

ANALISIS SENTIMENT TWITTER TERHADAP CAPRES INDONESIA 2019 DENGAN METODE K-NN

Abdul Malik Zuhdi¹⁾, Ema Utami²⁾, Suwanto Raharjo³⁾

^{1,2}Megister Teknik Informatika Universitas Amikom Yogyakarta

³Teknik Informatika, Fakultas Teknologi Industri, Institut Sains & Teknologi AKPRIND
Yogyakarta

Email : ¹zuhdi.smktaq@gmail.com, ²emma@nrar.net, ³wa2n@nrar.net

Abstrak

Twitter adalah media sosial di mana banyak orang dari seluruh dunia dapat mengungkapkan pendapat mereka. Di Twitter masyarakat ramai berbincang tentang pemilihan pilpres 2019, sehingga bermunculan opini yang bersifat positif maupun negatif tentang calon presiden. Penelitian ini diharapkan dapat bermanfaat membantu riset atas opini masyarakat yang mengandung sentimen positif, negatif dan netral. Tahapan penelitian dilakukan dengan pengumpulan data tweet, preprosesing, tokenisasi, penentuan class atribut, load directory, determine sentiment, kemudian dilanjutkan klasifikasi, dengan pengujian nilai k dan pengujian confusion matrix. Dari hasil penelitian dan pengujian, metode K-NN memiliki tingkat akurasi klasifikasi sentimen mencapai 83,33%.

Kata Kunci: Sentimen analisis, Twitter, K-NN, Capres.

1. PENDAHULUAN

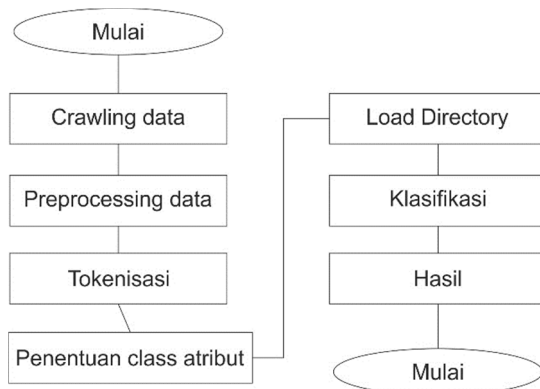
Twitter adalah media sosial di mana banyak orang dari seluruh penjurus dunia dapat mengungkapkan pendapat mereka. Data yang dihasilkan di Twitter bisa sangat berguna jika dianalisis karena data tersebut dapat diekstrak menjadi informasi penting melalui opini mining. Pendapat tentang berita atau peluncuran produk apapun atau bahkan jenis tren tertentu dapat diamati dengan baik di Twitter (Dubey, 2017). Maraknya Pilpres di Indonesia menyebabkan Twitter digunakan untuk kepentingan politik oleh masyarakat atau institusi politik. Twitter berperan aktif dalam proses komunikasi antara institusi politik dengan masyarakat terutama pada saat kampanye berlangsung (Rasyadi, 2017). Selama masa kampanye, banyak berita terbaru tentang capres yang bermunculan, menyebabkan masyarakat beropini positif atau negatif tentang sebuah berita. Hal ini diyakini menjadikan Twitter sebagai sumber data yang berpotensi dan efisien mengenai Pemilihan Presiden Tahun 2019.

Pada penelitian sentimen sebelumnya terhadap Twitter @SBYudhoyono dengan menggunakan metode Naive Bayes dihasilkan akurasi 79,42% (Aliandu, 2013). Mengacu pada penelitian sebelumnya, penelitian ini bertujuan mengetahui sentimen analisis

Twitter capres Indonesia tahun 2019. Pengumpulan data dilakukan berdasarkan kata kunci berupa nama capres, dengan jumlah tertentu dan pengambilan sampel di tanggal tertentu. Selama pengumpulan data, peneliti mengambil atribut isi teks kemudian menambahkan atribut sentimen. Setelah itu tahap *preprocessing*, yaitu normalisasi data dengan cara mengubah kata tidak baku menjadi baku, menghilangkan angka, menghilangkan tanda baca dan simbol, stemming data dengan menghilangkan imbuhan pada setiap kata, dan penghapusan stopword (Husin Muhamad, 2017), kemudian masuk ke tahap sentimen menggunakan metode K-NN dan dihitung jumlah sentimen positif, negative dan netral.

