

Analisis Sentimen Kurikulum 2013 pada Twitter menggunakan *Ensemble Feature* dan Metode *K-Nearest Neighbor*

M. Rizzo Irfan¹, Mochammad Ali Fauzi², Tibyani³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹m.rizzoirfan@gmail.com, ²moch.ali.fauzi@ub.ac.id, ³tibyani@ub.ac.id

Abstrak

Kurikulum 2013 merupakan kurikulum baru dalam sistem pendidikan Indonesia yang telah diberlakukan oleh pemerintah untuk menggantikan kurikulum 2006 atau Kurikulum Tingkat Satuan Pendidikan (KTSP). Diberlakukannya kurikulum ini pada beberapa tahun terakhir memicu berbagai kontroversi dalam dunia pendidikan Indonesia, seperti siswa yang dituntut lebih aktif, jam pelajaran yang ditambah dan hal-hal lainnya yang menyebabkan muncul berbagai opini yang berkembang di masyarakat terutama pada Twitter. Diperkirakan sekitar 200 juta pengguna Twitter melakukan *posting* 400 juta *tweet* per hari. Dalam penelitian ini, dilakukan analisis sentimen untuk mengetahui opini yang berkembang tersebut yang dibagi ke dalam opini positif atau opini negatif. Fitur dan metode yang digunakan adalah *ensemble feature* dan metode klasifikasi *K-Nearest Neighbor* (K-NN). *Ensemble feature* merupakan fitur gabungan, berupa fitur statistik *Bag of Words* (BoW) dan semantik (*twitter specific*, *textual features*, *PoS features*, *lexicon based features*). Berdasarkan serangkaian pengujian, kombinasi fitur berdampak dalam meningkatkan akurasi metode *K-Nearest Neighbor* (K-NN) untuk menentukan opini positif atau negatif. Penggabungan fitur ini dapat melengkapi kelemahan masing-masing fitur, sehingga hasil akhir akurasi yang didapatkan dengan menggabungkan kedua fitur tersebut mencapai 96%. Berbeda hal jika hanya menggunakan fitur secara independen saja, akurasi yang didapatkan hanya mencapai 80% pada fitur *Bag of Words* (BoW) dan 82% pada fitur *ensemble* tanpa *Bag of Words* (BoW).

Kata kunci: *analisis sentimen, kurikulum 2013, twitter, ensemble feature, k-nearest neighbor*

Abstract

The 2013 curriculum is a new curriculum in the Indonesian education system that has been enacted by the government to replace the 2006 curriculum or the Education Unit Level Curriculum. The implementation of this curriculum in recent years has sparked controversy in Indonesian education, students who are demanded more actively, added lessons and other matters that lead to various opinions that develop in the community, especially on Twitter. An estimated 200 million Twitter users post 400 million tweets per day. In this research, sentiment analysis is conducted to find out the developing opinion which is divided into positive opinion or negative opinion. The features and methods used are the ensemble feature and the K-Nearest Neighbor (K-NN) classification method. Ensemble feature is a combined feature, in the form of statistical Bag of Words (BoW) and semantic features (*twitter specific*, *textual features*, *PoS features*, *lexicon based features*). Based on a series of tests, the combination of features has an impact in improving the accuracy of the K-Nearest Neighbor (K-NN) method for determining positive or negative opinions. Merging this feature can complement the weaknesses of each feature, so the final result of accuracy gained by combining both features reaches 96%. In contrast to using only features independently, the accuracy achieved only reaches 80% on Bag of Words (BoW) features and 82% on ensemble features without Bag of Words (BoW).

Keywords: *sentiment analysis, curriculum 2013, twitter, ensemble feature, k-nearest neighbor*

1. PENDAHULUAN

Menurut survei yang telah dilakukan oleh

International Data Corporation (IDC) yaitu, sebuah badan *market research* di Amerika, pada tahun 2013 hingga 2020 jumlah pertumbuhan informasi digital akan terus tumbuh, yaitu dari 4

triliun *gigabytes* menjadi 44 triliun *gigabytes*. Informasi tersebut meningkat lebih dari dua kali lipat setiap dua tahun sekali. Hal ini sepadan dengan meningkatnya jumlah pengguna media sosial saat ini untuk dapat bertukar informasi dengan lebih cepat dari berbagai belahan dunia. Namun, tidak seluruh informasi yang ditampilkan selalu memiliki nilai opini yang baik. Terdapat opini yang beragam bisa bernilai positif atau negatif terhadap topik tertentu yang sedang hangat dibicarakan.

Salah satu informasi yang beredar luas saat ini adalah opini mengenai kurikulum 2013. Kurikulum 2013 merupakan kurikulum baru dalam sistem pendidikan Indonesia yang telah diberlakukan oleh pemerintah untuk menggantikan kurikulum 2006 atau yang sering disebut sebagai Kurikulum Tingkat Satuan Pendidikan (KTSP). Adanya kurikulum baru ini menuai berbagai kontroversi masyarakat. Beberapa penyebab disinyalir yang menjadi perbincangan seperti, siswa dituntut harus aktif, guru hanya menyampaikan materi lalu siswa harus mencari tahu sendiri, ada beberapa pelajaran yang dihapuskan, mewajibkan Pramuka dan hal-hal lain yang semakin memancing berbagai opini mengenai topik tersebut terutama di kalangan pengguna Twitter.

