

ANALISIS SENTIMEN TERHADAP ACARA TELEVISI BERDASARKAN OPINI PUBLIK

Aditia Rakhmat Sentiaji¹, Adam Mukaharil Bachtiar²

^{1,2}Teknik Informatika - Universitas Komputer Indonesia
Jl. Dipati Ukur No. 112-116, Bandung 40132

E-mail : aditia.rakhmat@outlook.com¹, adammb@outlook.com²

ABSTRAK

Banyak acara-acara televisi menjadi bahan perbincangan di media sosial, baik karena kualitas acaranya yang bagus ataupun sebaliknya. Sayangnya media sosial tidak mempunyai kemampuan untuk mengagregasi informasi mengenai suatu perbincangan yang ada menjadi sebuah kesimpulan. Salah satu cara untuk menarik kesimpulan dari hasil agregasi adalah menggunakan *text mining*. Karena salah satu fungsi dari *text mining* adalah untuk melakukan analisis sentiment. *Naïve bayes classifier* adalah algoritma yang dimanfaatkan untuk mengimplementasikannya. Agar mendukung performansi yang cepat maka basis data yang digunakan adalah *NoSQL*.

Analisis sentimen menggunakan menggunakan algoritma *naïve bayes classifier* memberikan hasil yang baik. Terbukti dengan pengujian menggunakan algoritma *percentage split* memperoleh akurasi $\pm 90\%$. Dibantu proses *preprocessing* yang bertujuan untuk menghapus bagian yang tidak penting dan juga mengubah bentuk dokumen yang berbentuk *tweet* ke bentuk standar sehingga pengklasifikasian yang dilakukan oleh *naïve bayes* menjadi lebih akurat. Sehingga hasil dari analisis sentimen acara televisi bisa dijadikan referensi dalam menentukan tontonan acara televisi.

Kata kunci : analisis sentiment, *teks mining*, *NoSQL*

1. PENDAHULUAN

Dunia pertelevisian Indonesia kini sedang berkembang. Terbukti dari semakin banyaknya stasiun televisi swasta baru mengudara baik yang bersifat nasional ataupun lokal. Banyaknya jumlah stasiun televisi tentu akan berbanding lurus dengan jumlah dan keberagaman acara televisi. Namun sayangnya dengan banyaknya jumlah acara televisi tidak diimbangi dengan kualitas acara tersebut. Banyak acara televisi yang lebih mengedepankan aspek hiburan akan tetapi tidak mendidik. Banyak acara yang menonjolkan kekerasan, saling menghina, mengumbar aib seseorang, mengeksploitasi kekurangan seseorang dan lain

sebagainya. Hal tersebut banyak dilakukan hanya untuk meningkatkan *rating* dan *share* acara tersebut. Senada dengan apa yang diungkapkan oleh dosen komunikasi Universitas Indonesia dan anggota tim panel pemantau KPI Pusat, Nina Mutmainah Armando. Menurut beliau, seringkali ditampilkan acara yang melanggar norma kesopanan melecehkan orang, menggunakan bahasa kasar dan menggunakan anak dalam *setting* film yang tidak pantas [1].

Sentimen publik bisa dijadikan sebagai indikator untuk melihat apakah acara tersebut berkualitas atau tidak. Media sosial merupakan media yang sering digunakan untuk menuangkan sentimen atau opini publik mengenai acara televisi tersebut. Banyak acara-acara televisi menjadi bahan perbincangan di media sosial, baik karena kualitas acaranya yang bagus ataupun sebaliknya. Sebagai contoh di twitter, acara televisi yang sering di-tweet dalam suatu waktu akan menjadi *trending topics*. Hal tersebut bisa membuat pengguna yang melihat tertarik untuk menonton acara televisi tersebut ataupun tidak sama sekali. Tergantung dari konteks yang diperbincangkan. Sebagai informasi, twitter memiliki sekitar 500 juta pengguna atau sekitar 4% dari seluruh pengguna twitter berasal dari Indonesia [2]. Angka-angka tersebut menggambarkan akan banyak sekali yang terpengaruh dari perbincangan yang lagi ramai dibahas seperti acara televisi.

