

ANALISIS SENTIMEN PASAR OTOMOTIF MOBIL: *TWEET* TWITTER MENGUNAKAN NAÏVE BAYES

Deden Rustiana

Program Studi Sistem Komputer
Perguruan Tinggi Raharja
Email: deden.rustiana@raharja.info

Nina Rahayu

Magister Teknologi Informatika
Perguruan Tinggi Raharja
Email: nina.rahayu@raharja.info

ABSTRAK

Sekarang ini orang tidak lepas dari internet, dari bangun pagi kebanyakan orang sudah disibukan untuk melakukan *update* status, *browsing* berita dan berkirim email. Dalam jejaring social seperti facebook atau yang lain nya kebanyakan orang menceritakan suasana yang dialaminya ke tempat jejaring social seperti menceritakan sedang liburan, makan, jalan-jalan pada facebook, twitter ataupun yang lainnya, sehingga data tersebut sekarang ini membentuk menjadi data yang sangat besar atau dikenal dengan "*Big Data*". Media sosial seperti twitter banyak dipakai untuk mencurahkan isi hati pemakai nya, sehingga data tersebut dapat menggambarkan sentimen. Naïve bayes merupakan metode klasifikasi analisi sentiment yang banyak digunakan karena ketangguhan, sederhana dan akurasi yang tinggi dalam mengklasifikasi data. Pada penelitian ini menganalisis data dari twitter dengan pencarian merk mobil yang terlaris berdasarkan data dari GAIKINDO (Gabungan Industri Kendaraan Bermotor Indonesia). Dengan menggunakan tiga kelas maka dapat menghasilkan nilai akurasi sebesar 93%, dan tingkat *precision* positif sebesar 90%, *precision* negatif sebesar 90% dan sentimen yang netral 100% sehingga naïve bayes mampu menganalisis dengan akurasi yang tinggi.

Kata kunci: *big data*, jejaring social, twitter, naïve bayes dan merk mobil.

ABSTRACT

Now these people can not be separated from the internet, from waking up in the morning most people are too preoccupied to do status updates, news browsing and emailing. In social networks like facebook or else his most telling atmosphere that happened to a social network such as communicating on vacation, eating, sightseeing on facebook, twitter or others, so that from these data is now shaping up to be a very large data or known as "Big Data". Social media such as twitter is widely used to confide its user, so that data can describe the sentiment. Naïve Bayes classification method of analysis is a sentiment that is widely used for toughness, simple and high accuracy in classifying data. In this study analyzed data from twitter by searching the best-selling brands of cars are based on data from GAIKINDO (Association of Indonesian Automotive Industries). By using a three-class accuracy, it can generate value by 93%, and the level of precision of 90% positive, 90% negative precision and neutral sentiment 100% so naïve Bayes able to analyze with high accuracy.

Keywords: *big data*, social networking, twitter, naïve Bayes and brands of cars.

1. PENDAHULUAN

Munculnya teknologi baru, perangkat, dan komunikasi seperti situs jejaring sosial mengakibatkan jumlah data yang dihasilkan oleh manusia tumbuh cepat setiap tahun. Jumlah data yang dihasilkan dari awal tahun 2003 sampai akhir tahun 2003 berjumlah 5 miliar *giga bytes*. Jika data disimpan dalam bentuk disk mungkin akan memenuhi seluruh lapangan sepak bola. Jumlah yang sama diciptakan pada setiap dua hari pada tahun 2011, dan di setiap sepuluh menit pada tahun 2013. Angka ini masih akan tumbuh pesat. Meskipun semua informasi yang diproduksi ini bermakna dan dapat berguna ketika diproses, tetapi masih banyak yang belum dimanfaatkan. Microblogging menjadi populer sebagai alat komunikasi antara pengguna internet. Pada pertengahan tahun 2010 Twitter memiliki pengguna lebih dari 106 juta pengguna diseluruh dunia dan terus meningkat setiap harinya sebanyak 300.000 pengguna dan Twitter setiap harinya mendapatkan lebih dari 3 juta request. Dari angka tersebut Indonesia menjadi negara yang menduduki peringkat 8 dalam mengakses situs Twitter. Twitter menerima tweet dari pengguna sebanyak

55 juta pesan setiap harinya [1]. Dengan demikian, kita dapat memanfaatkan data tweet tersebut untuk kepentingan tertentu seperti untuk mengetahui sentimen pada merk mobil sehingga kita bisa menggambarkan sentimen konsumen terhadap merk mobil. Text mining (penambangan teks) adalah penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, sesuatu yang tidak diketahui sebelumnya atau menemukan kembali informasi yang tersirat secara implisit, yang berasal dari informasi yang diekstrak secara otomatis dari sumber-sumber data teks yang berbedabeda [2]. *Text mining* merupakan teknik yang digunakan untuk menangani masalah klasifikasi, *clustering*, *information extraction* dan *information retrieval* [3].

Untuk memperluas pasar dan mengelola hubungan dengan konsumen Kami membutuhkan informasi sebagai pendukung keputusan salah satu informasi yang dapat digunakan adalah informasi dari media sosial. Twitter adalah raksasa utama media sosial, dengan jutaan pengguna dan memberikan wawasan yang unik, data yang ada didalam Twitter bisa membantu mengaetaahui sentimen pasar secara objektif. Dalam penelitian ini akan dikembangkan aplikasi dengan menggunakan teknik klasifikasi *naive bayesian classifiers*. Pendekatan ini merupakan pendekatan mengacu pada teorema Bayes yang merupakan prinsip peluang statistika untuk mengkombinasikan pengetahuan sebelumnya dengan pengetahuan baru. [2]. Klasifikasi yang digunakan menggunakan tiga kelas yaitu sentiment positif, negatif dan netral. Sehingga diharapkan dapat menghasilkan rancangan *prototype analytical* data dengan metode *naive bayes*.

