) { r("ate"); break; }
              if (ends("ator")) { r("ate"); break; }
              break;
    case 's': if (ends("alism")) { r("al"); break; }
              if (ends("iveness")) { r("ive"); break; }
              if (ends("fulness")) { r("ful"); break; }
              if (ends("ousness")) { r("ous"); break; }
              break;
    case 't': if (ends("aliti")) { r("al"); break; }
              if (ends("iviti")) { r("ive"); break; }
              if (ends("biliti")) { r("ble"); break; }
              break;
}
}

```



```

        case 'g': if (ends("logi")) { r("log"); break; }
    } }

    /* step4() deals with -ic-, -full, -ness etc. similar strategy to step3. */

private final void step4() { switch (b[k])
{
    case 'e': if (ends("icate")) { r("ic"); break; }
              if (ends("ative")) { r(""); break; }
              if (ends("alize")) { r("al"); break; }
              break;
    case 'i': if (ends("iciti")) { r("ic"); break; }
              break;
    case 'l': if (ends("ical")) { r("ic"); break; }
              if (ends("ful")) { r(""); break; }
              break;
    case 's': if (ends("ness")) { r(""); break; }
              break;
} }

    /* step5() takes off -ant, -ence etc., in context <c>vcvc<v>. */

private final void step5()
{ if (k == 0) return; /* for Bug 1 */ switch (b[k-1])
{ case 'a': if (ends("al")) break; return;
  case 'c': if (ends("ance")) break;
              if (ends("ence")) break; return;
  case 'e': if (ends("er")) break; return;
  case 'i': if (ends("ic")) break; return;
  case 'l': if (ends("able")) break;
              if (ends("ible")) break; return;
  case 'n': if (ends("ant")) break;
              if (ends("ement")) break;
              if (ends("ment")) break;
              /* element etc. not stripped before the m */
              if (ends("ent")) break; return;
  case 'o': if (ends("ion") && j >= 0 && (b[j] == 's' || b[j] == 't')) break;
              /* j >= 0 fixes Bug 2 */
              if (ends("ou")) break; return;
              /* takes care of -ous */
  case 's': if (ends("ism")) break; return;
  case 't': if (ends("ate")) break;
              if (ends("iti")) break; return;
  case 'u': if (ends("ous")) break; return;
  case 'v': if (ends("ive")) break; return;
  case 'z': if (ends("ize")) break; return;
} }

```

```

        default: return;
    }
    if (m() > 1) k = j;
}

/* step6() removes a final -e if m() > 1. */

private final void step6()
{
    j = k;
    if (b[k] == 'e')
    {
        int a = m();
        if (a > 1 || a == 1 && !cvc(k-1)) k--;
    }
    if (b[k] == 'l' && doublec(k) && m() > 1) k--;
}

/** Stem the word placed into the Stemmer buffer through calls to add().
 * Returns true if the stemming process resulted in a word different
 * from the input. You can retrieve the result with
 * getResultLength()/getResultBuffer() or toString().
 */
public void stem()
{
    k = i - 1;
    if (k > 1) { step1(); step2(); step3(); step4(); step5(); step6(); }
    i_end = k+1; i = 0;
}

public String stem(String s) {
    for(char c : s.toCharArray()) {
        add(c);
    }
    stem();
    return toString();
}

public static void main(String[] args) throws InvalidDataException {
    Stemmer stemmer = new Stemmer();

    String str = "watched out of tv";
    StringTokenizer st = new StringTokenizer(str);
    System.out.println("---- Split by space -----");
    StringBuilder sb = new StringBuilder();
    String temp="";
    while (st.hasMoreElements()) {

```



```
//System.out.println(st.nextElement());
//String id= st.nextToken();

//System.out.println(lemmatizer.lemmatize(st.nextToken()));
System.out.println("stem : "+stemmer.stem(st.nextToken()));

//temp = temp + lemmatizer.lemmatize()+"\n";

//sb.append( lemmatizer.lemmatize(st.nextToken())+" ");
}
}
}
```