Sayangnya media sosial tidak mempunyai kemampuan untuk mengagregasi informasi mengenai suatu bahasan yang ada menjadi sebuah kesimpulan. *Trending topics* yang ada di twitter pun hanya menampilkan topik yang sedang banyak diperbincangkan tanpa memberikan suatu kesimpulan. Diperlukan metode khusus agar informasi seperti acara televisi dapat menggunakan banyak sudut pandang yang bisa digunakan untuk mengambil sebuah kesimpulan tentang positif atau negatifnya suatu acara televisi.

Berdasarkan penjelasan sebelumnya, diperlukan sebuah cara agar dapat mengklasifikasikan opini publik menjadi pengetahuan baru berupa kesimpulan negatif atau positifnya mengenai acara televisi dari data yang ada di media sosial. Hal tersebut dimungkinkan dengan menggunakan *text mining*. *Text mining* yang juga dikenal dengan *text data*

mining adalah sebuah proses yang semi otomatis melakukan klasifikasi dari pola yang ada dari *database* yang tidak terstruktur [3]. Sehingga, hasil dari klasifikasi tersebut bisa menjadi media alternatif bagi masyarakat untuk memilih acara televisi yang berkualitas.

Berdasarkan hasil dari beberapa penelitian mengenai pengklasifikasian informasi subjektif atau yang sering disebut analisis sentimen. Dibutuhkan suatu algoritma untuk dapat mengklasifikasikan suatu opini ke dalam kelas negatif atau positif. Adapun algoritma yang dapat digunakan dalam pengklasifikasian adalah C45, *Support Vector Machine* (SVM), *Naïve Bayes Classifier* (NBC), *K-Nearest Neighbors*, *Information Fuzzy Networks*, dan masih banyak algoritma lainnya. Melihat hasil dari penelitian Fatimah Wulandini dan Anto Satriyo Nugroho (*Text Classification Using Support Machine for Webmining Based Spation Temporal Analysis of the Spread of Tropical Disease*, 2009) mendapatkan hasil bahwa algoritma SVM menunjukkan akurasi paling tinggi pada kategorisasi teks Bahasa Indonesia dengan presentase 92.5%, tidak jauh berbeda dengan algoritma NBC yang memiliki presentase 90% [4].

Apabila dilihat kompleksitasnya, NBC jauh lebih konvensional dan sederhana. Hal tersebut berpengaruh terhadap waktu komputasi yang dibutuhkan. NBC memerlukan waktu komputasi lebih singkat dibanding dengan SVM yang berkembang secara kuadratik seiring dengan perkembangan jumlah data latih. Berdasarkan dari kompleksitasnya, waktu yang dibutuhkan untuk komputasi yang lebih rendah, dan akurasi yang tidak terlalu jauh berbeda dari SVM, NBC lebih cocok untuk diimplementasikan dalam penelitian ini.

1.1 Text Mining

Text mining yang juga dikenal dengan *text data mining* atau pencarian pengetahuan di basis data textual adalah sebuah proses yang semi otomatis melakukan ekstraksi dari pola yang ada di *database*. Dari hasil ekstraksi tersebut munculah pengetahuan baru yang bisa dimanfaatkan untuk kepentingan pengambilan keputusan. *Text mining* mempunyai kesamaan dengan *data mining*. Keduanya memiliki tujuan yang sama yaitu untuk memperoleh informasi dan pengetahuan dari sekumpulan data yang sangat besar. Data tersebut bisa berbentuk sebuah *database*. Namun keduanya memiliki perbedaan jenis data. *Data mining* memiliki input data dari data yang sudah terstruktur sedangkan *text mining* dimulai dengan data yang tidak terstruktur.

Pemanfaatan dari *text mining* secara nyata sangatlah luas. Areanya seluas data tekstual yang terbentuk seperti di area hukum dengan data putusan pengadilan, penelitian dengan data artikel penelitian, keuangan dengan data laporan triwulan, teknologi dengan data arsip paten, pemasaran dengan data

komentar konsumen, dan di area lainnya. Sebagai contoh sebuah perusahaan membuat formulir yang biasa diisi apabila konsumennya ingin memuji, komplain, ataupun klaim garansi. Dari kartu formulir tersebut terbentuklah data yang sangat besar dan bisa digunakan untuk mengidentifikasi secara objektif produk dan layanan dari suatu perusahaan menggunakan *text mining*. Selain itu proses *text mining* yang dilakukan secara otomatis adalah di bidang komunikasi elektronik dan email. *Text mining* tidak hanya mengklasifikasikan dan menyaring email sampah, tetapi bisa juga memprioritaskan email secara otomatis berdasarkan tingkat kepentingannya [3].

