

ANALISA SENTIMEN KEBIJAKAN PEMERINTAH PADA KONTEN TWITTER BERBAHASA INDONESIA MENGGUNAKAN SVM DAN K-MEDOID CLUSTERING

Latifa Nurrachma Pradany¹, Chastine Faticah²

^{1,2}Institut Teknologi Sepuluh Nopember Surabaya

Email: latifadany@gmail.com¹, chastine@cs.its.ac.id²

Abstrak. Pada saat ini penggunaan Twitter semakin luas. Semua pengguna twitter yang dapat dengan bebas untuk berpendapat dan membagikan sudut pandang mereka mengenai isu tren dunia membuat konten twitter menjadi beragam dan menarik untuk dianalisa, termasuk dengan tren kebijakan politik yang ramai diperbincangkan di Indonesia setahun terakhir. Hal tersebut menjadi alasan kuat bahwa analisis opini publik dan sentimen tentang kebijakan pemerintah dari isi twitter Indonesia sangat penting untuk dilakukan. Namun, identifikasi sentimen di dalam pesan konten Twitter berbahasa Indonesia memiliki tantangan. Pertama, konten tidak memiliki kalimat dengan struktur baku. Kedua, domain dari setiap konten sangat luas dan heterogen, sehingga sulit untuk mengelompokkan topik tersebut dan mengklasifikasikan sentimen. Ketiga, penggunaan slang menyebabkan ambiguitas dan Out of Vocabulary (OOV). Di dalam penelitian ini diusulkan metode K-Medoid Clustering dan Support Vector Machine (SVM) untuk melakukan klasifikasi berdasarkan fitur sentiment score. Tujuan dari penggunaan K-Medoid clustering untuk mengeliminasi data yang tidak sesuai dan mengelompokkan data sesuai dengan homogenitas topik. Penelitian ini dilakukan dalam skenario uji coba perbandingan hasil analisa sentimen setelah dilakukan clustering berdasarkan fitur sentiment score terhadap fitur berbasis konten dengan parameter tingkat akurasi yang dianalisa dengan uji t. Hasil akhir menunjukkan bahwa metode yang diusulkan memiliki kualitas lebih baik daripada metode sebelumnya yang menggunakan fitur berbasis konten.

Kata Kunci: analisa sentimen, SVM, K-Medoid, clustering

Twitter merupakan media sosial yang banyak digunakan di Indonesia dan memiliki persebaran serta distribusi informasi yang sangat cepat. Pengguna Twitter yang berasal dari berbagai kalangan dan disiplin ilmu menyebabkan aliran komunikasi dan opini di media ini menjadi lebih beragam, salah satunya adalah kritikan dan komentar mengenai kebijakan pemerintah. Isu politik dan kebijakan pemerintah banyak dibahas oleh pengguna twitter Indonesia, sejak 2015 terpilihnya pasangan Joko Widodo-Jusuf Kalla dengan dengan prosentase akumulasi suara sebesar 53,15%, mengalahkan Prabowo-Hatta Rajasa, yang memperoleh akumulasi suara sebesar 46,85% [1]. Perbedaan persentase kemenangan yang begitu tipis memicu pro kontra di masyarakat, tak terkecuali pengguna Twitter. Oleh karena itu perlu dilakukan analisis sentimen terhadap konten Twitter berbahasa Indonesia. Hasil analisis sentimen dapat menjadi tolok ukur pengambilan kebijakan pemerintah selanjutnya, serta mengetahui bagaimana reaksi masyarakat terhadap kebijakan tersebut.

Proses klasifikasi sentimen dari konten media sosial memiliki beberapa tantangan. Pertama, bahasa yang digunakan di setiap konten media sosial sering kali tidak memiliki struktur formal dan baku, seperti penggunaan singkatan, perubahan huruf menjadi angka, kurangnya tanda baca, dan lain lain. Kedua, konten kalimat yang ada pada media sosial merupakan domain independen yang sangat luas (teknologi ekonomi, desain, dll), sehingga pengguna dapat membahas tentang banyak hal dan dalam domain yang berbeda. Hal ini membuat proses klasifikasi sentimen menjadi semakin sulit. Selain itu, penggunaan bahasa *slang* dapat menyebabkan ambiguitas dan OOV (Out of Vocabulary). Hal tersebut dapat menurunkan akurasi dari hasil klasifikasi sentimen. Oleh karena itu, diperlukan metode yang tepat untuk melakukan *preprocessing* data dan topik, serta pemilihan fitur yang dapat merepresentasikan isi dari konten media sosial.

