

**ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE**

Ghulam Asrofi Buntoro

Program Studi Informatika Fakultas Teknik
Universitas Muhammadiyah Ponorogo
Jl. Budi Utomo 10 Ponorogo 63471 Jawa Timur Indonesia
ghulamasrofibuntoro@gmail.com

ABSTRACT

Abstract— Today, social media, especially Twitter have enormous influence to the success or ruin the image of a person. Many movements are carried out in social media, especially Twitter, all of which can influence its success. There is a movement that aims good there is also a movement with malicious purposes, namely hatred to others. Usually the movement on Twitter was done using the hashtag (#), the latest movement there tagar Hatespeech (#HateSpeech), viewed from the name is already clear that hate speech. This study analyzes the hashtag proficiency level, all justified by the hashtag was the sentiment of hate. The classification process in this study using the method of classification Naive Bayes classifier (NBC) and Support Vector Machine (SVM) with the data preprocessing using tokenisasi, cleansing and filtering. The data used are in Indonesian tweet with the hashtag HateSpeech (#HateSpeech), with the number of datasets as much as 522 tweets were distributed evenly into two sentiments HateSpeech and GoodSpeech. The highest accuracy of results obtained when using the method of classification Support Vector Machine (SVM) with tokenisasi unigram, stopword list Indonesian and emoticons, with the average value reached 66.6% accuracy, precision value of 67.1%, 66.7% recall value TP value rate of 66.7% and 75.8% rate the value TN.

Keywords—Sentiment Analysis; HateSpeech; Twitter; Data Mining; Naive Bayes classifier (NBC); Support Vector Machine (SVM);

ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN *SUPPORT VECTOR MACHINE*
(Ghulam Asrofi Buntoro)

INTISARI

Sekarang ini, media sosial khususnya Twitter mempunyai pengaruh yang sangat besar bagi kesuksesan ataupun kehancuran citra seseorang. Banyak gerakan-gerakan yang dilakukan dalam media sosial khususnya Twitter yang semuanya sukses mampu mempengaruhi penggunanya. Ada gerakan yang bertujuan baik ada juga gerakan dengan tujuan jahat yaitu menebar kebencian kepada orang lain. Biasanya gerakan di Twitter itu dilakukan dengan menggunakan tagar (#), Gerakan terbaru ada tagar Hatespeech (#HateSpeech), dilihat dari namanya sudah jelas yaitu ucapan kebencian. Penelitian ini menganalisis tagar tersebut, benarkan semua yang diberi tagar itu sentimennya kebencian. Proses klasifikasi pada penelitian ini menggunakan metode klasifikasi *Naive Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM) dengan *preprocessing* data menggunakan tokenisasi, *cleansing* dan *filtering*. Data yang digunakan adalah *tweet* dalam bahasa Indonesia dengan tagar HateSpeech (#HateSpeech), dengan jumlah *dataset* sebanyak 522 *tweet* yang didistribusikan secara merata menjadi dua sentimen HateSpeech dan GoodSpeech. Hasil akurasi tertinggi didapatkan saat menggunakan metode klasifikasi *Support Vector Machine* (SVM) dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*, dengan nilai rata-rata akurasi mencapai 66,6%, nilai presisi 67,1%, nilai *recall* 66,7% nilai *TP rate* 66,7% dan nilai *TN rate* 75,8%.

Kata kunci – Analisis Sentimen; Sentiment Analysis; HateSpeech; Twitter; Data Mining; Naive Bayes Classifier (NBC); Support Vector Machine (SVM);.

PENDAHULUAN

Hate Speech adalah perkataan benci atau ucapan kebencian. Ucapan kebencian pada zaman dahulu diucapkan langsung kepada orang yang dibenci, tapi seiring dengan perkembangan zaman dan Teknologi Informasi, perkataan itu bisa diucapkan dengan banyak media[1].

Perkembangan media yang begitu pesat, memunculkan banyak media online dari media berita sampai media sosial. Media sosial saja sudah begitu banyak, dari Facebook, Twitter, Path, Instagram, Google+, Tumblr, LinkedIn dan sebagainya masih banyak lagi[2]. Media sosial sekarang ini tidak hanya digunakan sebagai sarana pertemanan, mencari teman, tapi sudah banyak digunakan untuk kegiatan lain. Promo dagangan, jual beli apa saja sampai promo partai politik atau kampanye calon calon legislative dan presiden,

Media sosial khususnya *Twitter* sekarang ini menjadi perangkat komunikasi yang sangat populer di kalangan pengguna internet. Pada konferensi resmi pengembang *Twitter* Chirp 2010, perusahaan tersebut menyampaikan statistik mengenai situs dan pengguna *Twitter*. Statistik tersebut menyebutkan bahwa pada bulan April 2010, *Twitter* memiliki 106 juta akun dan sebanyak 180 juta pengunjung unik setiap bulannya. Jumlah pengguna *Twitter* disebutkan terus meningkat 300.000 *user* setiap harinya[3]. *Digital Buzz blog* suatu situs yang menyediakan statistik infografik menyebutkan data statistik yang sama.

Tidak salah lagi bila sekarang ini media sosial *Twitter* menjadi tempat promo atau kampanye yang efektif dan efisien. Tim sukses sebuah pasangan calon kepala daerah atau presiden sekarang ini bisa menghalalkan segala cara dalam kampanye calonnya. Dahulu ada istilah *Black Campaign*, untuk sekarang ini setelah lewat masa kampanye, ada istilah baru yaitu *Hate Speech*.