2. METODE PENELITIAN

Metode yang dilakukan adalah metode observasi tidak langsung yaitu untuk memperoleh data dari subjek penelitian dengan mengambil data dari situs Twitter. Penelitian mengambil data *tweet* yang ada kemudian mengolah datanya untuk kemudian dijadikan subjek penelitian. Penelitian ini dilakukan dalam beberapa tahap, dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

3. TINJAUAN PUSTAKA

Sebelumnya telah banyak dilakukan penelitian tentang analisis sentimen, analisis sentimen kini digunakan untuk kepentingan politik, seperti pada penelitian Hadiyan Rasyadi tentang analisis sentimen pada pemilihan gubernur DKI Jakarta 2017, yang menghasilkan akurasi sebesar 60,60% menggunakan Naive Bayes. Penelitian ini menggunakan sampel data Twitter menggunakan Twitter API dari tanggal 4 sampai 19 April 2017. Data yang diambil menggunakan keyword @AhokDjarot dan @JktMajubersama. Model yang dibuat telah berhasil memprediksi setiap data pada data uji (Rasyadi, 2017).

Penelitian yang dilakukan oleh Nuvirta Monarizqa dkk tentang analisis sentimen Twitter sebagai pemberi rating menghasilkan nilai akurasi dari konfigurasi algoritma SVM. Dengan preprocessing yang diterapkan penelitian ini untuk analisis sentimen pada teks berupa *tweet* berbahasa Indonesia adalah sebesar 73.43%. Konfigurasi ini kemudian menjadi *core* aplikasi analisis sentimen. Aplikasi analisis sentimen dibangun dengan menggunakan Twitter API untuk mengambil data realtime, pustaka Lucene sebagai alat *stemming* pada *core* aplikasi, SVMlight, dan Java. Ketika konfigurasi diterapkan dalam aplikasi analisis sentimen yang dibuat, nilai akurasi ketika menggunakan kata kunci “Jokowi” sebesar 68%, “Prabowo” sebesar 56%, “kalimilk” sebesar 70% dan “sunmor” sebesar 74% (Antonius Rachmat C, 2016).

K-NN adalah metode untuk melakukan klasifikasi terhadap objek berdasarkan data training yang menggunakan jarak terdekat atau kemiripan pada objek tertentu. Menurut Deviyanto pada penelitiannya tentang analisis

sentimen pada pengguna Twitter menggunakan k-NN menghasilkan akurasi 67,2% (Deviyanto, 2018).

Penelitian ini bertujuan untuk menghasilkan sentiment terhadap dua calon presiden Indonesia tahun 2019 dengan menggunakan metode K-NN.

a. Analisis Sentimen

Analisis sentimen juga dikenal sebagai opinion mining atau emotion artificial intelligence adalah penggunaan pemrosesan bahasa alami, analisis teks, komputasi linguistik dan biometrik untuk mengidentifikasi, mengekstrak, menghitung dan mempelajari informasi subjektif secara sistematis (Rizal, 2017). Analisis Sentimen atau opinion mining mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan text mining yang bertujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu (Defri Rosdiyansyah, 2015).

Analisis sentimen adalah sebuah teknik untuk mendeteksi opini terhadap suatu subyek (misalnya individu, organisasi ataupun produk) dalam sebuah kumpulan data (Nuswaka, 2003). Analisis sentimen digunakan untuk memahami komentar yang diciptakan oleh pengguna internet dan menjelaskan bagaimana sebuah produk maupun brand diterima oleh mereka. (Cvijikj, 2013). Dari beberapa pendapat di atas, bisa diambil kesimpulan bahwa analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif atau negatif.

b. Twitter

Twitter didirikan pada bulan Maret 2006 oleh Jack Dorsey, dan situs jejaring sosialnya diluncurkan pada bulan juli. Sejak diluncurkan, Twitter telah menjadi salah satu dari sepuluh situs yang paling sering dikunjungi di internet, dan dijuluki dengan pesan singkat dari Internet (Rizal, 2017). Beberapa istilah dalam Twitter antara lain:

- 1) *Mention*, adalah menyebut atau memanggil pengguna Twitter lain dalam sebuah *tweet*. Mention dilakukan dengan

menuliskan '@' diikuti dengan nama pengguna lain.