Berdasarkan permasalahan diatas, Di dalam penelitian ini diusulkan metode K-Medoid Clustering dan Support Vector Machine (SVM) untuk melakukan klasifikasi berdasarkan fitur sentiment score. K-Medoids clustering

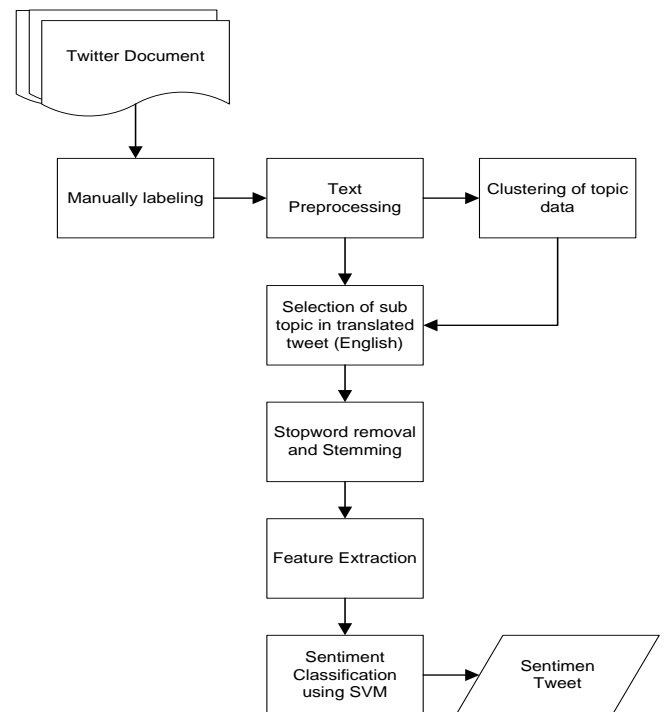
digunakan untuk membuang data-data yang tidak merepresentasikan topik yang diinginkan dan mengelompokkan data berdasarkan kesesuaian topik yang dibahas. Fitur tambahan yakni penggunaan kalimat negasi dan kata tanya juga diharapkan dapat meningkatkan presisi dan akurasi dari sistem. Hal ini berbeda dengan penelitian sebelumnya mengenai analisis sentimen media sosial Indonesia yang sering menggunakan fitur berdasarkan isi data teks twitter [2] [3].

Tujuan dari penelitian ini adalah mengembangkan sistem klasifikasi analisa sentimen menggunakan SVM (Support Vector Machine) dan K-medoid Clustering pada data twitter terhadap kebijakan pemerintah Indonesia. Dalam studi ini, kami mengembangkan sistem menjadi 4 bagian utama, *pre-processing* teks berderau, pengelompokan topik (*clustering*), ekstraksi fitur, kemudian klasifikasi sentimen. Proses pertama bertujuan untuk meminimalkan kosakata istilah yang digunakan dalam pesan teks yang cenderung menggunakan kata-kata informal yang umum untuk menggantikan kata-kata formal. Bagian kedua bertujuan untuk memilih data yang diinginkan pada topik kebijakan pemerintah, menghilangkan data yang tidak tepat. Ekstraksi fitur akan menghasilkan nilai fitur yang kemudian digunakan dalam tahapan klasifikasi. Tahapan terakhir adalah tahapan klasifikasi menggunakan SVM dan menganalisis bagaimana kinerja metode yang diusulkan dibandingkan dengan fitur konten berbasis teks

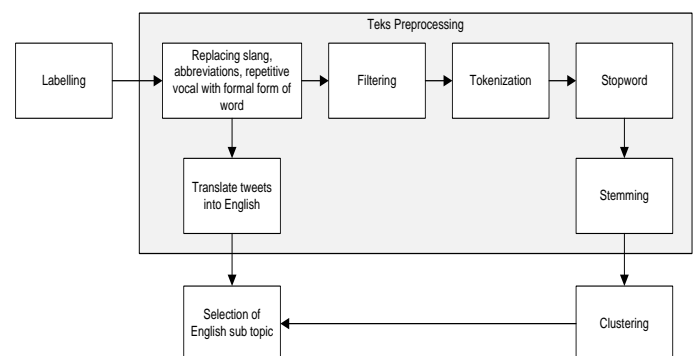
I. Metodologi

Sistem analisis sentimen dalam penelitian ini dibagi menjadi empat bagian, yaitu *preprocessing*, *data clustering*, ekstraksi fitur dan klasifikasi. Preprocessing digunakan untuk mempersiapkan kalimat tweet yang siap diolah pada tahapan selanjutnya, *data clustering* bertujuan untuk mengelompokkan tweet terkait dengan topik yang diinginkan. Setelah itu fitur konten diekstrak menggunakan *sentiment score*. Selama tahapan ekstraksi fitur, setiap konten tweet diekstraksi untuk mendapatkan vektor dimensi ke-n. Klasifikasi terdiri dari pelatihan dan pengujian menggunakan metode Support Vector Machine (SVM).