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Pada penelitian ini Analisis sentimen dilakukan untuk melihat sebuah opini seseorang itu apakah termasuk *HateSpeech* atau tidak. *HateSpeech* bisa dimasukkan kategori opini negatif. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian dan aplikasi berbasis analisis sentimen berkembang pesat. Bahkan di Amerika terdapat sekitar 20-30 perusahaan yang memfokuskan pada layanan analisis sentimen [4].

PENELITIAN TERKAIT

Penelitian oleh Mesut *et al*[2] menggunakan *machine learning* untuk mengklasifikasikan berita politik Turki. Penelitian ini melakukan klasifikasi sentimen terhadap berita politik Turki dan menentukan apakah berita politik Turki tersebut memiliki sentimen positif ataukah negatif. Fitur yang berbeda dari berita politik Turki tersebut diekstrak dengan algoritme *machine learning Naïve Bayes Classifier* (NBC), *Maximum Entropy* (ME) dan *Support Vector Machine* (SVM) untuk menghasilkan model klasifikasi. Penelitian ini memperoleh Akurasi 72,05% untuk *Naïve Bayes Classifier* (NBC), Akurasi 69,44% *Maximum Entropy* dan 66,81% untuk SVM pada penggunaan *bigram*.

Pak dan Paurobek[3] menggunakan *emoticon* untuk membangun *corpus* berbahasa Inggris dari Twitter dengan sentimen positif, negatif dan netral. Untuk kelas netral Pak dan Paurobek mengambil data *training* dari *tweet* akun media berbahasa Inggris. Metode yang digunakan adalah *Naïve Bayes* dengan *n-gram*. Performasi terbaik dihasilkan ketika menggunakan *bigram*.

Read[5] menggunakan *emoticons* seperti “:-)” dan “:- (“ untuk membentuk suatu *training* set klasifikasi sentimen. Untuk tujuan ini Read mengumpulkan teks yang berisi *emoticons* dari *Usenet Newsgroup*. Dataset dibagi atas sampel positif (teks dengan *happy emoticon*) dan sampel negatif (teks dengan *sad* atau *angry emotions*). Read[5] melakukan percobaan klasifikasi berdasarkan topik *dependency*. Untuk kategori *mixed* topik pada data *training* dan *mixed* topik pada data testing, Read menghasilkan akurasi sebesar 84,6% untuk metode *Naïve Bayes* dan 81,1% untuk metode SVM. Untuk klasifikasi topik yang lain read memperoleh akurasi berkisar pada 62% hingga 70% pada test set untuk kedua metode. Read juga memperoleh akurasi 81,5% untuk metode SVM dan akurasi 78,9% untuk metode *Naïve Bayes*, ketika menggunakan *domain dependency* pada klasifikasi sentimen.

ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE
(Ghulam Asrofi Buntoro)

Penelitian oleh Pang et al[6] menggunakan *machine learning* untuk mengklasifikasikan *movie reviews*. Penelitian ini melakukan klasifikasi sentimen terhadap *review* film dan menentukan apakah *review* film tersebut memiliki sentimen positif ataukah negatif. Fitur yang berbeda dari *review* tersebut diekstrak dan digunakan algoritme *machine learning Naïve Bayes* dan *Support Vector Machine (SVM)* untuk menghasilkan model klasifikasi. Mereka memperoleh akurasi antara 78,7% ketika menggunakan *Naïve Bayes* pada penggunaan *unigram*. Akurasi yang diperoleh ketika menggunakan SVM dengan *unigram* adalah 72,8%.

Frangky dan Manurung[7] mencoba untuk mengulangi eksperimen klasifikasi sentiment *movie review* oleh Pang et al[6] untuk Bahasa Indonesia. Berkaitan dengan ketidaktersediaannya *training corpora* untuk Bahasa Indonesia, maka diaplikasikan *machine translation tools* untuk mentranslasikan *corpus* Bahasa Inggris yang dibuat Pang et al[6] yang asli ke Bahasa Indonesia dan hasil translasinya digunakan untuk *train* klasifikasi. Berbagai pilihan *machine translation* digunakan mulai dari *commercial tool* hingga translasi sederhana kata demi kata dan metode klasifikasi teks dicoba. Hasil akurasi rata-rata yang diperoleh untuk metode *Naïve Bayes* adalah 74,6% dan 75,62% untuk metode SVM. Hasil terbaik diperoleh sama dengan yang diperoleh ketika menggunakan eksperimen dalam Bahasa Inggris.

METODOLOGI

Langkah-langkah penelitian sesuai dengan alur penelitian adalah sebagai berikut:

1. Mengumpulkan data tweet

Data *tweet* diambil mulai bulan September - November 2015 dari media sosial *Twitter*. Data yang diambil hanya *tweet* dalam bahasa Indonesia, yaitu *tweet* dengan Tagar *Hate Speech (#HateSpeech)*. Data diambil secara acak baik dari user biasa ataupun media *online* di *Twitter*.

2. Mengubah data tweet ke dalam format ARFF

Data *tweet* yang dikumpulkan yang berbentuk teks, kemudian dijadikan file ARFF[8]. Untuk membuat file ARFF dengan cara manual.