Twitter merupakan salah satu media sosial yang besar dan digemari oleh masyarakat, pengguna dapat melakukan *posting* status atau pesan yang disebut sebagai *tweet* yang tidak lebih dari 140 karakter. Sebagian besar kasus, para pengguna menuliskan pesan mereka jauh lebih sedikit dari batasan karakter yang telah ditentukan. Twitter merupakan salah satu media sosial penyumbang data terbesar dan dinamis berdasarkan konten buatan pengguna. Diperkirakan sekitar 200 juta pengguna Twitter melakukan *posting* 400 juta *tweet* per hari (Da Silva, Hruschka and Hruschka, 2014).

Dalam penelitian kali ini, analisis sentimen dilakukan dengan berbagai tahapan agar dapat memproses data *tweet* tersebut, akan dilakukan tahapan dalam *text mining* seperti *preprocessing*, lalu diikuti proses pengambilan nilai fitur hingga proses pengklasifikasian. Pada tahapan pengambilan nilai fitur dilakukan menggunakan *ensemble feature*. Berdasarkan penelitian sebelumnya yang dilakukan oleh Siddiqua, Ahsan dan Chy pada tahun 2016 mengusulkan *ensemble feature*, yaitu fitur gabungan secara statistik dan semantik. Terdapat 5 kategori fitur yang digunakan, yaitu dari sisi semantik kata menggunakan *twitter specific features*, *textual*

features, *Parts of Speech (PoS) features*, *lexicon based features* dan dari sisi statistik kata menggunakan *Bag of Words (BoW)*. Hasil dari penelitian tersebut menunjukkan penggunaan dari fitur *ensemble* secara lengkap dapat meningkatkan akurasi dan mendapati akurasi tertinggi jika dibandingkan dengan penggunaan fitur secara terpisah. Penggunaan fitur secara terpisah, yaitu *Bag of Words (BoW)* saja hanya mencapai akurasi sebesar 73,8%, sedangkan jika menggabungkan *Bag of Words (BoW)* dengan keseluruhan fitur semantik mencapai akurasi sebesar 87,7%. Penggunaan fitur *ensemble* ini memiliki akurasi yang lebih baik dibandingkan dengan fitur lainnya seperti *unigram+bigram* 83,3%, *label propagation* 84,7%, *sentiment topic features* 86,7%, *sentistrength* 81,7%, *meta level features* 81,6%, *semantria (online system)* 78,1% (Siddiqua, Ahsan and Chy, 2016).

Dalam hal pengklasifikasian opini, metode yang digunakan adalah *K-Nearest Neighbor*. *K-Nearest Neighbor (K-NN)* merupakan metode yang melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang kemiripannya paling dekat dengan objek tersebut (Hardiyanto and Rahutomo, 2016). Perhitungan kemiripan dapat dilakukan dengan menggunakan *cosine similarity*. Pengukuran kemiripan tersebut melibatkan variabel pada data uji dengan variabel yang terdapat di dalam data *training*. Sehingga, dengan melakukan pengukuran kemiripan akan didapatkan nilai dari kedekatan semua titik dari data *training* tersebut dengan data pada data uji. Pada penelitian sebelumnya yang dilakukan oleh Bijalwan et al. pada tahun 2014 menggunakan tiga metode untuk mengetahui akurasi hasil pengklasifikasian dari setiap metode. Pada penelitian tersebut metode yang digunakan adalah *Naive Bayes*, *Term Graph* dan *K-Nearest Neighbor (K-NN)*. Dari hasil penelitian, metode *K-Nearest Neighbor (K-NN)* menghasilkan nilai akurasi tertinggi ketika dibandingkan dengan *Naive Bayes* dan *Term Graph*. Rata-rata hasil akurasi didapatkan sebesar 98,95% untuk metode *K-Nearest Neighbor (K-NN)*, 62,66% untuk *Naive Bayes* dan 98,72% untuk *Term Graph*. Menurut penelitian tersebut *K-Nearest Neighbor (K-NN)* akan lebih cocok digunakan dibandingkan *Naive Bayes* dan *Term Graph* untuk melakukan *mining* pada suatu dokumen teks (Bijalwan et al., 2014).

2. LANDASAN KEPUSTAKAAN

Bagian ini berisikan dasar teori mengenai Kurikulum 2013, Twitter, *text mining*, *ensemble feature*, *vector space model* dan metode *K-Nearest Neighbor* (K-NN).

2.1. Kurikulum 2013

Kurikulum 2013 merupakan kurikulum baru dalam sistem pendidikan Indonesia yang telah diberlakukan oleh pemerintah untuk menggantikan kurikulum 2006 atau yang sering disebut sebagai Kurikulum Tingkat Satuan Pendidikan (KTSP). Kurikulum 2013 merupakan suatu model kurikulum yang dapat mengintegrasikan *skill*, *theme*, *concepts and topics* baik dalam bentuk *within single disciplines*, *across several disciplines* dan *within and across learners* (Poerwati and Amir, 2013).