Sistem arsitektur analisis sentimen ditunjukkan pada Gambar 1. Pada Gambar 1, fase preprocessing terdiri dari beberapa proses, seperti parsing json, pelabelan, mengubah kata-kata *slang* ke dalam kata-kata standar, clustering, dan terjemahan ke Bahasa Inggris. Output dari fase preprocessing adalah kumpulan data yang akan disebut sebagai sub topik. Dalam ekstraksi fitur, dilakukan proses penrhitungan bobot (*score*) dengan standar SentiWordNet, dan memeriksa apakah terdapat kalimat pertanyaan. Kemudian, pada tahapan klasifikasi dilakukan pelatihan dan proses pengenalan menggunakan Support Vector Machine (SVM).



Gambar 1. Arsitektur sistem



Gambar 2. Tahapan text processing

Pemberian label (*Labelling*)

Pelabelan dilakukan secara manual dengan memberikan label setiap tweet yang dengan positif, negatif, atau netral. Tweet yang telah disimpan dalam file kemudian ditambahkan dengan label pada awal tweet dan dipisahkan oleh karakter tab (`\t`).

Pemrosesan text (*Text processing*)

Text preprocessing merupakan salah satu proses penting dalam analisis sentimen. Tujuan dari *text processing* adalah untuk membersihkan data teks sehingga siap untuk diekstraksi. Adapun pemrosesan teks yang dilakukan adalah sebagai berikut.

- **Mengubah Bahasa Slang**

Didalam tweet berbahasa indonesia, pengguna lebih suka memakai bahasa slang dari pada bahasa baku. Yang dimaksud bahasa slang di penelitian ini adalah penggunaan angka sebagai pengganti alphabet seperti kata "ga21", pengulangan karakter vokal seperti "naaiiik", dan penggunaan kata tidak baku seperti "cemungudh". Semua kata slang akan dimasukkan dalam kamus slang yang kemudian digunakan untuk mengganti kata-kata slang didalam tweet menjadi kata yang baku sesuai dengan KBBI [4].

- **Terjemah ke bahasa Inggris**

Tweet dalam bahasa Indonesia akan diterjemahkan ke dalam bahasa Inggris. Tujuan dari terjemahan ini adalah untuk mendapatkan nilai dari bahasa Inggris SentiWordNet dengan mudah.

- **Filtering**

Pada tahapan filtering, data dibersihkan dari kata words atau character yang tidak sesuai. Kata yang akan dihilangkan adalah hashtag yang dilambangkan dengan karakter "#", "RT" (retweet), username yang ditandai dengan karakter "@", and the menghilangkan tanda baca seperti koma (,) dan titik (.) [5].

- **Tokenization**

Dalam proses tokenization, data tweet yang berupa kalimat akan dipisahkan per kata berdasarkan karakter spasi dan tanda baca. Setelah itu, kata yang memiliki karakter non-abjad dan angka dihilangkan. Selain karakter abjad dan angka, garis bawah juga dipisahkan sesuai dengan posisi karakter [5].

- **Stopword removal**

Ada dua proses stopwords dalam penelitian ini, Indonesia stopwords dan stopwords Inggris. Dalam proses stopwords Indonesia, pasalnya dokumen akan dihapus, seperti "di", "oleh", "pada", "sebuah", "karena" dan lain-lain. Jika kata termasuk dalam stoplist maka kata ini akan dihapus dari deskripsi, maka kata-kata yang tersisa dalam deskripsi dianggap sebagai kata-kata yang menjadi ciri isi dokumen atau kata kunci. Setelah itu kata-kata diuraikan ke dalam kata-kata dasar. Sementara dalam proses stopwords Inggris, menghapus artikel dalam bahasa Inggris juga telah dilakukan dari susunan kalimat kata-kata, seperti "a", "an", "the" [6].

- **Stemming**

Proses Stemming digunakan untuk mendapatkan bentuk dasar dari semua kata yang membentuk tweet. Algoritma yang digunakan pada penelitian ini merupakan *library* dari Apache Lucene [5]. Di dalam proses stemming, ada juga proses penghapusan stopwords. Stopword removal sangat berguna untuk menghilangkan kata-kata yang dianggap tidak terlalu penting dalam tweet. Contoh daftar stopwords adalah 'pada,' adalah ',' tidak ',' lalu ', dan seterusnya.