- 2) *Hashtag*, digunakan untuk menandai sebuah topik pembicaraan di Twitter. Penulisan hashtag dimulai dengan tanda '#' diikuti dengan topik yang sedang dibahas. *Hashtag* biasa digunakan untuk meningkatkan visibilitas *tweet* pengguna.
- 3) *Emoticon*, adalah ekspresi wajah yang direpresentasikan dengan kombinasi antara huruf, tanda baca dan angka. Pengguna biasa menggunakan *emoticon* untuk mengekspresikan mood yang sedang mereka rasakan.
- 4) *Trending Topics*, Jika *hashtag* adalah cara untuk menandai sebuah topik pembicaraan di Twitter, maka *trending topics* adalah kumpulan dari topik pembicaraan yang sangat populer di Twitter.

c. Twitter API

API (*Application Programming Interface*) merupakan sekumpulan perintah, fungsi, dan protokol yang dapat digunakan dalam membangun perangkat lunak untuk sistem operasi tertentu, juga merupakan suatu dokumentasi yang terdiri dari antar muka, fungsi, kelas, struktur untuk membangun *software*. Seorang programmer akan lebih mudah dalam membongkar suatu *software* untuk kemudian dapat dikembangkan atau diintegrasikan dengan perangkat lunak yang lain melalui API. Ada beberapa jenis Twitter API :

- 1) Twitter REST API, Terdiri dari Twitter REST dan Twitter Search. Twitter REST memberikan *core* data dan *core* Twitter *objects*. Twitter search berfungsi untuk melakukan pencarian mengenai suatu *instance* objek Twitter maupun mencari *trend*.
- 2) Twitter Streaming API, API ini biasa digunakan untuk penggalian data karena melalui API ini informasi bisa didapatkan secara *realtime* dengan volume yang sangat tinggi.

d. Text Mining

Text mining adalah bidang interdisipliner yang mengacu pada pencarian informasi, pertambangan data, pembelajaran mesin, statistik, dan komputasi *linguistik*. Dikarenakan kebanyakan informasi (perkiraan umum mengatakan lebih dari 80%)

saat ini disimpan sebagai teks, text mining diyakini memiliki potensi nilai komersial tinggi. Text mining merupakan teknik yang digunakan untuk menangani masalah klasifikasi, *clustering*, *information extraction* dan *information retrieval* (Syaputra, 2016).

Text mining mempunyai 5 tahapan yaitu:

- 1) Tahapan *Tokenizing* adalah proses pemotongan string masukan berdasarkan tiap kata yang menyusunnya. Pada prinsipnya proses ini adalah memisahkan setiap kata yang menyusun suatu dokumen.
- 2) Tahapan *Filtering* adalah suatu proses dimana diambil sebagian data dari data tertentu, dan membuang data pada frekuensi yang lain.
- 3) Tahapan *Stemming* adalah proses pemetaan dan penguraian berbagai bentuk (*variants*) dari suatu kata menjadi bentuk kata dasarnya.
- 4) Tahapan *Tagging* adalah kata yang belum lama dilahirkan. Dahulu sebelum ada tagging, dunia informasi yang ada di internet berserakan dan tidak tersusun berdasarkan kategorinya. Hal itu bagaikan, perpustakaan tanpa ada pengurusnya atau pustakawan.
- 5) Tahapan *Analyzing* yaitu untuk mencari seberapa jauh keterhubungan antar kata-kata setiap dokumen.

e. Text Preprocessing

Data *tweet* yang telah diambil dari Twitter masih berupa data mentah maka dari itu dilakukan tahap *preprocessing* untuk mendapatkan data bersih agar dapat diproses ke tahap selanjutnya. Tahapan yang dilakukan adalah :

- 1) *Cleansing Data*, Dilakukan untuk mengurangi noise pada data *tweet*. Kata-kata yang tidak penting dihilangkan seperti URL, *hashtag* (#), *username* (@username), email, *emoticon* (:@, :*, :D), tanda baca seperti koma (,), titik (.) dan juga tanda baca lainnya.
- 2) *Case Folding*, merupakan tahapan yang mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf a sampai dengan z yang diterima.
- 3) *Tokenizing*, adalah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya. Pada dasarnya proses

tokenizing adalah pemenggalan kalimat menjadi kata.