Kurikulum 2013 merupakan kurikulum terpadu sebagai sebuah konsep sistem dan pendekatan pembelajaran yang melibatkan beberapa disiplin ilmu atau mata pelajaran/bidang studi untuk memberikan pengalaman yang bermakna dan luas kepada peserta didik (Nurdiana, 2015).

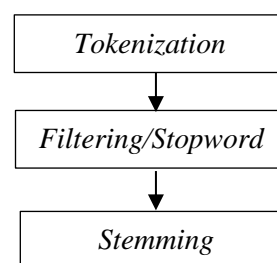
2.2. Twitter

Twitter merupakan layanan *microblogging* yang populer, pengguna dapat melakukan *posting* status atau pesan yang di sebut sebagai *tweet* yang tidak lebih dari 140 karakter. Sebagian besar kasus, para pengguna menuliskan pesan mereka jauh lebih sedikit dari batasan karakter yang telah ditentukan. Twitter merupakan salah satu media sosial penyumbang data terbesar dan dinamis berdasarkan konten buatan pengguna. Diperkirakan sekitar 200 juta pengguna Twitter melakukan *posting* 400 juta *tweet* per hari. *Tweet* ini dapat berupa ekspresi pendapat terhadap berbagai macam topik yang dapat membantu agar dapat berbagi pendapat terhadap para pengguna, identifikasi penyimpangan, kejadian yang menimbulkan ketidaknyamanan, prediksi perilaku politik dan olahraga, penerimaan atau penolakan politik dan semua komunikasi yang diutarakan dari mulut ke mulut (Da Silva, Hruschka and Hruschka, 2014).

2.3. Text Mining

Text mining atau pemrosesan teks merupakan suatu cara mencari untuk mendapatkan informasi dari suatu dokumen.

Dokumen tersebut diidentifikasi dan dieksplorasi. Proses untuk mendapatkan informasi tersebut melalui tahapan tertentu. Tahapan awal yang dapat dilakukan dalam *text mining* adalah dengan membakukan kata, hal ini dapat dilakukan untuk kasus data tertentu seperti data yang berasal dari *twitter* dengan bantuan kamus kata baku (Siddiqua, Ahsan and Chy, 2016). Setelah proses pembakuan kata selesai maka dapat dilanjutkan proses dalam *text mining* seperti pada umumnya. Secara umum proses untuk mendapatkan informasi dalam *text mining* mirip seperti yang terdapat pada data *mining*. Keduanya dilakukan dengan tahap *preprocessing* terlebih dahulu (Feldman and Sanger, 2007). Berikut alur dari *preprocessing* pada Gambar 1.



Gambar 1. Diagram alir *preprocessing*

2.4. Ensemble Feature

Ensemble Feature merupakan fitur gabungan dengan memperhatikan berbagai aturan-aturan tertentu. Terdapat fitur F1 – F37, keseluruhan fitur dapat dilihat pada Tabel 1.

Tabel 1. *Ensemble Feature*

Type	ID	Feature Description
Twitter Specific	F1	Whether the tweet contains a #hashtag or not.
	F2	Whether the tweet is a retweet or not.
	F3	Whether the tweet contains a user name or not.
	F4	Whether the tweet contains a URL or not.
Textual Features	F5	<i>TweetLength</i> : Number of words in the tweet.
	F6	<i>AvgWordLength</i> : Average character length of words.
	F7	Number of question marks available in the tweet.
	F8	Number of exclamation marks available in the tweet.
	F9	Number of quotes available in the tweet.
	F10	Number of words start with the uppercase letter in tweet.
	F11	Whether the tweet contains a

	F12	<i>positive emoticon or not. Whether the tweet contains a negative emoticon or not.</i>
PoS Features	F13	<i>Number of noun PoS available in the tweet.</i>
	F14	<i>Number of adjective PoS available in the tweet.</i>
	F15	<i>Number of verb PoS available in the tweet.</i>
	F16	<i>Number of adverb PoS available in the tweet.</i>
	F17	<i>Number of interjection PoS available in the tweet</i>
	F18	<i>Percentage of noun PoS in the tweet.</i>
	F19	<i>Percentage of adjective PoS in the tweet.</i>
	F20	<i>Percentage of verb PoS in the tweet.</i>
	F21	<i>Percentage of adverb PoS in the tweet.</i>
	F22	<i>Percentage of interjection PoS in the tweet.</i>
Lexicon Based Features	F23	<i>Number of positive words available in the tweet.</i>
	F24	<i>Number of negative words available in the tweet.</i>
	F25	<i>Number of positive words with adjective PoS.</i>
	F26	<i>Number of negative words with adjective PoS.</i>
	F27	<i>Number of positive words with verb PoS.</i>
	F28	<i>Number of negative words with verb PoS.</i>
	F29	<i>Number of positive words with adverb PoS.</i>
	F30	<i>Number of negative words with adverb PoS.</i>
	F31	<i>Percentage of positive words with adjective PoS.</i>
	F32	<i>Percentage of negative words with adjective PoS.</i>
	F33	<i>Percentage of positive words with verb PoS.</i>
	F34	<i>Percentage of negative words with verb PoS.</i>
	F35	<i>Percentage of positive words with adverb PoS.</i>
	F36	<i>Percentage of negative words with adverb PoS.</i>
	F37	<i>Number of intensifier words available in the tweet.</i>