Pengelompokan data (*Clustering*)

Tweet diperoleh dalam setiap topik kebijakan pemerintah tidak selalu homogen terdapat tweet outlier yang ikut terambil. K-Medoid *Clustering* digunakan untuk memperoleh data yang memiliki tingkat kesamaan yang tinggi (homogen). Jumlah cluster (k) yang digunakan dalam penelitian ini berkisar antara 2 sampai 30. Hasil cluster akan dianggap sebagai sub topik. Namun, sub-topik yang berisi kurang dari 20 baris atau memiliki nilai Average Width Silhouette (ASW) [7] yang kecil akan dihapus.

Pada saat pengelompokan tweet dengan menggunakan K-medoids, perhitungan kesamaan dilakukan untuk setiap tweet yang sedang diproses dengan semua tweets yang ada di dalam cluster. Secara umum, proses pengelompokan dokumen twitter yang ditunjukkan pada Gambar 3. Fungsi kesamaan yang digunakan adalah pencocokan kesamaan berbasis uni-gram ditunjukkan dalam Persamaan 1 [8].

$$sim(s_i, s_j) = \frac{2 * |s_i \cap s_j|}{|s_i| + |s_j|} \tag{1}$$

Dimana s_i dan s_j adalah tweet ke-i dan ke-j. $|s_i| \cap |s_j|$ merupakan jumlah kata yang sesuai. $|s_i|$ adalah jumlah kata yang membentuk tweet i. $|s_j|$ adalah panjang tweet ke-j

Algoritma of K-medoids
 Input : $D = \{d1, d2, \dots, dn\}$
 1. Tentukan jumlah kluster yang akan dibentuk (k)
 2. Pilih k objek secara acak pada himpunan objek sebagai medoid
 3. Hitung kemiripan (*similarity*) antara non-medoids dengan k medoids
 4. Pilih jarak yang terbaik antara non-medoids object dengan medoids object
 5. Perbaharui medoid dengan memilih secara acak non-medoids object
 6. Hitung total cost S (current total cost – past total cost)
 7. Jika current total cost > past total cost maka ubah medoid, else medoids tidak diubah
 8. Ulangi langkah 3 to 7 hingga tidak ada perubahan dari cost value.

Gambar 3. Pseudocode K-Medoid Clustering

Klasifikasi sentimen

Proses klasifikasi dibagi menjadi dua proses, proses pelatihan dan proses pengenalan. Metode yang digunakan untuk klasifikasi adalah Support Vector Machine (SVM). Jenis SVM dan kernel yang digunakan dalam penelitian ini adalah C-SVC dengan kernel Radial Basis Function (RBF). Beberapa tweet dilatih untuk membentuk model klasifikasi yang akan menjadi model yang digunakan dalam proses pengenalan. Parameter pelatihan yang digunakan adalah gamma (γ) dan cost (C). Kisaran parameter gamma (γ) yang digunakan bernilai 0,1, 0,3, 0,5, 0,7, dan 0,9. Sedangkan cost (C) yang digunakan berada pada nilai 100, 300, 500, 700, dan 900.

II. Hasil dan Pembahasan

Dalam bagian ini, hasil penelitian dan pembahasan akan ditampilkan. Bagian pertama, adalah penjelasan dari dataset dan pengolahannya. Setelah itu penjelasan dari skenario dan hasil evaluasi akan dijelaskan. Kemudian ditampilkan hasil pengukuran kualitas kluster. Dan bagian terakhir adalah hasil perbandingan antara fitur *sentiment score*

dan fitur berbasis konten akan disajikan menggunakan parameter akurasi dengan evaluasi uji-t.

Pemrosesan Dataset

Data uji yang digunakan dalam penelitian ini diunduh dari data Twitter berbahasa Indonesia. Tweet diperoleh dari proses *crawling* dengan menggunakan Topsy’s APIs dengan bahasa pemrograman Java. Konten data diambil dari websites <http://otter.topsy.com> yang secara teknis memiliki akses unduh pengunduhan data twitter. Tahapan pengumpulan data dilakukan dengan melakukan *filtering* atau seleksi data dengan menggunakan kata kunci spesifik. Kata kunci yang digunakan dalam penelitian ini dibagi menjadi empat topic terkait kebijakan pemerintah Indonesia. Table 1 adalah penjelasan detail topic dan kata kunci yang digunakan selama pengunduhan data konten Twitter. Sedangkan Tabel 2 merupakan contoh data konten tweet yang sudah dibubuhi label sentiment.