3. Preprocessing Data

Melakukan *preprocessing* data *tweet*. *Preprocessing* meliputi tokenisasi *cleansing* dan *filtering*. Tokenisasi dilakukan untuk memecah *tweet* menjadi beberapa kata atau kumpulan kata yang berdiri sendiri. Penelitian ini menggunakan 3 metode tokenisasi, yaitu *unigram*, *bigram*, dan *n-gram* dengan nilai *n* minimum = 1 dan *n* maksimum = 3. Proses tokenisasi menggunakan menu yang ada dalam WEKA. Pada *tokenizer* pilih *choose* kemudian pilih metode tokenisasi yang akan digunakan.

Kemudian masuk proses *cleansing*, yaitu proses menghapus simbol-simbol yang kurang penting dalam data *tweet* yang bisa mengganggu proses klasifikasi nantinya. Proses ini dilakukan dengan menggunakan menu *delimiters* pada WEKA.

Tahapan akhir setelah tokenisasi dalam penelitian ini adalah *filtering*. *Filtering* dilakukan untuk menghapus kata-kata yang kurang penting atau kurang berpengaruh terhadap proses klasifikasi nantinya. Proses ini dilakukan dengan menggunakan *stopword list*. *Stopword list* yang digunakan dalam penelitian ini adalah *stopword list* WEKA dan *stopword list* Bahasa Indonesia yang dibuat oleh Tala[9].

4. Data tweet diubah ke dalam bentuk vektor

Data *tweet* kemudian diubah ke dalam bentuk vektor[10]. Dengan cara memilih *StringToWordVector* pada tool WEKA. Hasilnya dapat dilihat pada Gambar 1.

relation: Data_Capres-weka_filters_unsupervised_attribute_StringToWordVector-R1-W1000-prune-rate-1.0-T-1-N0-stemmerweka_core_st

38: Capres Numeric	39: Cina Numeric	40: Daftar Numeric	41: Demi Numeric	42: Dengan Numeric	43: Dia Numeric	44: Diantaranya Numeric	45: Diseret Numeric	46: Doi Numeric	47: Dosa Numeric	48: Fakta Numeric
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3.052966	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.234608
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Gambar 1. Contoh vektor dari data *tweet*

Pada Gambar 1. bagian yang ada dalam kotak merah merupakan kata-kata yang ada dalam data *tweet*. Untuk tiap baris mewakili tiap-tiap data *tweet*. Pada baris ke-3 dalam kotak berwarna biru dapat dilihat bahwa kata “Capres” memiliki nilai 3,052966 dan kata “fakta” memiliki nilai 4,234608. Sedangkan yang lainnya bernilai 0,0 itu artinya ada kata “Capres” dan “fakta” pada data *tweet* ke-3.

5. Pemberian Bobot

Memberikan bobot pada tiap-tiap kata (*term*). Pembobotan dilakukan untuk mendapatkan nilai dari kata yang berhasil diekstrak. Metode yang digunakan untuk pemberian bobot dalam penelitian ini adalah TF-IDF (*Term Frequency – Inverse Document Frequency*).

6. Klasifikasi

Masuk pada proses klasifikasi. Proses klasifikasi menggunakan WEKA 3.7.11. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)*. Pada proses klasifikasi data diuji dengan menggunakan metode *10 fold cross validation*[11]. Jadi dataset akan dibagi menjadi dua, yaitu 10 bagian dengan 9/10 bagian digunakan untuk proses *training* dan 1/10 bagian digunakan untuk proses *testing*. *Iterasi* berlangsung 10 kali dengan variasi data *training* dan *testing* menggunakan kombinasi 10 bagian data.

7. Evaluasi Hasil

Melakukan evaluasi performa Akurasi, Presisi dan *Recall* dari eksperimen yang telah dilakukan. Evaluasi dilakukan dengan menggunakan *true positive rate (TP rate)*, *true negative rate (TN rate)*, *false positive rate (FP rate)* dan *false negative rate (FN rate)* sebagai indikator. *TP rate* adalah persentase dari kelas positif yang berhasil diklasifikasi sebagai kelas positif, sedangkan *TN rate* adalah persentase dari kelas negatif yang berhasil diklasifikasi sebagai kelas negatif. *FP rate* adalah kelas negatif yang diklasifikasi sebagai kelas positif. *FN rate* adalah kelas positif yang diklasifikasi sebagai kelas negatif[12].

UJI COBA DAN PEMBAHASAN

Dataset pada penelitian ini menggunakan format ARFF yang dikumpulkan dari Twitter. Data yang diambil hanya *tweet* dalam bahasa Indonesia, yaitu *tweet* dengan Tagar *Hate Speech (#HateSpeech)*. Data diambil secara acak baik dari user biasa ataupun media *online* di *Twitter*.

Dataset yang digunakan sebanyak 522 *Tweet*, data dibagi secara seimbang (*balanced*) setiap kelasnya, karena dengan data yang tidak seimbang (*imbalanced*), klasifikasi yang dibangun memiliki kecenderungan untuk mengabaikan *minority class*[25]. Data dibagi menjadi *HateSpeech* 257 *Tweet* dan *GoodSpeech* 265 *Tweet*. Pemberian label dilakukan secara manual dengan bantuan ahli Bahasa Indonesia.