- 4) *Stopword Removal*, adalah proses penghilangan kata-kata yang tidak berkontribusi banyak pada isi dokumen. Kata-kata yang termasuk ke dalam *stopword* dihilangkan karena memberikan pengaruh yang tidak baik dalam proses *text mining* seperti kata-kata "dan", "saya", "kamu", "dengan", "dia" dan lain-lain.
- 5) *Stemming*, adalah tahapan untuk membuat kata berimbuhan menjadi kata dasar sesuai dengan aturan bahasa Indonesia yang benar.
- 6) *Convert Negation*, dalam Bahasa Indonesia terdapat kata "tidak", "nggak", "tak", "kurang", "tanpa" yang disebut kata negasi yaitu kata yang dapat membalikan arti dari kata yang sebenarnya.

f. Klasifikasi

Klasifikasi adalah proses menemukan model dari sebuah data. Tujuan dari klasifikasi adalah untuk mengambil suatu keputusan dengan memprediksi suatu kasus berdasarkan hasil klasifikasi yang diperoleh. Dalam proses pengklasifikasian terdapat 2 proses yang dilakukan yaitu :

- 1) Proses *training* Pada proses ini dilakukan *training set* yang sudah diketahui label-labelnya untuk membangun model.
- 2) Proses *testing* Proses ini untuk mengetahui keakuratan model yang dibangun pada proses *training*, umumnya digunakan data yang disebut data *test set* untuk memprediksi label.

g. K-Nearest Neighbor (KNN)

Algoritma *K-Nearest Neighbor* (KNN) adalah metode yang digunakan untuk mengelompokkan objek berdasarkan contoh pelatihan terdekat di ruang fitur. *K-Nearest Neighbor* merupakan jenis yang paling dasar dari contoh *based learning* atau *lazy learning* juga termasuk kelompok *instance-based learning*. *K-Nearest Neighbor* dilakukan dengan mencari kelompok objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data *testing*. Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama (*data training*), diantaranya *Euclidean distance* dan

manhattan distance (*city block distance*) yang sering digunakan adalah *euclidean distance* seperti pada persamaan 1, yaitu :

$$d(x, y) = \sqrt{\sum_i^n (a_i - b_i)^2 + (a_2 - b_2)^2, \dots \dots (a_n - b_n)^2} \quad (1)$$

Dimana

a = $a_1, a_2, \dots \dots a_n$

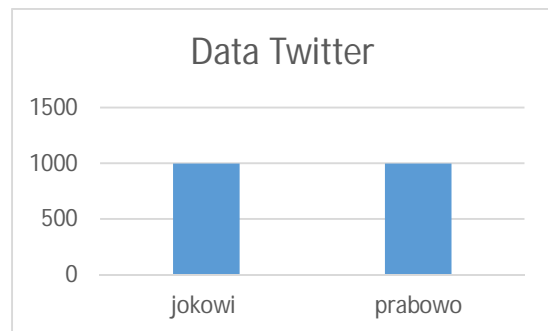
b = $b_1, b_2, \dots \dots b_n$ mewakili n nilai *atribut* dari dua *record*.

4. HASIL DAN PEMBAHASAN

Dalam melakukan sentimen capres kami melakukan beberapa langkah penelitian, diantaranya adalah :

a. *Crawling* data

Data *tweet* di ambil dengan metode *crawling* dari Twitter dengan *keyword* jokowi dan prabowo dan *tweet* dalam bahasa indonesia. Pengambilan data dilakukan pada bulan Desember 2019 dengan jumlah 1000 data seperti gambar 2 di bawah ini.