Sumber: Siddiqua et al. (2016)

Fitur tersebut terbagi menjadi beberapa kategori fitur yaitu, *twitter specific features*, *textual features*, *Parts of Speech (PoS) features*, *lexicon based features*, dan *Bag of Words (BoW)*. Fitur tersebut akan menggunakan

lexicon berbahasa indonesia, kemudian pada pengambilan nilai fitur *emoticon* akan digunakan kamus *emoticon*. Fitur juga memanfaatkan *hashtag (#)* dan *retweet* untuk mengekstrak fitur spesifik. Pada kategori *textual features*, pengambilan nilai fitur yang dilakukan hanya berdasarkan informasi eksplisit yang terdapat dalam teks, seperti jumlah kata, jumlah kemunculan tanda petik/*quotes*, tanda seru, *emoticon* dan lainnya. Untuk bagian fitur *Bag of Words (BoW)* akan dilakukan pembobotan TF-IDF (Siddiqua, Ahsan and Chy, 2016).

2.5. Vector Space Model

Vector space model adalah model perolehan informasi yang menggunakan dokumen dan *query* sebagai vektor bermultidimensi, dimensi digunakan sebagai istilah untuk membentuk indeks yang mewakili setiap dokumen. *Vector space model* digunakan sebagai sarana untuk pencarian informasi atau *information retrieval*, *indexing* dan *relevancy ranking* yang menghasilkan pemodelan informasi yang dapat bekerja dengan baik pada banyak kasus tersebut. Penggunaan *vector space model* ini untuk membantu mendapatkan nilai dari fitur *Bag of Words (BoW)* dan terbagi dalam 3 tahapan, yaitu *document indexing*, *term weighting*, *cosine similarity* (Singh and Dwivedi, 2012).

Pada tahapan *document indexing* akan dilakukan ekstraksi tiap-tiap *term* atau kata yang terdapat dalam dokumen (Singh and Dwivedi, 2012), sehingga kata yang sebelumnya tidak terstruktur akan dapat terindeks dengan baik dan menjadi terstruktur.

Tahapan berikutnya adalah *term weighting*, yaitu pemberian bobot pada tiap-tiap *term* dilakukan untuk memberikan informasi tambahan dalam pencarian dokumen relevan (Singh and Dwivedi, 2012). Pembobotan dasar dilakukan dengan menghitung frekuensi kemunculan *term* dalam dokumen (*tf*) dan bentuk *log normalization (wf)*, lalu diikuti perhitungan *idf* dan *wf idf*.

Rumus perhitungan *term weighting* yang meliputi *wf*, *idf* dan *wf idf* (Manning, Raghavan and Schütze, 2009) adalah sebagai berikut.

1. Perhitungan *Wf*

$$w_{f,t,d} = 1 + \log_{10}(t_{f,t,d}), \text{ if } t_{f,t,d} > 0 \quad (1)$$

$$0, \text{ otherwise}$$

Keterangan:

$w_{f,t,d}$ = *term frequency* dengan pendekatan *log normalization*.

$t_{f,t,d}$ = frekuensi kemunculan kata t pada dokumen d

t = kata atau *term*

d = dokumen

2. Perhitungan *Inverse Document Frequency* (IDF)

$$idf_t = \log \frac{N}{df_t} \quad (2)$$

Keterangan:

idf_t = banyaknya dokumen yang mengandung *term* atau kata t

N = menyatakan banyaknya dokumen yang ada

df_t = banyak dokumen yang memuat kata t

3. Perhitungan *Wf Idf Weighting*

$$wf-idf_{t,d} = w_{f,t,d} \times idf_t \quad (3)$$

Keterangan:

$wf-idf_{t,d}$ = bobot *term* atau kata t pada dokumen ke d

$w_{f,t,d}$ = *term frequency* dengan pendekatan *log normalization*

idf_t = *inverse document frequency*

Setelah tahapan *term weighting* dilakukan, maka tahapan selanjutnya adalah perhitungan *cosine similarity*, tahapan ini digunakan untuk mendapatkan nilai kedekatan antar dokumen atau kemiripan dokumen yang direpresentasikan dalam ruang v -dimensional vektor, dimana v merupakan ukuran atau dimensi vektor (Singh and Dwivedi, 2012).

Rumus perhitungan nilai *similarity* dapat dinyatakan dalam persamaan berikut, CosSim (d_j, q):

$$\frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^t (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^n w_{ij}^2 \cdot \sum_{i=1}^n w_{iq}^2}} \quad (4)$$

Keterangan:

\vec{d}_j = Vektor dokumen ke j

\vec{q} = Vektor *query* atau dokumen lain (uji)

w_{ij} = Bobot *term* ke i dari koleksi dokumen ke j / $wf-idf_{t,d}$

w_{iq} = Bobot *term* ke i dari *query* atau dokumen uji / $wf-idf_{t,q}$

2.6. *K-Nearest Neighbor* (K-NN)

K-Nearest Neighbor (K-NN) merupakan metode pengelompokan yang mengelompokkan data baru berdasarkan jarak data baru itu ke beberapa data/tetangga (*neighbor*) terdekat (Santosa, 2007). Untuk mengimplementasikan metode *K-Nearest Neighbor* (K-NN) dapat dengan melakukan langkah-langkah berikut ini:

1. Mulai *input* data *training*, label data *training*, k , data testing.
2. Untuk semua data testing, hitung kemiripannya ke setiap data *training*.
3. Tentukan k data *training* yang kemiripannya paling dekat dengan data testing.
5. Periksa label dari k data ini.
6. Tentukan label yang frekuensinya paling banyak.
7. Masukkan data testing ke kelas dengan frekuensi paling banyak.
8. Berhenti.