Gambar 2. Data *tweet*

ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN *SUPPORT VECTOR MACHINE*
 (Ghulam Asrofi Buntoro)

Pada Gambar 2. bagian yang berwarna merah menunjukkan Tweet HateSpeech dan yang berwarna hijau menunjukkan Tweet GoodSpeech. Dari gambar diatas menunjukkan bahwa dataset yang digunakan dalam penilitan ini seimbang (*balanced dataset*). Hasil akurasi yang diperoleh dari penelitian yang menggunakan *dataset* yang seimbang lebih baik dari yang menggunakan dataset yang tidak seimbang[13]. Pada penelitian ini menggunakan data yang seimbang untuk klasifikasi menggunakan metode *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)*. Tujuan utamanya adalah ingin membuktikan apakah hasilnya sudah cukup baik apa belum.

Pada penelitian ini metode yang digunakan untuk melakukan klasifikasi data adalah *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* dengan menggunakan perangkat lunak WEKA versi 3.7.11 untuk melakukan klasifikasi. WEKA menggunakan tipe dokumen *Atribut-Relation File Format (ARFF)* sebagai masukan untuk melakukan klasifikasi data.

Hasil dari proses klasifikasi kemudian diuji dengan menggunakan metode *10 fold cross validation*, data dibagi mejadi 10 bagian dengan 9/10 bagian digunakan untuk proses *training* dan 1/10 bagian digunakan untuk proses *testing*. Iterasi berlangsung 10 kali dengan variasi data *training* dan *testing* menggunakan kombinasi 10 bagian data.

Pengujian	Dataset									
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

Gambar 3. Ilustrasi *10 fold cross validation*

Bagian yang berwarna hitam pada Gambar 4.2 menunjukkan data uji. Pada gambar di atas *dataset* dibagi menjadi 10 *subset*. Pada pengujian pertama, *subset* yang diuji adalah *subset* yang pertama, sedangkan *subset* kedua hingga kesepuluh digunakan sebagai data *training*. Pada pengujian kedua, *subset* yang diuji adalah *subset* yang kedua, sedangkan *subset* pertama, ketiga sampai kesepuluh digunakan sebagai data *training* dan begitu seterusnya.

Dari hasil klasifikasi dengan menggunakan metode klasifikasi *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* tokenisasi *unigram*, *bigram* dan *n-gram*, *stopword list* WEKA dan *stopword list* Bahasa Indonesia) dengan menggunakan *emoticons* dan menghapus *emoticons*, *cleansing* dan *filtering* yang berbeda, maka dapat diperoleh tabel perbandingan Hasil Klasifikasi *Naïve Bayes Classifier (NBC)* dengan Hasil Klasifikasi *Support Vector Machine (SVM)* seperti Tabel 1.