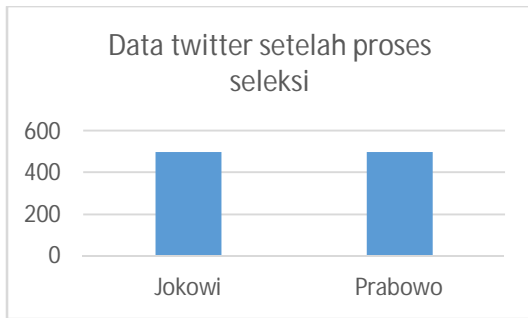


Gambar 2. Data Tweet

b. *Preprocessing* Data

Pada tahap *preprocessing*, dilakukan langkah sebagai berikut:

- 1) Seleksi Komentar, Pada tahap ini, dilakukan seleksi komentar yang mengandung kata jokowi dan prabowo. Pada tahap ini kami melakukan penghapusan retweet dan duplikasi. Setelah proses cleansing, dari 1000 data *tweet* jokowi menghasilkan 562 data dan prabowo 511 data. Kami hanya mengambil 500 data seperti terlihat pada gambar 3.



Gambar 3. Data Twitter setelah proses seleksi

- 2) *Cleansing* Data, dilakukan untuk mengurangi noise pada data *tweet*. Proses ini menghilangkan kata yang tidak penting seperti URL, hashtag (#), username (@username), email, emoticon (:@, :*, :D), tanda baca seperti koma (,), titik (.) dan juga tanda baca lainnya.
- 3) *Parsing*, proses memecah dokumen menjadi sebuah kata dengan melakukan analisa terhadap kumpulan kata dengan memisahkan kata tersebut dan menentukan struktur sintaksis dari tiap kata tersebut (Putro, 2011).
- 4) *Normalisasi Kalimat*, Bertujuan untuk menormalkan kalimat sehingga kalimat gaul menjadi normal (Adiyasa, 2011), sehingga bahasa gaul tersebut dapat dikenali sebagai bahasa yang sesuai dengan KBBI. Hal yang harus dilakukan dalam normalisasi kalimat adalah sebagai berikut:
 - Meregangkan tanda baca (*punctuation*) dan symbol selain alphabet.
 - Mengubah huruf menjadi kecil.
 - Normalisasi kata
 - Menghilangkan huruf yang berluang.
 - Menghilangkan emoticon

c. *Tokenisasi*

Setelah normalisasi kalimat, selanjutnya kalimat tersebut dipecah kedalam token-token menggunakan pembatas atau delimiter spasi. Token yang digunakan dalam penelitian ini yaitu: N-gram: token yang terdiri dari tiga kata setiap satu token, contohnya: Presiden Republik Indonesia.

d. *Penentuan class attribute*

Data Twitter yang sudah dilakukan Preprocessing kemudian ditentukan *class attribute*, ada 3 class yaitu, positif, negatif

dan netral. Seperti terlihat pada tabel 1 dibawah ini.

Tabel 1. *Penentuan class attribute*

	Class
Positif	1
Negatif	-1
Netral	0

e. *Load Dictionary*

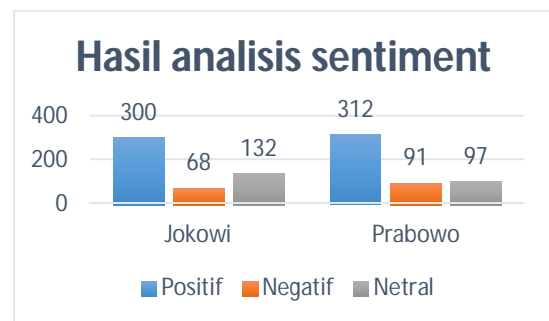
Setelah dilakukan tokenisasi dan ditentukan *class attribute*, langkah selanjutnya adalah *load dictionary*. Banyak jenis kamus yang dapat digunakan, misalnya kamus kata kunci sentimen positif dan kamus kata kunci sentimen negatif.