1.2 Analisis Sentimen

Sejarah analisis sentimen pertama kali muncul pada sebuah jurnal karya Das, Chen, dan Tong pada tahun 2001, bahasan yang mereka angkat sesuai dengan minat mereka yaitu menganalisis sentimen pasar. Analisis sentimen adalah mengekstraksi pendapat, sentimen, evaluasi, dan emosi orang tentang suatu topik tertentu yang tertulis menggunakan teknik pemrosesan bahasa alami. Sejumlah karya-karya besar lainnya menyebutkan analisis sentimen fokus pada aplikasi spesifik yang mengklasifikasikan mengenai sifat yang berlawanan (antara positif dan negatif). Dari pengertian tersebut menjadi sebuah fakta yang menyebabkan beberapa penulis bahwa istilah analisis sentimen mengacu pada tugas yang sempit atau spesifik. Namun saat ini banyak yang menafsirkan istilah analisis sentiment lebih luas lagi yang berarti cara pengkomputasian pendapat, sentimen, dan subjektivitas pada teks [5].

1.3 Preprocessing

Tahap *preprocessing* diperlukan untuk membersihkan data dari yang tidak diperlukan, dengan tujuan pada tahap masuk ke dalam metode *Naïve Bayes Classifier* lebih optimal dalam perhitungannya. Pada tahap ini melibatkan rekognisi dari isi dan struktur teksnya. Adapun tahapan-tahapan dari *preprocessing*.

1. Convert Emoticon

Emoticon merupakan salah satu cara pengungkapan ekspresi perasaan secara tekstual. Hal tersebut akan membantu dalam menentukan sentiment dalam suatu kalimat atau tweet sekalipun. Setiap *emoticon* akan dikonversikan ke dalam *string* yang bersesuaian. Jenis *emoticon* yang akan diproses oleh adalah *emoticon western style*, dikarenakan jenis tersebut banyak digunakan atau menjadi standar di semua platform seperti *web* ataupun *mobile*. Pada Tabel 1 dijelaskan hasil pengklasifikasian arti dari *emoticon western style* secara umum.

Tabel 1 Emoticon yang akan Dikonversi

Emoticon	Deskripsi
:~) :) :o) :) :3 :c) :> =] 8)	Senang
=) :) :^) :~)	
:D :D 8-D 8D x-D xD X-D XD =- D=D =-3 =3 B^D	Ketawa
:- :@ >:(Benci
:\$ >:[:-(:(- :~c :c :~< :~C :< :- [:[:{ :(- :(- :(-D :< D: D8 D; D= DX v.v D-':	Kecewa
<3 :-) ;) *-) *) ;~] ;] ;D ;^) :-,	Suka

2. Cleansing

Cleansing merupakan proses membersihkan kata-kata yang tidak diperlukan untuk mengurangi noise. Kata yang dihilangkan adalah URL, *hashtag* (#), *username* (@username), dan email. Selain itu juga tanda baca seperti titik(.), koma(,), dan tanda baca yang lainnya akan dihilangkan. Pada Tabel 2 dijelaskan contoh penerapan dari *cleansing*:

Tabel 2 Penerapan Cleansing

Data Latih	
Input	Output
Saya suka video @YouTube http://youtu.be/wjLPx8J1dDs?a MATA NAJWA - Memilih Wakil Rakyat - Iwan Fals, Ahok, JK & DS (Full)	Saya suka video MATA NAJWA Memilih Wakil Rakyat Iwan Fals Ahok JK DS Full
Stand Up Comedy Anang! stand up comedy Anang Lucu Banget Terbaru 2014 indonesia metro tv http://www.youtube.com/watch?v=JoSg0a4Q8vY ...	Stand Up Comedy Anang stand up comedy Anang Lucu Banget Terbaru indonesia metro tv

3. Case Folding

Case folding merupakan tahapan merubah bentuk kata-kata menjadi sama bentuknya, baik semuanya menjadi *lower case* ataupun menjadi *upper case*. Pada Tabel3 dijelaskan contoh dari penerapan proses *case folding*.