3. DATA PENELITIAN

Data yang digunakan pada penelitian ini adalah data berupa *tweet*. Sumber *datasets* yang digunakan dalam penelitian ini berasal dari situs www.twitter.com, *datasets* terdiri dari data latih dan data uji. Sebanyak 200 *tweet* diambil yang memuat kata kunci ‘kurikulum2013’. Dari 200 data tersebut, 100 data merupakan *tweet* yang positif dan 100 lagi merupakan *tweet* negatif. Penentuan kelas positif dan negatif dianotasi secara manual dengan bantuan pakar, pakar adalah guru Bahasa Indonesia. Sebanyak 150 data *tweet* dijadikan sebagai data latih (75 data latih kelas positif, 75 data latih kelas negatif) dan 50 sebagai data uji (25 positif dan 25 negatif).

Selain data tersebut juga terdapat data lainnya yang digunakan dalam penelitian ini, yaitu pada tahapan *preprocessing*, terdapat data *stopword list* berbahasa Indonesia dari Fadillah Z. Tala yang digunakan pada saat proses *filtering* dan data kamus kata dasar berbahasa Indonesia dari *library* Sastrawi yang digunakan pada proses *stemming*.

Data lainnya yang juga digunakan pada tahapan untuk membantu pengambilan nilai fitur yaitu, data kamus atau *lexicon* kata yang memiliki *tagging* atau *Part of Speech* (PoS) berbahasa Indonesia (29.081 kata), data *list emoticon*, data kamus kata penguat atau

intensifier word (38 kata), data kamus kata positif (1.183 kata), kamus kata negatif (2.402 kata), data kamus bahasa baku dan tidak baku. Data tersebut dikumpulkan dengan cara *crawling* satu per satu dan beberapa data berasal dari penelitian terdahulu seperti untuk data *Part of Speech* (PoS) berbahasa Indonesia dan kamus kata positif dan negatif yang di bagikan di situs www.github.com.

4. PERANCANGAN SISTEM

Bagian ini menjelaskan langkah-langkah yang ditempuh dalam pembuatan analisis sentimen dengan tujuan mendapatkan label *tweet* apakah termasuk label kelas positif atau negatif, alur kerja utama sistem dapat dilihat pada Gambar 2. Seperti terlihat pada Gambar 2, proses pertama yang dilakukan adalah sistem akan mengambil *tweet* yang dimasukkan oleh pengguna sebagai *tweet* uji, setelah itu *tweet* tersebut menuju proses pembakuan kata (*standardization of words*), setelah melakukan pembakuan kata, *tweet* akan menuju proses pengambilan nilai pada *ensemble feature*, pada proses ini terdapat bagian *twitter specific features*, *textual features*, *Parts of Speech* (PoS) *features* dan *lexicon based features*.

Pada bagian *twitter specific features*, bagian ini akan mengambil nilai fitur berdasarkan spesifikasi tertentu yang dimiliki oleh *tweet*, seperti apakah *tweet* tersebut memiliki *hashtag* (#), *retweet* (RT), *username* (@) dan URL.

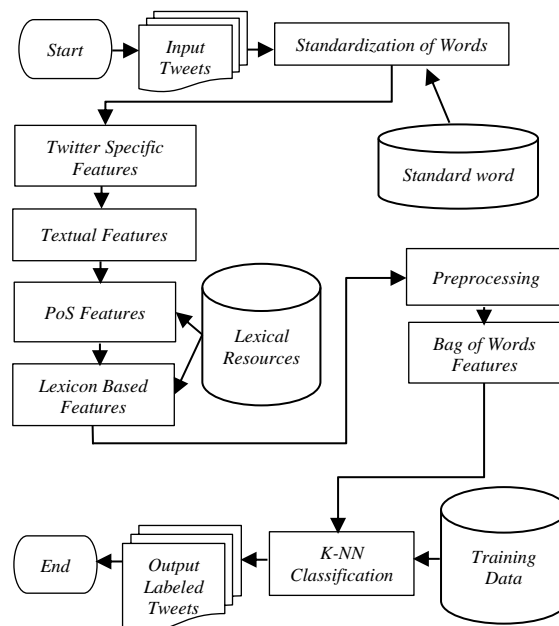
Selanjutnya bagian *textual features*, yaitu pengambilan nilai fitur berdasarkan informasi eksplisit yang terdapat dalam suatu *tweet*, seperti jumlah kata dalam *tweet*, jumlah tanda tanya, seru dan sebagainya.

Setelah itu masuk bagian *Parts of Speech* (PoS) *features* dan *lexicon based features* yang melakukan pengambilan fitur menggunakan bantuan *lexicon* atau kamus.