Tabel 1. Topik dan kata kunci pada *dataset*

Topic of governme nt policy	Keyword	Date of data retrieval	Number of tweet
increasin g of fuel price	"kenaikan+bbm",	12-11-2012	1672
	"pengalihan+subsidi",	until 21-11-2015	
diversion of 12 kg LPG support	"kenaikan+bbm",	11-2015	97
	"pengalihan+subsidi+bbm"	01-04-2015	
KPK vs POLRI polemic	"lpg+12kg",	01-04-2015	2580
	"subsidi+lpg+12kg",	until. 20-04-2015	
President Jokowi’s resolutio ns	"subsidi+lpg", "lpg+12kg",	10-01-2015	2640
	"subsidi+lpg+12kg",	until 26-01-2015	
President Jokowi’s resolutio ns	"kpk+vs+polri",	10-01-2015	2640
	"save+kpk",	until 26-01-2015	
	"polemik+kpk+polri",	01-09-2014	
	"pemberantasan+korupsi"	until 28-02-2015	
President Jokowi’s resolutio ns	"anti+jokowi",	01-09-2014	2640
	"jokowi+hater",	until 28-02-2015	
	"jokowi+lover",	02-2015	
	"kebijakan+jokowi",		
	"100+hari+kerja+jokowi",		
	"respon+kebijakan+Jokowi"		

Tabel 2. Contoh konten tweet yang sudah diberi label

Label	Indonesian tweet
2	Demo Mahasiswa Tolak Kenaikan BBM Diwarnai Aksi Dorong
1	Komisi VI DPR Minta Alasan Kenaikan BBM
3	Jokowi: Manfaat Kenaikan BBM akan Nampak Tahun Depan

Note: 1: neutral, 2: negative, 3: positive

Proses crawling data diberi batasan hanya pada pengambilan konten twitter berbahasa Indonesia dengan menambahkan batasan *geolocation* Indonesia pada pengaturannya yakni (94, -11, 141, 6). Proses pengunduhan data dilakukan secara berkala dengan total data sebanyak 6989 data. Data tweet kemudian dipisahkan dan dikelompokkan topic yang berbeda sesuai dengan kata kunci.

Tahapan selanjutnya adalah dilakukannya pembersihan *dataset* dengan metode *text processing* yang hasilnya disimpan dalam bentuk.txt file. Kemudian dilakukan tahapan *clustering* untuk perbaikan domain data dan penyeleksian kandidat data terbaik.

Setelah proses *clustering* dilakukan pengujian kualitas tweet, kemudian tweet dibagi menjadi beberapa sub topik yang sesuai untuk kemudian dilakukan ekstraksi fitur.

Evaluasi 1: Pengukuran kualitas kluster

Tujuan dari evaluasi 1 adlah mengukur kualitas kulster yang terbentuk dari proses *clustering* menggunakan algoritma K-medoid. *Clustering* adalah satu metode yang disulkan dalam penelitian ini yang bertujuan untuk mengelompokkan data menjadi lebih homogen.

Skenario uji coba pengujian kualitas kluster dilakukan dengan mencoba beberapa variasi nilai centroid (*k*) untuk mendapatkan hasil yang paling optimal. Pengukuran kualitas kluster diukur dengan menggunakan metode *Silhouette evaluation*.

Salah satu faktor yang mempengaruhi hasil *clustering* adalah penentuan nilai centroid (*k*). Pada penelitian ini, kasus uji ditentukan dengan variasi nilai *k*, dimulai dari *k* = 2 sampai *k* = 30. Kemudian hasil *clustering* dari setiap nilai *k* tersebut dihitung dan divalidasi menggunakan Silhouette (*s*) tweet ke *i* sesuai dengan persamaan 2 dan persamaan 3.

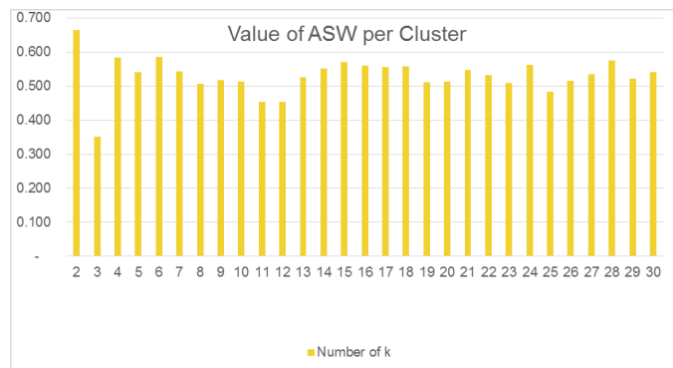
$$b(x) = \max_{C_i \neq A} d(x, C_i) \tag{2}$$

$$S_i = \frac{a(x) - b(x)}{\max(a(x), b(x))} \tag{3}$$

Dimana titik *x* (*a_i*) di kluster A, dan *a(x)* adalah jarak rata-rata antara titik *a_i* ke anggota titik lain yang ada pada kluster A, dan *b(x)* adalah nilai tertinggi dari jarak rata-rata titik *a_i* dengan anggota titik lainnya pada kluster yang berbeda [20].