Berikut adalah contoh kamus dan isinya :

1. Kata positif : baik, hebat, jujur, cerdas.
2. Kata negatif : korupsi, jahat, jelek.
3. Kata negasi : jauh, bukan, siapa, kemana

f. *Determine Sentiment*

Proses ini menentukan sentimen suatu kalimat opini, penentuan dilakukan dengan penjumlahan n skor polaritas kata opini p yang mengomentari fitur f. Skor polaritas suatu kata opini p akan bernilai 1 jika kata tersebut adalah kata opini positif, bernilai -1 jika kata tersebut adalah kata opini negative dan bernilai 0 jika kata tersebut adalah kata netral. Dari penelitian yang dilakukan didapatkan hasil sebagai berikut:



Gambar 4. Hasil analisis sentiment

g. *Pengujian pengaruh nilai k*

Pengujian pengaruh nilai *k* dilakukan untuk mengetahui nilai *k* optimal untuk melakukan proses klasifikasi KNN terhadap hasil akurasi sistem. Pada pengujian ini menggunakan nilai *k* yang bervariasi. Setiap nilai *k* yang digunakan, akan menjadi parameter dari pengujian, sehingga dapat diketahui pengaruh nilai *k* tersebut terhadap

akurasi sistem. Pada pengujian pengaruh nilai k , nilai konstanta α dan β yang digunakan masing-masing 0,5. Uji coba perhitungan nilai k dilakukan sebanyak 10 kali sampai sejumlah data latih yaitu 500. Hasil dari pengujian pengaruh nilai k terdapat pada Tabel 2.

Tabel 2. Pengujian Nilai k

No	Nilai k	Accuracy	Precision	Recall	F-Measure
1	2	71%	63,76%	100%	75,52%
2	3	81,83%	73,28%	100%	82,90%
3	4	72,59%	65,51%	100%	78,22%
4	5	72,66%	64,82%	100%	78,11%
5	6	72,53%	65,51%	100%	77,95%
6	7	71,16%	64,82%	100%	79,51%
7	8	74,33%	66,21%	100%	76,56%
8	9	74,56%	66,93%	100%	74,43%
9	10	68,16%	62,85%	100%	72,79%
10	15	66,83%	59,45%	100%	72,77%

Berdasarkan tabel 2, pengujian nilai k dapat dilihat bahwa nilai k terbaik untuk melakukan proses klasifikasi KNN adalah $k=3$, dengan tingkat akurasi mencapai 81,83%, precision mencapai 73,28%, recall mencapai 100%, dan f-measure mencapai 82,90%

h. Pengujian akurasi

Pada pengujian akurasi di perbandingkan data latih dan data uji yang digunakan yaitu 70% dan 30%. Jumlah data latih dan data uji yang digunakan yaitu 350 dan 150.

Nilai k yang digunakan untuk melakukan pengujian akurasi sistem adalah $k=3$, dan nilai konstanta yang digunakan $\alpha=0,8$ dan $\beta=0,2$. *Confusion Matrix* dan hasil dari pengujian akurasi sistem dapat dilihat pada Tabel 3 dan 4.

Tabel 3. Confusion matrix pengujian akurasi

	Tekstual	Non-Tekstual	Gabungan
True positif	55	60	58
False positif	16	48	18
True negatif	44	12	42
False negatif	5	0	2

Tabel 4. Hasil pengujian akurasi

	Tekstual	Non-Tekstual	Gabungan
Accuracy	82,50%	60%	83,33%

Precision	77,46%	55,55%	76,31%
Recall	91,66%	100%	96,66%
F-Measure	83,96%	71,42%	85,29%

5. KESIMPULAN DAN SARAN

a. Kesimpulan

Dari penelitian yang telah dilakukan, dapat ditarik sekimpulan bahwa:

- 1) analisis sentiment dapat digunakan untuk mengetahui opini masyarakat di Twitter terhadap capres Indonesia tahun 2019.
- 2) Hasil analisis sentimen, memperlihatkan bahwa masyarakat memiliki opini baik terhadap kedua capres yaitu opini positif 300, negatif 68, netral 132 untuk jokowi dan positif 312, negatif 91, netral 97 untuk prabowo.
- 3) Penggabungan pembobotan tekstual dan pembobotan nontekstual dapat meningkatkan akurasi. Akurasi ketika menggunakan pembobotan tekstual adalah 82,50%, ketika menggunakan nontekstual adalah 60%, sedangkan penggabungan keduanya adalah 83,33%.