Pada bagian fitur *Bag of Words* (BoW), terdapat proses khusus sebelum dilakukan pengambilan nilai fitur, yaitu *preprocessing* yang meliputi *tokenization*, *filtering*, *stemming* dapat dilihat pada Gambar 1 diagram alir *preprocessing*.

Setelah seluruh proses pengambilan nilai fitur dilakukan, maka akan dilanjutkan menuju proses klasifikasi dengan metode *K-Nearest Neighbor* (K-NN). Tahapan ini melakukan perhitungan atau testing dengan data *training* yang telah diproses lebih dulu, sehingga dari hasil klasifikasi dapat ditentukan label kelas

tweet, yaitu positif atau negatif.



Gambar 2. Alur kerja utama sistem

5. PENGUJIAN DAN ANALISIS

Pada bagian ini dilakukan pengujian dan analisis yang bertujuan untuk mengetahui akurasi dari metode *K-Nearest Neighbor* (K-NN), terdapat 4 macam skenario pengujian yang dilakukan, pengujian pertama merupakan pengujian dari nilai *k* terhadap akurasi metode *K-Nearest Neighbor* (K-NN), lalu pengujian berikutnya adalah merupakan pengujian dari kombinasi fitur meliputi pengujian hanya fitur *Bag of Words* (BoW), lalu pengujian fitur *ensemble* tanpa BoW (*twitter specific features*, *textual features*, *PoS features*, *lexicon based features*), terakhir pengujian fitur *ensemble* secara lengkap, yaitu *Bag of Words* (BoW) dan *ensemble* (*twitter specific features*, *textual features*, *PoS features*, *lexicon based features*).

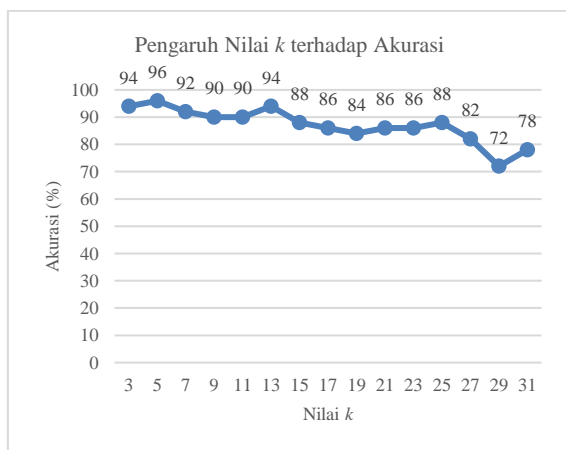
6.1 Hasil dan Analisis Pengujian Nilai *k* terhadap Akurasi Metode *K-Nearest Neighbor* (K-NN)

Pengujian berdasarkan nilai *k* pada analisis sentimen digunakan untuk mengetahui banyaknya *k* yang memiliki nilai akurasi terbaik. Pada pengujian ini, mode fitur yang digunakan adalah fitur *ensemble* secara lengkap. Data yang digunakan adalah sebanyak 200 *datasets*, yaitu 150 data merupakan data *training* dan 50 data merupakan data *testing*. Pengujian dilakukan secara bergantian dengan mengganti nilai *k*, nilai *k* yang diambil hanya nilai dari bilangan ganjil saja agar dapat dilakukan pengambilan

keputusan kelas positif atau negatif. Banyaknya nilai k yang diuji coba adalah dimulai dari 3 hingga 31, sehingga terdapat variasi pengujian nilai k sebanyak 15 kali.

Dapat dilihat pada Gambar 3 grafik pengujian pengaruh nilai k terhadap akurasi menunjukkan bahwa, terjadi fluktuasi nilai akurasi. Ketika nilai k terlalu sedikit, yaitu nilai $k=3$ akurasi belum mencapai titik maksimal dikarenakan ada beberapa data terdekat yang sangat relevan tidak dilibatkan dalam perhitungan *voting* kelas pada data tersebut, sebaliknya ketika nilai k meningkat, yaitu lebih dari 13 maka akurasi mengalami penurunan secara perlahan dikarenakan terdapat data dengan tingkat kemiripan yang jauh atau data yang tidak relevan akan dilibatkan dalam perhitungan *voting* kelas pada data tersebut.

Nilai akurasi terbaik didapatkan ketika k bernilai 5. Akurasi yang didapatkan sebesar 96%, yaitu dari 50 data uji, terdapat 48 *tweet* uji yang diklasifikasikan secara benar oleh metode *K-Nearest Neighbor* (K-NN) dan sisanya 2 *tweet* uji saja yang salah. Oleh karena itu, nilai k yang akan diambil adalah sebanyak 5 untuk kemudian digunakan pada pengujian berikutnya.



Gambar 3. Pengaruh nilai k terhadap akurasi

6.2 Hasil dan Analisis Pengujian Fitur *Bag of Words* (BoW)

Pengujian berdasarkan fitur *Bag of Words* (BoW) digunakan untuk mengetahui seberapa besar pengaruh fitur ini terhadap akurasi metode *K-Nearest Neighbor* (K-NN) dari sisi statistik kata. Pengujian dilakukan menggunakan nilai k terbaik, yaitu bernilai 5. Dalam pengujian ini, nilai k statis dan akan digunakan pada setiap data uji yang berjumlah 50 *tweet* uji. Pengujian dilakukan satu persatu untuk mengetahui klasifikasi kelas positif dan negatif yang bernilai

benar dan salah.