Then, dilakukan pengukuran nilai ASW (Average Silhouette Width) pada setiap kluster sesuai dengan persamaan 4. Hasil *clustering* dengan nilai *k* terbaik akan memiliki ASW yang tertinggi.

Pada tahapan ini dilakukan pengujian pada setiap subtopic (*cluster*) menggunakan *Silhouette Coefficient*. Subtopik yang memiliki nilai $ASW \geq 0.5$ akan dipilih menjadi kandidat data latih dan data uji. Namun, jika tidak ada subtopik yang memiliki nilai $ASW \geq 0.5$, maka akan dipilih subtopik yang memiliki nilai ASW tertinggi dan memiliki jumlah data yang paling banyak. Nilai rata-rata ASW untuk setiap kasus uji ditampilkan pada Gambar 4.



Gambar 4. Rata- rata ASW per kasus uji (*k* ∈ 2,3,4, ...,30)

Penentuan kualitas kluster mengacu pada penelitian yang dilakukan oleh Rousseeuw [20] seperti yang ditampilkan pada Tabel 3.

Tabel 3. Interpretasi nilai ASW (Average Silhouette Width)

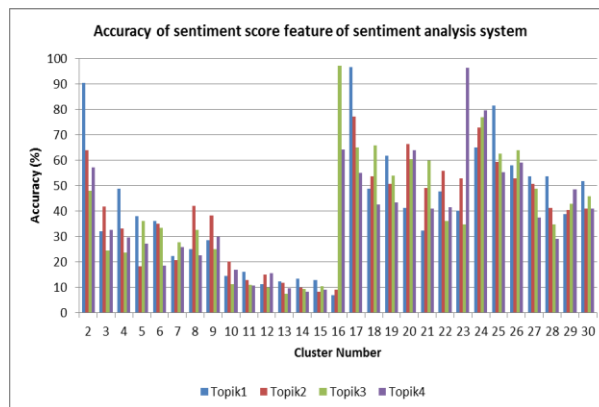
ASW	Intepretasi
$0.71 \leq ASW \leq 1$	A strong structure has been found
$0.51 \leq ASW < 0.71$	Reasonable, A reasonable structure has been found (2.12)
$0.26 \leq ASW < 0.51$	The structure is weak and could be artificial. Try additional methods of data analysis (2.13)
$ASW < 0.26$	No substantial structure has been found (2.14)

Gambar 4 menunjukkan kualitas kluster dari setiap kasus uji hampir sama. Kasus uji dengan jumlah kluster $k \in 2,3, \dots, 10$ memiliki kualitas yang baik (**good**) dan masuk akal (**reasonable**). Kasus uji dengan jumlah kluster $k \in 13, \dots, 24, 27, 28, 29, 30$ juga memiliki kualitas yang baik (**good**) and masuk akal (**reasonable**). Namun, kluster $k \in 3, 11, 12, \text{ and } 25$ tidak lah memiliki kualitas yang baik (**quite good**) dan lemah (**weak**). Dengan demikian diseleksi lah data dengan kualitas kluster yang baik untuk menjadi data latih dan data uji.

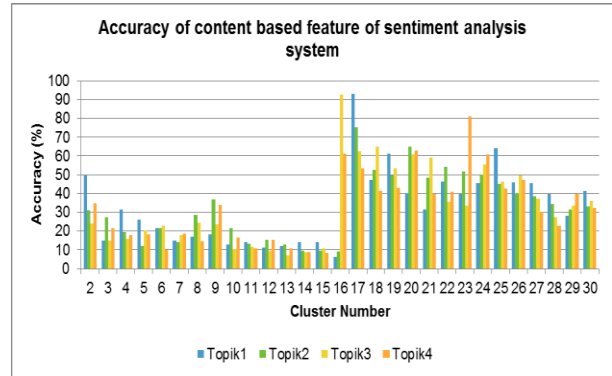
Evaluasi 2: Evaluasi parameter pada SVM menggunakan tingkat akurasi

Pada bagian ini akan ditampilkan analisa dan hasil evaluasi terhadap parameter SVM menggunakan tingkat akurasi dan uji-t dari metode yang diusulkan dengan menggunakan *sentiment score feature* dan penelitian terdahulu yang menggunakan *content based feature*.