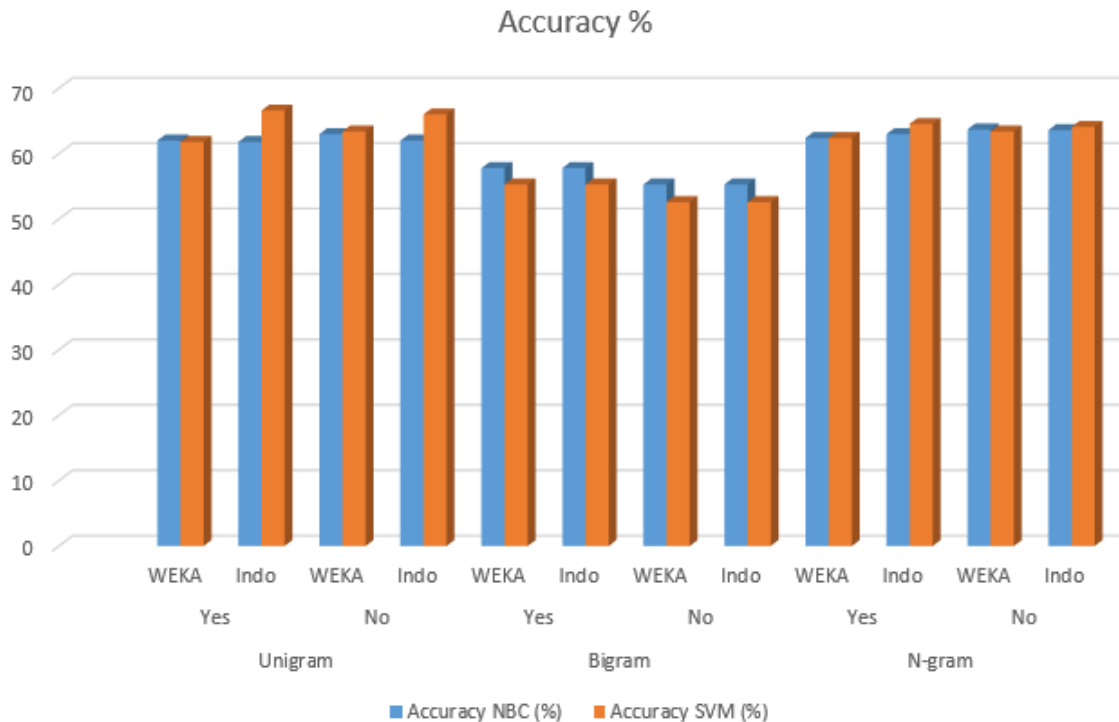
Tabel 1. Perbandingan Hasil Klasifikasi

Naïve Bayes Classifier (NBC)							
Metode Tokenisasi	Emoticons	Stopword List	Accuracy (%)	Precision (%)	Recall (%)	TP Rate (%)	TN Rate (%)
Unigram	Yes	WEKA	62	62,1	62,1	62,1	60
		Indo	61,8	62,1	61,9	61,9	56,6
	No	WEKA	63	63,1	63	63	60,4
		Indo	62	62,4	62,1	62,1	55,5
Bigram	Yes	WEKA	57,8	59,2	57,9	57,9	78,9
		Indo	57,8	59,2	57,9	57,9	78,9
	No	WEKA	55,3	56,9	55,4	55,4	81,5
		Indo	55,3	56,9	55,4	55,4	81,5
N-gram	Yes	WEKA	62,4	62,7	62,6	62,6	61,9
		Indo	63	63,2	63	63	59,6
	No	WEKA	63,7	63,9	63,8	63,8	61,9
		Indo	63,6	63,7	63,6	63,6	60
Support Vector Machine (SVM)							
Metode Tokenisasi	Emoticons	Stopword List	Accuracy (%)	Precision (%)	Recall (%)	TP Rate (%)	TN Rate (%)
Unigram	Yes	WEKA	61,8	62	61,9	61,9	69,1
		Indo	66,6	67,1	66,7	66,7	75,8
	No	WEKA	63,4	63,5	63,4	63,4	69,8
		Indo	66	66,5	66,1	66,1	75,5
Bigram	Yes	WEKA	55,3	57,5	55,4	55,4	84,5
		Indo	55,3	57,5	55,4	55,4	84,5
	No	WEKA	52,6	53,9	52,7	52,7	85,7
		Indo	52,6	53,9	52,7	52,7	85,7
N-gram	Yes	WEKA	62,4	63,4	62,5	62,5	77,4
		Indo	64,55	66,2	64,6	64,6	81,5
	No	WEKA	63,4	64,3	63,4	63,4	77
		Indo	64,1	65,8	64,2	64,2	81,1

*) Nilai *Precision* dan *Recall* merupakan nilai rata-rata dari nilai kelas positif dan kelas negatif.

Tabel 1. Dan Tabel 2. berisi informasi mengenai nilai akurasi, presisi, *recall*, *TP rate* dan *TN rate* dari masing-masing uji coba yang telah dilakukan. Bagian kolom berisi informasi mengenai jenis metode tokenisasi, penggunaan *emoticons*, *stopword list* yang digunakan dalam masing-masing uji coba, nilai akurasi, presisi, *recall*, *TP rate* dan *TN rate*. Sedang bagian baris berisi nilai akurasi, presisi, *recall*, *TP rate* dan *TN rate* dari masing-masing uji coba yang telah dilakukan. Penelitian ini menggunakan 3 metode tokenisasi, yaitu *unigram*, *bigram*, dan *n-gram*. Pada data *preprocessing* masing-masing metode tokenisasi menggunakan 2 jenis *stopword list* (*stopword list* WEKA dan *stopword list* Bahasa Indonesia) dengan menggunakan *emoticons* dan menghapus *emoticons*. Dari proses data *preprocessing* menghasilkan sejumlah token yang kemudian digunakan sebagai input sebuah proses klasifikasi. Proses klasifikasi dilakukan menggunakan metode *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)*. Dari proses klasifikasi diperoleh nilai akurasi, presisi, *recall*, *TP rate* dan *TN rate* dari masing-masing uji coba.

ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN *SUPPORT VECTOR MACHINE*
 (Ghulam Asrofi Buntoro)