Tabel 3 Penerapan Case Folding

Data Latih	
Input	Output
Saya suka video MATA NAJWA Memilih Wakil Rakyat Iwan Fals Ahok JK DS Full	saya suka video mata najwa memilih wakil rakyat iwan fals ahok jk ds full

Stand Up Comedy Anang stand up comedy Anang Lucu Banget Terbaru di indonesia metro tv	stand up comedy anang stand up comedy anang lucu banget terbaru di indonesia metro tv
---	---

4. Convert Negation

Seperti halnya ilmu matematika, dalam bahasa terdapat kata yang dapat membalikan arti dari kata tersebut atau bersifat negasi. Kata-kata yang bersifat negasi adalah “kurang”, “tidak”, “enggak”, “ga”, “nggak”, “tak”, dan “gak”. Pada Tabel 4 dijelaskan contoh penerapannya.

Tabel 4 Penerapan Convert Negation

Sebelum Convert Negation	Setelah Convert Negation
gak seru final indonesia idol tanpa virzha sabar virzha sukses pulang	ga_seru final indonesia idol tanpa virzha sabar virzha sukses pulang

5. Tokenizing

Tokenizing bekerja untuk mengidentifikasi kata-kata dalam teks menjadi beberapa urutan yang terpotong oleh spasi atau karakter spesial. Berikut contoh penerapan dari *tokenizing*.

Tabel 5 Penerapan Tokenizing

Data Latih	
Input	Output
stand up comedy anang stand up comedy anang lucu banget terbaru di indonesia metro tv	[stand, up, comedy, anang, stand, up, comedy, anang, lucu, banget, terbaru, di, indonesia, metro, tv]

6. Stopping

Stopping berperan untuk membuang kata-kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansinya dengan teks. Kata-kata yang akan dibuang tersebut didefinisikan dalam *stopword list*. Contoh beberapa kata yang sering masuk ke dalam *stopword list* adalah “sebuah”, “yang”, dan “itu”. Berikut adalah beberapa daftar *stoplist* yang disimpan dalam *database*.

Tabel 6 Stoplist yang digunakan

masih	Dong	Ke	Ada	Yoi
malam	Ya	Loe	Pada	Yang
Ini	Dan	Juga	Kita	Saya
untuk	Dari	Bagi	Iya	di mana
kapan	Bisa	Mana	Itu	Sih
sudah	Bikin	dengan	Anda	Begitu

entah	Lalu	Yuk	Aku	Adalah
gue	Nanti	tunggu	Tau	Kemarin

Sebelum #jum'atan stay di trans7 lihat video2 lucu ala @CCTV_T7

7. Stemming

Stemming adalah tahapan untuk membuat kata yang berimbuhan kembali ke bentuk asalnya. Contohnya kata “memberikan” setelah melewati tahap ini maka akan menjadi “beri” [6].

Tabel 8 Penerapan *Stemming*

Data Latih	
Input	Output
menginspirasi	inspirasi

1.4 Algoritma Naïve Bayes

Naïve bayes classifier adalah penggolong menggunakan statistik sederhana berdasarkan teorema bayes yang mengasumsikan bahwa keberadaan atau ketiadaan dari suatu fitur tertentu dari suatu kelas tidak berhubungan dengan keberadaan atau ketiadaan fitur lainnya. Sifatnya yang sebagai model probabilitas, *naïve bayes classifier* bisa dilatih dengan efisien sebagai *supervised learning*. Pada pengaplikasiannya, parameter estimasi untuk model naïve bayes menggunakan metode kemungkinan maksimum. Dengan kata lain, masih bisa bekerja dengan model naïve bayes tanpa harus mempedulikan bayesian probabilitas atau metode bayesian lainnya. Berikut adalah model matematis untuk *naïve bayes classifier*:

$$p(C|F_1, \dots, F_n) = \frac{p(C)p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)} \quad (1)$$

2. ISI PENELITIAN

2.1 Sumber Data

Data yang digunakan berbentuk *tweet* yang diambil langsung memanfaatkan *API twitter*. Berikut contoh data yang digunakan.