Grafik tingkat akurasi yang ditampilkan pada Gambar 5 merupakan nilai akurasi dari metode yang diusulkan yakni analisa sentiment dengan menggunakan *sentiment score feature*. Sedangkan Gambar 6 adalah grafik yang menunjukkan tingkat akurasi dari metode terdahulu yakni dengan menggunakan *content based feature*.



Gambar 5. Nilai akurasi analisa sentiment menggunakan *sentiment score feature*



Gambar 5. Nilai akurasi analisa sentiment menggunakan *content based feature*

Pada Gambar 5 terlihat bahwa nilai akurasi tertinggi untuk Topik1, Topik2, Topik3, dan Topik4 secara berturut-turut terletak pada jumlah kluster 17, 17, 16, dan 23 sedangkan akurasi terendah secara berturut-turut terletak pada jumlah kluster 16, 15, 13, dan 14. Secara umum akurasi dari semua topik bersifat fluktuatif atau naik-turun. Nilai akurasi antar topik dalam satu kluster bervariasi. Bervariasinya nilai akurasi topik dalam satu kluster bisa ditunjukkan dengan nilai standar deviasi yang tinggi. Hanya pada jumlah kluster 16 yang menunjukkan nilai standart deviasi tinggi.

Sama halnya seperti hasil akurasi pada Gambar 5, secara umum akurasi dari semua topik pada Gambar 6 juga bersifat fluktuatif atau naik-turun. Akurasi tertinggi untuk Topik1, Topik2, Topik3, dan Topik4 secara berturut-turut terletak pada jumlah kluster 17, 17, 16, dan 23 sedangkan akurasi terendah secara berturut-turut terletak pada jumlah kluster 16, 16, 13, dan 15. Nilai akurasi antar topik dalam satu kluster pada Gambar 4.2 juga bervariasi. Bervariasinya nilai akurasi topik dalam satu kluster bisa ditunjukkan dengan nilai standar deviasi yang tinggi. Nilai standar deviasi tinggi ditunjukkan pada jumlah kluster 16.

Untuk menentukan sistem yang lebih baik tidak dapat hanya dilihat dari nilai akurasi semata. Untuk membuktikan bahwa nilai akurasi dari metode yang diusulkan yaitu analisa sentimen dengan *sentiment score feature* memiliki kualitas tingkat akurasi lebih baik dibandingkan metode terdahulu yang menggunakan *content based feature*, diperlukan uji statistik. Uji statistik yang dipakai pada penelitian ini adalah uji t dua kelompok berpasangan. Uji t digunakan untuk

mendapatkan apakah ada perbedaan signifikan antara akurasi analisis sentimen berdasarkan *sentiment score feaure* dengan akurasi analisis sentimen berdasarkan *content based feature*. Jika ada perbedaan signifikan antara akurasi berbasis fitur dengan akurasi berbasis konten, maka salah satu dari kedua akurasi tersebut mempunyai rata-rata akurasi yang lebih tinggi.

Nilai alpha yang digunakan adalah 5%. Hipotesis nol (H_0) yang diambil adalah akurasi analisis sentimen berdasarkan *sentiment score feature* lebih kecil sama dengan akurasi analisis sentimen berdasarkan *content based feature*. Hipotesis satu (H_1) adalah akurasi analisis sentimen berdasarkan *sentiment score feature* lebih besar dari akurasi analisis sentimen berdasarkan *content based feature*. Setelah dilakukan perhitungan statistik didapatkan nilai t hitung dan nilai t kritis dari nilai akurasi analisis sentimen seperti yang ditunjukkan pada Tabel 4.

Tabel 4. Nilai t hitung dan nilai t kritis

	Topik 1	Topik 2	Topik 3	Topik 4
Nilai t hitung	5.94	5.47	6.48	6.28
Nilai t kritis	1.7	1.7	1.7	1.7

Pada Tabel 4 dapat terlihat bahwa semua topik memiliki nilai t hitung yang lebih tinggi dari nilai t kritis. Dengan demikian dapat diambil keputusan bahwa H_0 ditolak sehingga dapat disimpulkan bahwa akurasi analisis sentimen berdasarkan *sentiment score feature* lebih tinggi dari akurasi analisis sentimen berdasarkan *content based feature* pada semua topik yang diujikan.