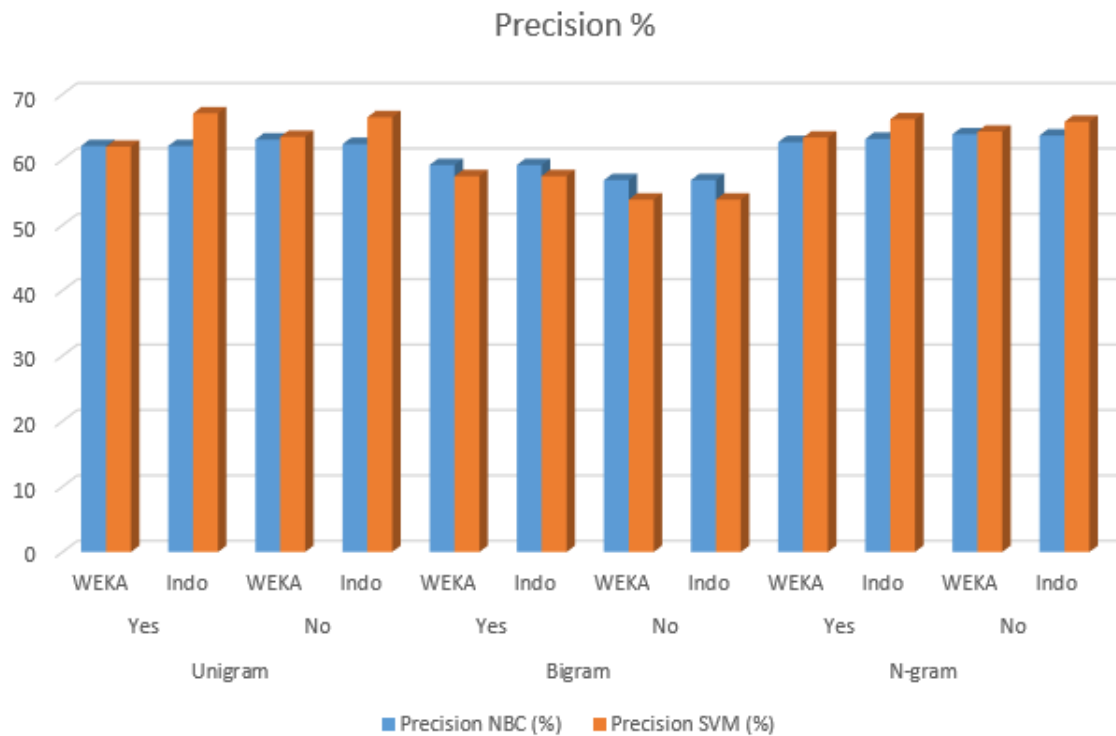


Gambar 4. Grafik tingkat akurasi *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)*

Dari Gambar 4. dapat dilihat bahwa nilai akurasi dicek dengan dua metode klasifikasi yaitu *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* dan dua jenis *stopword list (stopword list WEKA dan stopwords list Bahasa Indonesia)* dengan menggunakan *emoticons* dan menghapus *emoticons*. Pada penelitian ini akurasi tertinggi diperoleh dengan menggunakan metode klasifikasi *Support Vector Machine (SVM)*, tokenisasi *unigram*, *stopword list Bahasa Indonesia dengan emoticons*. kenapa saat menggunakan metode klasifikasi *Support Vector Machine (SVM)* mendapatkan akurasi paling tinggi, karena metode ini mengenali persebaran pola kata dalam kalimat. Dengan menggunakan tokenisasi *unigram* metode klasifikasi *Support Vector Machine (SVM)* akan lebih akurat dalam mengenali pola kata dalam kalimat karena dipecah menjadi satu kata tiap token. Sedangkan nilai akurasi terendah didapat saat menggunakan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *bigram*, kenapa tokenisasi *bigram* mendapatkan akurasi paling rendah, karena dipecah jadi dua kata tiap tokennya, karena nilai n minimum 2 dan n maksimum 2 kata per token. Dengan 2 kata tiap token akan membuat kombinasi pola yang muncul menjadi terlalu banyak sehingga sulit mencari fungsi klasifikasi yang optimal. Untuk metode klasifikasi *Naïve Bayes Classifier (NBC)* hasilnya cenderung lebih stabil, tidak terlalu tinggi dan tidak terlalu rendah karena metode ini berbasis probabilitas munculnya kata dalam setiap kalimat. Nilai akurasi adalah salah satu parameter penilaian dari metode yang telah digunakan, nilai akurasi didapat dari jumlah banyaknya data yang berhasil diklasifikasi dengan benar sesuai kelas sentimennya dari seluruh jumlah data yang diklasifikasi. Nilai akurasi yang tinggi didapat ketika banyak data yang berhasil diklasifikasi dengan benar sesuai kelas sentimen.

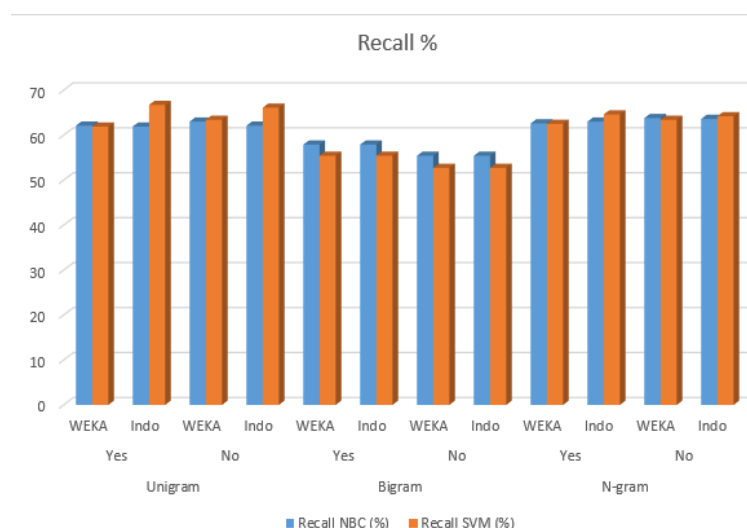
Gambar 4. di atas menunjukkan bahwa penggunaan *stopword list Bahasa Indonesia* pada klasifikasi *Naïve Bayes Classifier (NBC)* tidak mampu meningkatkan akurasi, bahkan sebaliknya yang terjadi, saat menggunakan *stopword list Bahasa Indonesia* akurasi menjadi turun. Karena *stopword list Bahasa Indonesia* menghapus kata-kata negasi, misalnya, "tidak", "bukan", dan sebagainya, jadi kata yang positif bisa menjadi negatif dan sebaliknya, contoh kalimat positif, "tidak jahat" tapi karena kata "tidak" dihapus, kata itu menjadi negatif. Sebaliknya, *stopword list Bahasa Indonesia* saat digunakan dalam metode klasifikasi *Support Vector Machine (SVM)* mampu meningkatkan akurasi. Karena berbasis pola persebaran kata dalam kalimat, *stopword list Bahasa Indonesia* justru meningkatkan akurasi karena data yang digunakan memang dalam Bahasa Indonesia.