Dari 50 data *tweet* uji yang diujikan, terdapat 40 data uji yang diklasifikasikan dengan benar oleh metode *K-Nearest Neighbor* (K-NN), sisanya terdapat 10 data *tweet* uji yang salah. Setelah dilakukan analisis, dari 10 data *tweet* uji yang diklasifikasikan dengan salah tersebut, ada beberapa *tweet* yang pendek, yaitu hanya memiliki beberapa kata saja dalam sebuah *tweet*. Hal ini menyebabkan pengambilan nilai fitur pada data uji hanya bergantung pada sedikit kata tersebut, akibatnya ada beberapa kata yang terdapat pada data *tweet* uji namun kata tersebut tidak ada pada *tweet* data latih sehingga data uji tersebut akan sulit dikenali yang berdampak pada hasil pengklasifikasian. Ini menunjukkan penggunaan fitur ini sangat bergantung pada statistik kata yang terdapat pada data latih.

Oleh karena itu, penggunaan fitur *Bag of Words* (BoW) belum dapat bekerja secara maksimal untuk kasus diatas dan hasil akurasi yang didapatkan sebesar 80%.

6.3 Hasil dan Analisis Pengujian Fitur *Ensemble Tanpa BoW (Twitter Specific Features, Textual Features, POS Features, Lexicon Based Features)*

Pengujian berdasarkan fitur *ensemble (twitter specific features, textual features, PoS features, lexicon based features)* digunakan untuk mengetahui seberapa besar pengaruh fitur ini terhadap akurasi metode *K-Nearest Neighbor* (K-NN) dari sisi semantik kata. Pengujian ini juga menggunakan k bernilai 5. Dalam pengujian ini, juga dilakukan pengujian satu persatu terhadap 50 data *tweet* uji dan nilai k juga statis untuk setiap data uji.

Dari 50 data *tweet* uji yang diujikan, terdapat 41 data uji yang diklasifikasikan dengan benar oleh metode *K-Nearest Neighbor* (K-NN), sisanya terdapat 9 data uji *tweet* yang salah, hal ini wajar dikarenakan fitur ini sangatlah bergantung pada kamus atau *lexicon* yang digunakan. Kata pada data *tweet* uji yang belum dapat dikenali dengan baik atau yang tidak terdapat didalam *lexicon* akan mempengaruhi nilai fitur sehingga berdampak pada hasil pengklasifikasian. Untuk penggunaan fitur ini sudah sedikit lebih baik dari penggunaan yang hanya melibatkan *Bag of Words* (BoW) saja dan fitur ini tidak memiliki ketergantungan pada satatistik kata saja, namun terdapat fitur lain serta *lexicon* yang membantu dalam mengenali sebuah *tweet* uji.

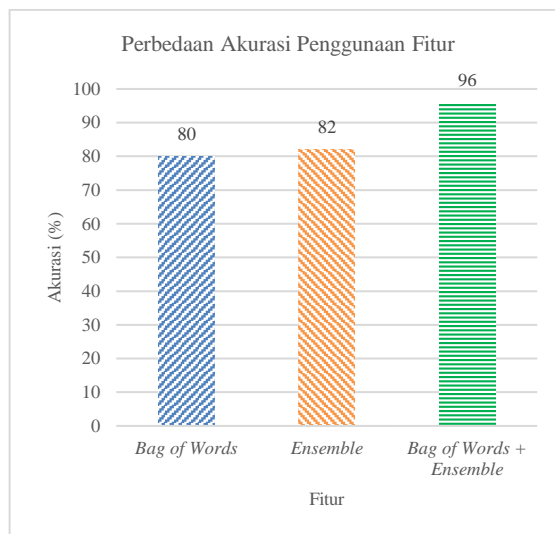
Berdasarkan grafik pada Gambar 4 akurasi perbedaan fitur, terlihat bahwa akurasi yang didapatkan sebesar 82 % untuk penggunaan fitur *ensemble* (*twitter specific features, textual features, PoS features, lexicon based features*), terdapat peningkatan sebesar 2 % dari penggunaan fitur *Bag of Words* (BoW) pada pengujian sebelumnya.

6.4 Hasil dan Analisis Pengujian Fitur *Bag of Words* (BoW) dan *Ensemble* (*Twitter Specific Features, Textual Features, PoS Features, Lexicon Based Features*)

Pengujian berdasarkan fitur *Bag of Words* (BoW) dan *ensemble* (*twitter specific features, textual features, PoS features, lexicon based features*) digunakan untuk mengetahui seberapa besar pengaruh fitur *ensemble* secara lengkap terhadap akurasi metode *K-Nearest Neighbor* (K-NN), yaitu dari menggabungkan sisi statistik kata dengan semantik kata. Dalam pengujian ini, juga dilakukan pengujian satu persatu terhadap 50 data *tweet* uji dan nilai $k=5$ yang statis untuk setiap data uji.