b. Saran

Setelah melakukan penelitian, kami menyadari masih banyak kekurangan dalam penelitian ini, seperti kurangnya data dan metode yang digunakan sehingga menghasilkan akurasi yang didapat. Untuk penelitian selanjutnya perlu dicoba dikembangkan dengan menggunakan data yang lebih banyak dan Real Time. Perlu dikembangkan metode lain yang mampu meningkatkan akurasi dalam analisis sentiment.

6. REFERENSI

Adiyasa, N. (2011, 8 2). *Analisis Sentimen Pada Opini Berbahasa Indonesia Menggunakan Pendekatan Lexicon-Based*. Retrieved 12 1, 2018, from <https://adiyasan.wordpress.com/2013/02/08/sentiment-analysis-menggunakan-pendekatan-lexicon-based/>

Aliandu, P. (2013). Twitter Used by Indonesian President : An Sentiment Analysis of Timenline. *ISICO*, 713-717.

Antonius Rachmat C, Y. L. (2016). *Klasifikasi Sentimen Komentar Politik*

- dari Facebook Page Menggunakan Naive Bayes. *JUISI*, 2(2), 1-9.
- Cvijikj, I. P. (2013). *Online engagement factors on Facebook brand pages*. New York: Springer Vienna Publisher.
- Defri Rosdiyansyah, S. A. (2015). Analisis Sentimen Twitter Menggunakan Metode K-Nearest Neighbor dan Pendekatan Lexicon. *UIN SUSKA*, 1-6.
- Deviyanto, A. (2018). *Sentimen Analisis Tweet Pornografi Kaum Homoseksual Sentimen Analisis Tweet Pornografi Kaum Homoseksual Sentimen Analisis Tweet Pornografi Kaum Homoseksual*. Yogyakarta: UIN Kalijaga.
- Dubey, S. A. (2017). Clustering and Sentiment Analysis on Twitter Data. *International Conference on Telecommunication and Networks (TEL-NET)*, 5.
- Hamilton, H. (2018, 7 9). *Confusion Matrix*. Retrieved 12 1, 2018, from http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html
- Husin Muhamad, C. A. (2017). OPTIMASI NAÏVE BAYES CLASSIFIER DENGAN MENGGUNAKAN PARTICLE SWARM OPTIMIZATION PADA DATA IRIS. *JTIK*, 4(3), 180-184.
- Liu, B. H. (2005). Opinion observer: analyzing and comparing opinions on. *international conference on World Wide Web*, (pp. 342-451). ACM.
- Manongga, A. M. (2018). Sentimen Analisis Tweet Pornografi Kaum Homoseksual Indonesia Di Twitter Dengan Naive Bayes. *SIMETRIS*, 313-318.
- Nuswaka, T. &. (2003). *Sentiment analysis: capturing favorability using natural*. New York: ACM.
- Putro, M. Y. (2011). *Analisis Sentimen pada Dokumen Berbahasa Indonesia*. Jakarta: Binus.
- Rasyadi, M. H. (2017). *Analisis Sentimen Pada Twitter Menggunakan Metode Naïve Bayes (Studi Kasus Pemilihan Gubernur Dki Jakarta 2017)*. Bogor: Institut Pertanian Bogor.
- Rizal, M. (2017). *Analisis Sentimen Pengguna Twitter Terhadap Objek Pariwisata Di Indonesia Menggunakan Algoritma Pengolahan Deep Natural Language Dari Ibm Insights Untuk Twitter*. Makasar: UIN Alauddin Makasar.
- Sismono, H. (2005). *Pengantar Logika Informatika*. Yogyakarta: Andi Offset.
- Syaputra, R. &. (2016). Analisis Sentimen Pada Sosial Media Twitter Terhadap Politik Di Indonesia Menggunakan Text Mining Dengan Metode Naive Bayes Classifier. *IJCCS*, 1-8.