Gambar 4. juga menunjukkan bahwa penggunaan *emoticons* pada klasifikasi *Support Vector Machine (SVM)* dan tokenisasi *unigram* mampu meningkatkan nilai akurasi, sedangkan pada saat digunakan dalam metode klasifikasi *Naïve Bayes Classifier (NBC)* dengan tokenisasi *unigram*, akurasi saat menggunakan *emoticons* tidak lebih tinggi dibandingkan yang tanpa *emoticons*.



Gambar 5. Grafik tingkat presisi

Dari Gambar 5. dapat diketahui nilai precision tertinggi diperoleh dengan metode klasifikasi *Support Vector Machine (SVM)* menggunakan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*, sedangkan untuk nilai presisi terendah didapat saat menggunakan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *bigram*, *stopword list* WEKA dan *no emoticons*. Nilai akurasi semakin tinggi akan diikuti nilai presisi yang tinggi juga, begitu sebaliknya, karena nilai presisi adalah jumlah kelas positif yang benar diklasifikasi sebagai kelas positif dibagi total data yang diklasifikasi sebagai kelas positif. Jadi kenapa presisi *unigram* tinggi, karena tokenisasi *unigram* menghasilkan banyak kelas data yang diklasifikasi dengan benar.

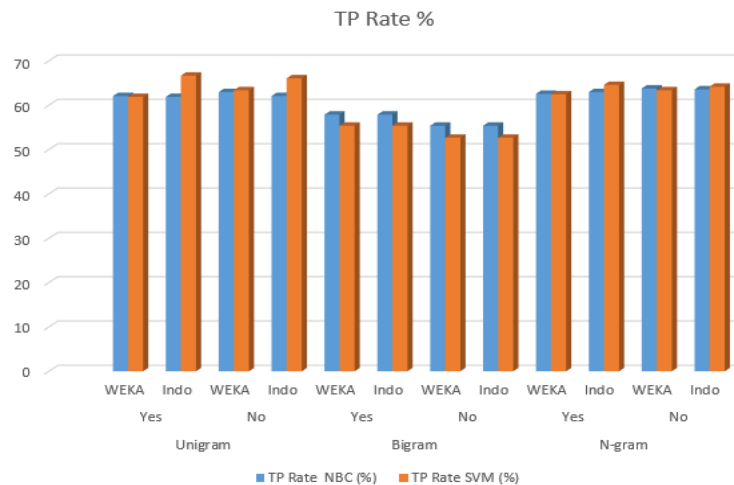


ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN *SUPPORT VECTOR MACHINE*
 (Ghulam Asrofi Buntoro)

Gambar 6. Grafik *recall*

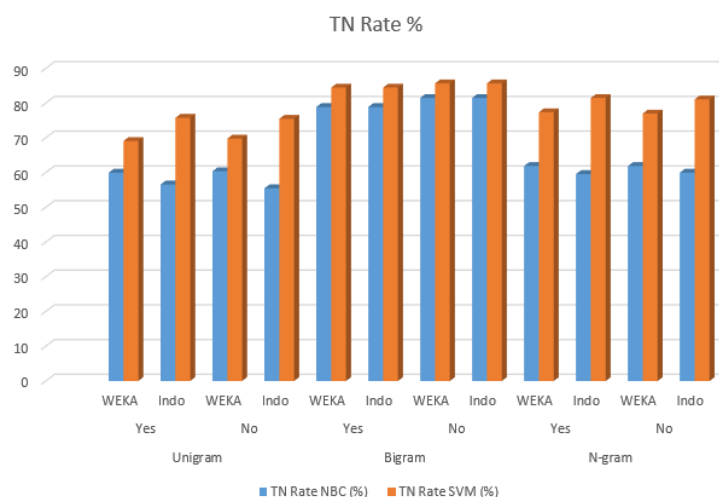
Dari Gambar 6. menunjukkan bahwa nilai *recall* tertinggi didapat saat klasifikasinya menggunakan *Support Vector Machine (SVM)* dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*. Nilai *recall* adalah jumlah kelas positif yang benar diklasifikasi sebagai kelas positif dibagi jumlah kelas positif sebenarnya.

Selain itu, Gambar 6. juga menunjukkan bahwa nilai *recall* yang diperoleh dengan menggunakan tokenisasi *unigram* lebih tinggi dibandingkan nilai *recall* yang diperoleh dengan metode tokenisasi lainnya, nilai *recall* terendah didapatkan saat menggunakan metode tokenisasi *bigram*. Kenapa *bigram* mendapatkan nilai *recall* yang paling rendah, karena dengan tokenisasi ini metode klasifikasi yang digunakan kurang mampu mengklasifikasi kelas data dengan benar, Nilai *recall* didapatkan dari menjumlahkan data yang diklasifikasi benar positif dibagi dengan seluruh data dengan kelas positif.



Gambar 7. Grafik *TP Rate*

Dari Gambar 7. kita dapat mengetahui *TP Rate* tertinggi didapatkan saat menggunakan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*. Nilai *TP Rate* adalah nilai data *tweet* positif yang diklasifikasi dengan benar sesuai kelas sentimennya, yaitu positif.