Dari 50 data *tweet* uji yang diujikan, terdapat 48 data uji yang diklasifikasikan dengan benar oleh metode *K-Nearest Neighbor* (K-NN), sisanya terdapat 2 data uji *tweet* yang salah, sehingga didapatkan akurasi sebesar 96% untuk penggunaan fitur gabungan ini, terdapat peningkatan sebesar 14% dari pengujian sebelumnya. Ini menunjukkan penggabungan fitur memiliki dampak yang signifikan dalam peningkatan akurasi, kelemahan yang ada pada penggunaan fitur hanya secara *Bag of Words* (BoW) dapat dilengkapi oleh *ensemble* begitu juga sebaliknya.

Berdasarkan grafik pada Gambar 4, terlihat bahwa akurasi tertinggi didapatkan dengan menggunakan fitur *ensemble* secara lengkap, yaitu gabungan *Bag of Words* (BoW) dan *ensemble* (*twitter specific features, textual features, PoS features, lexicon based features*). Dengan menggabungkan *Bag of Words* (BoW) dan *ensemble* tersebut, akan didapatkan hasil yang lebih baik yaitu, mencapai akurasi sebesar 96%. Berbeda hal jika hanya menggunakan fitur secara independen saja, akurasi yang didapatkan jika menggunakan fitur secara independen hanya mencapai 80% pada fitur *Bag of Words* (BoW) dan 82% pada fitur *ensemble* (*twitter specific features, textual features, PoS features, lexicon based features*).



Gambar 4. Perbedaan akurasi penggunaan fitur

6. KESIMPULAN

Pada penelitian ini, telah dilakukan berbagai skenario pengujian untuk mengetahui pengaruh nilai k dan pengaruh kombinasi fitur terhadap akurasi metode K-NN dalam analisis sentimen kurikulum 2013.

Nilai k sangat berpengaruh penting dalam akurasi metode K-NN, nilai k terbaik yang didapatkan setelah dilakukan pengujian adalah 5 data terdekat. Nilai k tersebut didapatkan dengan menguji nilai k secara berganti-ganti dan dilakukan pengambilan nilai k dengan akurasi tertinggi yaitu, sebesar 96%. Nilai k yang terlalu sedikit menyebabkan akurasi yang didapatkan belum mencapai titik maksimal sebaliknya nilai k yang terlalu banyak akan menyebabkan akurasi menurun.

Selain dari nilai k , kombinasi fitur juga memiliki pengaruh penting yang signifikan dalam meningkatkan akurasi dari metode K-NN, penggabungan fitur *Bag of Words* (BoW) dan *ensemble* (*twitter specific features, textual features, PoS features, lexicon based features*) dapat meningkatkan akurasi metode K-NN jika dibandingkan dengan menggunakan fitur secara independen saja. Penggabungan fitur ini dapat melengkapi kelemahan masing-masing fitur, akurasi yang didapatkan dengan menggabungkan kedua fitur tersebut mencapai 96%.

7. DAFTAR PUSTAKA

Bijalwan, V., Kumar, V., Kumari, P. and Pascual, J., 2014. KNN based Machine Learning Approach for Text and Document Mining. *International Journal of Database Theory and*

Application, [online] 7(1), pp.61–70. Available at:

<<http://dx.doi.org/10.14257/ijdta.2014.7.1.06>>.

Feldman, R. and Sanger, J., 2007. Introduction to text mining. *The text mining handbook: Advanced approaches to analyzing unstructured data*, pp.1–10.

Hardiyanto, E. and Rahutomo, F., 2016. Studi Awal Klasifikasi Artikel Wikipedia Bahasa Indonesia Dengan Menggunakan Metoda K Nearest Neighbor. *PROSIDING*, 1, pp.158–165.

Manning, C.D., Raghavan, P. and Schütze, H., 2009. *An Introduction to Information Retrieval*. Cambridge University Press, England, .

Nurdiana, Y.I., 2015. Studi Komparasi Implementasi Kurikulum 2013 Di Kelas X antara Sekolah Menengah Atas Negeri 1 Taman Sidoarjo dengan Madrasah Aliyah Negeri Sidoarjo.

Poerwati, L.E. and Amir, S., 2013. Panduan Memahami Kurikulum 2013. *Jakarta: PT. Prestasi Pustakaraya*, p.28.

Santosa, B., 2007. Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis. *Yogyakarta: Graha Ilmu*.

Siddiqua, U.A., Ahsan, T. and Chy, A.N., 2016. Combining a Rule-based Classifier with Ensemble of Feature Sets and Machine Learning Techniques for Sentiment Analysis on Microblog. pp.16–21.

Da Silva, N.F.F., Hruschka, E.R. and Hruschka, E.R., 2014. Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, [online] 66, pp.170–179. Available at: <<http://dx.doi.org/10.1016/j.dss.2014.07.003>>.

Singh, J. and Dwivedi, S., 2012. Analysis of Vector Space Model in Information Retrieval. *IJCA Proceedings on National Conference on Communication Technologies & its impact on Next Generation Computing 2012*, [online] pp.14–18. Available at: <<http://ijcaonline.net/proceedings/ctngc/number/2/9056-1016>>.