Tabel 10 Contoh Data yang Digunakan

Tweet
@OVJ_Trans7 gak seru, gak lucu.! Kalo gak ada kang @sule_ovj
@OVJ_Trans7 gak ada #Sule udah gak lucu , Garingg bgt sekarang .. Huekkkk
YKS ga ada kang wendi, Ga seru aah :(@WCIndonesia
@DEV0TEES bentar lg ada Rising Star Indonesia :-) bakal lebih bagus dr Idol deh,,
ILK (Indonesia lawak klub) bikin ngakak :D :D

2.2 Implementasi Preprocessing

Sebelum masuk proses utama yaitu pengklasifikasian tweet menggunakan algoritma naïve bayes, data akan diolah terlebih dahulu pada tahap *preprocessing* agar pada tahap klasifikasi hasilnya bisa jauh lebih optimal. Pada Tabel 11 menunjukkan hasil *preprocessing* dari data pada Tabel 10.

Tabel 11 Data Hasil *Preprocessing*

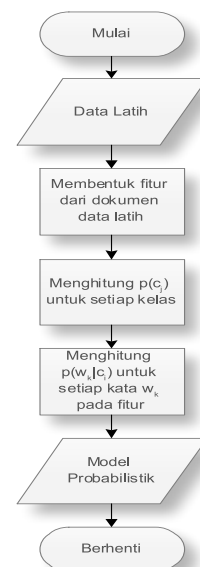
Tweet
[ga_seru, ga_lucu, ga_ada, kang]
[ga_ada, sule, udah, ga_lucu, garing, bgt, huekkkk]
[yks, ga_ada, kang, wendi, ga_seru]
[bentar, rising, star, indonesia, lebih, bagus, idol]
[ilk, indonesia, lawak, klub, bikin, ngakak, ketawa, ketawa]
[stay, trans, lihat, video, lucu, ala]

2.3 Implementasi Algoritma Naïve Bayes

Tahap ini merupakan tahap yang paling esensial dari tahap yang lainnya. Pada tahap ini proses pengklasifikasian berdasarkan sentimen yang ada di dalam dokumen dimulai. Tahap ini mempunyai dua proses, berikut prosesnya:

1. Proses *learning naïve bayes classifier*

Naïve bayes classifier sebagai algoritma *supervised learning* harus diberi pengetahuan awal terlebih dahulu, sebagai acuan untuk dapat mengklasifikasikan suatu dokumen berdasarkan sentimennya.



Gambar 1 Flowchart Training *Naïve Bayes*

Ada tiga langkah dalam proses *learning*. Berikut ketiga tahapan pada proses *learning*.

a. Membentuk Fitur

Dalam penelitian ini yang dimaksud dengan fitur adalah kata kunci yang akan menjadi parameter satuan data latih, yaitu dokumen (*tweet*) untuk diklasifikasikan ke dalam kelas yang telah ditentukan (positif atau kelas negatif). Dalam kata lain fitur adalah kata yang memiliki nilai sentimen. Tabel 10 berisikan pembentukan fitur dari data latih yang sudah dilakukan proses *preprocessing* terlebih dahulu (Tabel 10).

Tabel 11 Pembentukan Fitur dari Data Latih

Data	Fitur(Kemunculan)	Kelas Sentimen
D1	ga_seru(1), ga_lucu(1)	Negatif
D2	ga_seru(1)	Negatif
D3	ga_lucu(1)	Negatif
D4	bagus(1)	Positif
D5	ngakak(1),ketawa(2)	Positif
D6	lucu(1)	Positif

b. Menghitung probabilitas $p(c_i)$

Setelah dibentuk fitur dengan kemunculannya dari data latih. Selanjutnya menghitung probabilitas dari setiap kelas dengan cara sebagai berikut:

$$p(c_i) = \frac{fd(c_i)}{|D|} \quad (2)$$

Keterangan :