III. Simpulan

Penelitian ini telah memaparkan sebuah kerangka kerja (*framework*) baru dalam melakukan analisa sentimen pada kebijakan pemerintah Indonesia pada konten data Twitter berbahasa Indonesia. Penelitian ini dibangun dengan algoritma Support Vector Machine (SVM) yang dikombinasikan combined with K-Medoid Clustering. Tujuan dari penelitian ini adalah menawarkan fitur *sentiment score* yang berdasarkan pada SentiWordNet untuk

mengklasifikasikan sentiment dari data Twitter berbahasa Indonesia yang dapat menghasilkan nilai akurasi lebih tinggi dibandingkan metode sebelumnya yang menggunakan fitur konten.

Tahapan *clustering* menggunakan K-Medoid dilakukan untuk melakukan eliminasi data yang tidak sesuai dan mendapatkan data yang lebih homogen sesuai topik kebijakan pemerintah yang ingin diteliti.

Tahapan pengujian dan evaluasi dilakukan dengan menggunakan data sejumlah 6989 tweet. Data tersebut kemudian dikelompokkan menjadi kasus uji berdasarkan jumlah kluster ($k \in 2,3,4, \dots, 30$). Kemudian dilakukan pengujian parameter SVM untuk setiap kasus kasus uji seperti yang telah dibahas pada bab II.

Dari hasil uji coba menggunakan tingkat akurasi yang diperkuat dengan pengujian statistik dapat diketahui bahwa metode yang diusulkan dengan menggunakan *sentiment score feature* lebih unggul dibandingkan metode sebelumnya yakni *content based feature*. Hal ini dibuktikan dengan evaluasi uji- t . Pada masing-masing pengujian evaluasi tingkat akurasi, baik metode yang diusulkan maupun metode terdahulu sama-sama memiliki kecenderungan yang hampir sama untuk setiap kasus uji, yakni memiliki nilai akurasi yang tinggi ketika jumlah kluster k juga tinggi.

Dapat dilihat secara jelas setelah dilakukan uji- t nilai alpha 5% dibuktikan bahwa tingkat akurasi dari metode analisa sentimen yang diusulkan menggunakan *sentiment score feature* lebih tinggi dari tingkat akurasi *content based feature*. Metode yang diusulkan memiliki tingkat akurasi yang lebih unggul di semua topik kebijakan pemerintah yang menjadi kasus uji. Dengan demikian dapat disimpulkan bahwa analisa sentimen dengan menggunakan *sentiment score feature* memiliki kinerja yang lebih baik daripada fitur berbasis konten.

Dalam penelitian ini proses pelabelan setiap tweet yang masih melakukan secara manual dan masih menggunakan Senti WordNet berbahasa Inggris, untuk pengembangan selanjutnya akan dikembangkan sumber leksikal SentiWordNet berbahasa Indonesia. Selain itu pelabelan secara otomatis

juga akan dikembangkan sesuai dengan konten twitter berbahasa Indonesia.

IV. Daftar Pustaka

- [1] Lingkar Survey Indonesia, "100 Hari Jokowi: 3 Rapor Merah, 2 Rapor Biru", Jakarta, 29th January 2015, <http://lsi.co.id/lsi/2015/01/29/100-hari-jokowi-3-rapor-merah-2-rapor-biru/>.
- [2] Naradhipa, Aqsath Rasyid, and Ayu Purwarianti. "Sentiment classification for Indonesian message in social media." *Cloud Computing and Social Networking (ICCCSN), 2012 International Conference on*. IEEE, 2012.
- [3] Harlili and Wibisono, Yudi. "Sistem Analisis Opini Microblogging Berbahasa Indonesia". Konferensi Nasional Sistem Informasi. Bali. Indonesia, 2012
- [4] Lunando, Edwin, and Ayu Purwarianti. "Indonesian social media sentiment analysis with sarcasm detection." *Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference on*. IEEE, 2013.
- [5] Tala, Fadillah Z. "A study of stemming effects on information retrieval in Bahasa Indonesia." *Institute for Logic, Language and Computation Universeit Van Amsterdam* (2003).
- [6] arkar, Kamal. "Sentence Clustering-based Summarization of Multiple Text Documents." *International Journal of Computing Science and Communication Technologies* 2.1 (2009): 325-335.
- [7] Kaufman, Leonard, and Peter J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis*. Vol. 344. John Wiley & Sons, 2009.
- [8] Erkan, Günes, and Dragomir R. Radev. "LexRank: Graph-based lexical centrality as salience in text summarization." *Journal of Artificial Intelligence Research* (2004): 457-479.