Gambar 8. Grafik *TN Rate*

Dari Gambar 8. dapat diketahui nilai *TN Rate* tertinggi didapatkan saat menggunakan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia

dan *emoticons*. Nilai *TN Rate* adalah nilai data *tweet* negatif yang diklasifikasi dengan benar sesuai kelas sentimennya, yaitu negatif.

Dari penelitian yang telah dilakukan, diketahui bahwa hasil akurasi tertinggi didapatkan saat menggunakan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*. Meskipun menghasilkan akurasi yang cukup baik, tapi model yang dibangun masih melakukan sedikit kesalahan pada saat proses klasifikasi kelas positif, sedikit kesalahan ini ditunjukkan dengan nilai *TN Rate* yang lebih besar dari nilai *TP Rate*, nilai *TN Rate* sebesar 75,8% dan nilai *TP Rate* sebesar 66,7%. Selisih nilai yang sedikit ini karena data yang digunakan seimbang, jadi tidak menyebabkan banyak kesalahan dalam klasifikasi, bila menggunakan data yang tidak seimbang akan menyebabkan data *minority class* yang salah diklasifikasi sebagai data *majority class*[9] pada akhirnya menjadikan selisih nilai menjadi besar.

KESIMPULAN

Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan bahwa Analisis sentimen twit Bahasa Indonesia dengan tagar Hatespeech dapat membantu menentukan sentimen pada twit opini Bahasa Indonesia yang ada di Twitter. Setelah dilakukan analisis sentimen terlihat banyak twit opini yang sebenarnya tidak masuk kategori Hatespeech tapi diberi tagar Hatespeech. Nilai akurasi tertinggi didapat dengan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*, dengan nilai rata-rata akurasi mencapai 66,6%, nilai presisi 67,1%, nilai *recall* 66,7% nilai *TP rate* 66,7% dan nilai *TN rate* 75,8%. Dalam penelitian ini juga dapat diketahui metode klasifikasi *Support Vector Machine (SVM)* lebih tinggi akurasinya untuk klasifikasi sentiment tweet HateSpeech Bahasa Indonesia dibandingkan metode klasifikasi *Naïve Bayes Classifier (NBC)*. Untuk penelitian selanjutnya perlu dikembangkan *stopword list* dan *stemmer* Bahasa Indonesia yang mampu meningkatkan akurasi dalam analisis sentiment.

REFERENCES

- [1] Pengertian *Hate Speech* (Ucapan Penghinaan atau Kebencian)
<https://hatespeechgroup.wordpress.com/pengertianhatespeech/>
- [2] Top Media Sosial <http://www.evadollzz.com/2014/09/top-10-social-networkings-terpopuler.html>
- [3] Marian Radke Yarrow, John A. Clausen and Paul R. Robbins (2010). *The Social Meaning of Mental Illness*. Journal of Social Issues. Volume 11, Issue 4, pages 33–48, Fall 1955.
- [4] Go, A., Huang, L., & Bhayani, R. (2009). Twitter Sentiment Analysis. Final Project Report, Stanford University, Department of Computer Science.
- [5] Mesut Kaya, Guven Fidan, Ismail H. Toroslu (2012). *Sentiment Analysis of Turkish Political News*. IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology.
- [6] Pak, A., dan Paurobek, P., (2010). *Twitter as a Corpus for Sentiment Analysis and Opinion Mining*, Universite de Paris-Sud, Laboratoire LIMSI-CNRS.
- [7] Jennifer Yang Hui (2014). *Indonesian Presidential Election: Will Social Media Forecasts Prove Right?*
- [8] Read, J. (2005). *Using Emoticons to reduce Dependency in Machine Learning Techniques for Sentiment Classification*. Meeting of the Association for Computational Linguistics - ACL, 43.1-6.
- [9] Pang, B., Lee, L., & Vithyanathan, S. (2002). *Thumbs Up? Sentiment Classification Using Machine Learning Techniques*. Proceedings of The ACL-02 conference on empirical methods in natural language processing (pp. 79-86).

ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE
NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE
(Ghulam Asrofi Buntoro)

- [10] Franky dan Manurung, R., (2008). *Machine Learning-based Sentiment Analysis of Automatic Indonesian n Translations of English Movie Reviews*. In Proceedings of the International Conference on Advanced Computational Intelligence and Its Applications.
- [11] Olson, David L.; & Delen, Dursun (2008); *Advanced Data Mining Techniques*, Springer, 1st edition (February 1, 2008), page 138, ISBN 3-540-76916-1.
- [12] Tala, F. Z. (2003). *A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia*. M.S. thesis. M.Sc. Thesis. Master of Logic Project. Institute for Logic, language and Computation. Universiteti van Amsterdam The Netherlands.
- [13] ARFF files from Text Collections. <http://weka.wikispaces.com/ARFF+files+from+Text+Collections>.
- [14] Class StringToWordVector.
<http://weka.sourceforge.net/doc.de.v/weka/filters/unsupervised/attribute/StringToWordVector.html>.
- [15] Ian H. Witten. (2013) *Data Mining with WEKA*. Department of Computer Science University of Waikato New Zealand.
- [16] Kohavi, & Provost. (1998) *Confusion Matrix*
http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html.