$fd(c_i)$ = Jumlah dokumen yang termasuk c_i

$|D|$ = Jumlah data latih / jumlah *tweet*

Tabel 13 Probabilitas Kelas *Training Naïve Bayes*

Kelas Sentimen (c)	Data(j)						$fd(C_j)$	$P(C_j)$
	D 1	D 2	D 3	D 4	D 5	D 6		
Positif	0	0	0	1	1	1	3	$\frac{3}{6}$
Negatif	1	1	1	0	0	0	3	$\frac{3}{6}$

c. Menentukan probabilitas $p(w_k|c_i)$

Setelah didapat probabilitas dari setiap kelas, selanjutnya menghitung probabilitas setiap fitur pada setiap kelas dengan cara sebagai berikut:

$$p(w_k|c_i) = \frac{f(w_{ki}, c_i) + 1}{f(c_i) + |W|} \quad (3)$$

Keterangan :

$f(w_{ki}, c_i)$ = Nilai kemunculan kata w_{ki} pada kelas c_i

$f(c_i)$ = Jumlah keseluruhan kemunculan kata pada kelas c_i

$|W|$ = Jumlah keseluruhan dari w_k

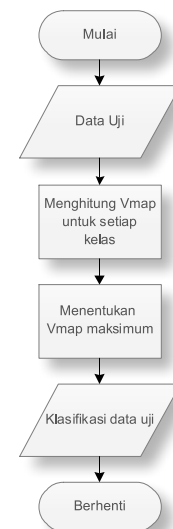
Dari rumus 3, maka diperoleh model probabilistik untuk setiap fitur pada 12.

Tabel 14 Model Probabilitas dari Data Latih

Data $f(W_{ki}, C_i)$	Kelas Sentimen (c)	
	Positif	Negatif
ga_seru	$\frac{0+1}{5+7} = \frac{1}{12}$	$\frac{2+1}{4+7} = \frac{3}{11}$
ga_lucu	$\frac{0+1}{5+7} = \frac{1}{12}$	$\frac{2+1}{4+7} = \frac{3}{11}$
Ketawa	$\frac{2+1}{5+7} = \frac{3}{12}$	$\frac{0+1}{4+7} = \frac{1}{11}$
bagus	$\frac{1+1}{5+7} = \frac{2}{12}$	$\frac{0+1}{4+7} = \frac{1}{11}$
lucu	$\frac{1+1}{5+7} = \frac{2}{12}$	$\frac{0+1}{4+7} = \frac{1}{11}$
ngakak	$\frac{1+1}{5+7} = \frac{2}{12}$	$\frac{0+1}{4+7} = \frac{1}{11}$

2. Proses klasifikasi naïve bayes classifier

Berikut alur proses dari proses klasifikasi menggunakan *naïve bayes classifier*.



Gambar 2 Flowchart Klasifikasi Naïve Bayes

Berikut adalah contoh satu *tweet* yang akan dijadikan data uji menggunakan model probabilitas fitur Tabel 4:

Tabel 15 Data Uji Klasifikasi

Tweet	
Sebelum <i>Preprocessing</i>	Setelah <i>Preprocessing</i>
Jd inget wkt msk studio Dahsyat rcti ,teh siapa gitu blg gini "guys! disini	inget studio dahsyat rcti lucu ga_lucu ketawa maksa gumprengr

lucu ga lucu hrs ketawa ya" maksa gumpreng	
--	--

Dari alur proses klasifikasi ada dua tahapan penting. Berikut penjelasan tahapan-tahapannya:

a. Menghitung Vmap

Vmap adalah perhitungan yang digunakan *naïve bayes classifier* untuk menentukan probabilitas data uji dari masing-masing kelas berdasarkan dari proses *learning*. Nilai probabilitas yang terbesar akan dipilih. Berikut perhitungannya:

$$Vmap = \operatorname{argmax}_c p(c) \prod p(w_k | c) \times p(c)$$

Berdasarkan dari hasil *training*, berikut hasil perhitungannya:

$$Vmap = \operatorname{argmax}_{\{\text{positif, negatif}\}} p(w_k | c) \times p(c) \quad (4)$$

$$Vmap = \operatorname{argmax}_{\{j \in \{\text{positif, negatif}\}\}} p(c_i) p(\text{"inget"} | c_i) p(\text{"studio"} | c_i) p(\text{"dahsyat"} | c_i) p(\text{"rcti"} | c_i) p(\text{"lucu"} | c_i) p(\text{"ga_lucu"} | c_i) p(\text{"ketawa"} | c_i) p(\text{"maksal"} | c_i) p(\text{"gumpreng"} | c_i)$$

1) Vmap untuk sentimen positif

$$\begin{aligned} Vmap(\text{"positif"}) &= P(\text{"positif"}) P(\text{"inget"} | \text{"positif"}) P(\text{"studio"} | \text{"positif"}) P(\text{"dahsyat"} | \text{"positif"}) \\ &P(\text{"rcti"} | \text{"positif"}) P(\text{"lucu"} | \text{"positif"}) P(\text{"ga_lucu"} | \text{"positif"}) P(\text{"ketawa"} | \text{"positif"}) \\ &P(\text{"maksal"} | \text{"positif"}) P(\text{"gumpreng"} | \text{"positif"}) \\ &= \frac{1}{2} \times \frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} \times \frac{2}{12} \times \frac{1}{12} \times \frac{3}{12} \times \frac{1}{12} \times \frac{1}{12} \\ &= 0.00000000058 \end{aligned}$$

2) Vmap untuk sentimen negatif

$$\begin{aligned} Vmap(\text{"negatif"}) &= P(\text{"negatif"}) P(\text{"inget"} | \text{"negatif"}) P(\text{"studio"} | \text{"negatif"}) \\ &P(\text{"dahsyat"} | \text{"negatif"}) P(\text{"rcti"} | \text{"negatif"}) P(\text{"lucu"} | \text{"negatif"}) P(\text{"ga_lucu"} | \text{"negatif"}) \\ &P(\text{"ketawa"} | \text{"negatif"}) P(\text{"maksal"} | \text{"negatif"}) P(\text{"gumpreng"} | \text{"negatif"}) \\ &= \frac{1}{2} \times \frac{1}{11} \times \frac{1}{11} \times \frac{1}{11} \times \frac{1}{11} \times \frac{1}{11} \times \frac{3}{11} \times \frac{1}{11} \times \frac{1}{11} \times \frac{1}{11} \\ &= 0.00000000064 \end{aligned}$$

b. Menentukan Vmap maksimum

Dari hasil perhitungan Vmap diatas didapatkan bahwa nilai Vmap negatif lebih besar dibanding dengan Vmap positif. Bisa disimpulkan bahwa *tweet* tersebut diklasifikasikan ke dalam sentimen negatif.

3. PENUTUP

Dari hasil penelitian yang telah dilakukan terlihat bahwa algoritma *naïve bayes* dapat mengklasifikasikan suatu opini yang berupa *tweet* ke dalam dua kelas yaitu positif dan negatif dengan akurat. Tingkat keakurasian dari pengklasifikasian tersebut sangat dipengaruhi oleh proses *training*. Sehingga dapat disimpulkan dari hasil

pengklasifikasian yang disajikan dalam bentuk grafik dapat terlihat dengan jelas informasi sentimen publik terhadap suatu acara televisi dan dapat dijadikan sebagai referensi untuk menonton acara televisi.

DAFTAR PUSTAKA

- [1] F. Rozy, "Rakyat Merdeka Online," Indonesian Online Media Syndicate, 2 November 2009. [Online]. Available: <http://www.rakyatmerdeka.co.id/news/2009/11/02/83360/Pakar:-Televisi-Cenderung-Tampilkan-Wajah-Buruk>. [Diakses 19 Maret 2014].
- [2] J. Bernstein, "Social Media in 2013: By the Numbers," Social Media Today Community, 6 November 2013. [Online]. Available: <http://socialmediatoday.com/jonathan-bernstein/1894441/social-media-stats-facts-2013>. [Diakses 27 Januari 2014].
- [3] E. Turban, R. Sharda dan D. Delen, Decision Support and Business Intelligence Systems, 2011: Pearson, New Jersey.
- [4] F. Wulandini dan A. S. Nugroho, "ext Classification Using Support Vector Machine for Webmining Based Spation Temporal Analysis of the Spread of Tropical Diseases," International Conference on Rural Information and Communication Technology, Jakarta, 2009.
- [5] B. Pang dan L. Lee, "Opinion Mining and Sentiment Analysis," *Foundation and Trends In Information Retrieval*, vol. 2, p. 10, 2008.
- [6] W. B. Croft, D. Metzler dan T. Strohman, "Document Parsing," dalam *Search Engines Information Retrieval in Practice*, Boston, Pearson, 2010, pp. 86 - 101.