1.1 Landasan Teori

1.1.1 Text Mining

Text mining (penambangan teks) adalah penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, sesuatu yang tidak diketahui sebelumnya atau menemukan kembali informasi yang tersirat secara implisit, yang berasal dari informasi yang diekstrak secara otomatis dari sumber-sumber data teks yang berbedabeda [2].

1.1.2 Text Preprocessing

Tahap *text preprocessing* adalah tahap awal dari text mining. Tahap ini mencakup semua rutinitas, dan proses untuk mempersiapkan data yang akan digunakan pada operasi *knowledge discovery* sistem *text mining* [2]. Tindakan yang dilakukan pada tahap ini adalah *toLowerCase*, yaitu mengubah semua karakter huruf menjadi huruf kecil dan *Tokenizing* yaitu proses penguraian deskripsi yang semula berupa kalimat-kalimat menjadi kata-kata dan menghilangkan delimiter seperti tanda titik (.), koma (,), spasi dan karakter angka yang ada pada kata tersebut [4].

1.1.3 Feature Selection

Tahap seleksi fitur (*feature selection*) bertujuan untuk mengurangi dimensi dari suatu kumpulan teks, atau dengan kata lain menghapus kata-kata yang dianggap tidak penting atau tidak menggambarkan isi dokumen sehingga proses pengklasifikasian lebih efektif dan akurat [2]. Pada tahap ini tindakan yang dilakukan adalah menghilangkan stopword (*stopword removal*) dan *stemming* terhadap kata yang berimbuhan [3].

1.1.4 Sentiment Analysis

Sentiment analysis 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 [5].

Tugas dasar dalam analisis sentimen adalah mengelompokkan teks yang ada dalam sebuah kalimat atau dokumen kemudian menentukan pendapat yang dikemukakan dalam kalimat atau dokumen tersebut apakah bersifat positif, negatif atau netral [6]. *Sentiment analysis* juga dapat menyatakan perasaan emosional sedih, gembira, atau marah.

1.1.5 Twitter API

Twitter API (*Application Programming Interface*) merupakan sejumlah fungsi yang dapat digunakan pengembang perangkat lunak untuk mengolah data saat membangun perangkat lunak. Twitter API menyediakan beberapa fungsi untuk melakukan suatu tugas tertentu, sehingga pengembang

perangkat lunak hanya memanggil fungsi tersebut di dalam perangkat lunak yang dibangun. Twitter API menggunakan arsitektur REST (*Representational State Transfer*) sehingga Twitter API dapat digunakan pada format data yang beragam seperti XML maupun JSON. Twitter API terdiri atas Twitter Search API dan Twitter Streaming API. Perbedaan keduanya yaitu, Twitter Search API menitikberatkan fungsi pencarian ke masa lampau sedangkan Twitter Streaming API menitikberatkan fungsi pencarian ke masa yang akan datang [7].

1.1.6 Naïve Bayes Classifier

Algoritma *naive bayes classifier* merupakan algoritma yang digunakan untuk mencari nilai probabilitas tertinggi untuk mengklasifikasi data uji pada kategori yang paling tepat [2]. Dalam penelitian ini yang menjadi data uji adalah dokumen *tweets*. Ada dua tahap pada klasifikasi dokumen. Tahap pertama adalah pelatihan terhadap dokumen yang sudah diketahui kategorinya. Sedangkan tahap kedua adalah proses klasifikasi dokumen yang belum diketahui kategorinya.

2. METODOLOGI PENELITIAN

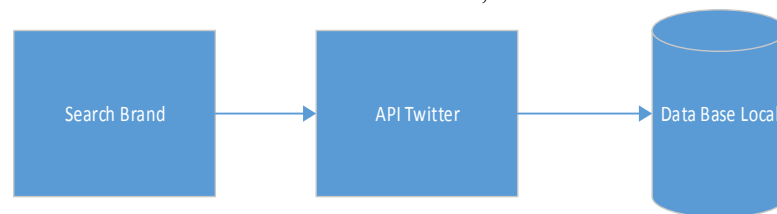
2.1 Pengumpulan Data

Pengambilan data dilakukan berdasarkan pencarian atas term objek pada aplikasi yang terhubung pada Twitter API. Hasil query berupa tweet kotor baik untuk data pelatihan maupun hasil query pengguna mengalami *preprocessing* yang sama.

Dalam penelitian ini kata kunci yang digunakan yaitu 10 brand mobil berdasarkan jumlah penjualan terbanyak pada tahun 2014. Data yang digunakan diambil dari situs GAIKINDO (Gabungan Industri Kendaraan Bermotor Indonesia), Lihat pada table 1.

Tabel 1. Kumpulan data brand⁸¹

<i>Brand</i>	<i>Sales</i>
Toyota	399,119
Daihatsu	185,226
Honda	159,147
Suzuki	154,923
Misubishi	141,923
Nissan	33,789
Hino	28,493
Isuzu	28,278
Datsun	20,520
Ford	12,008

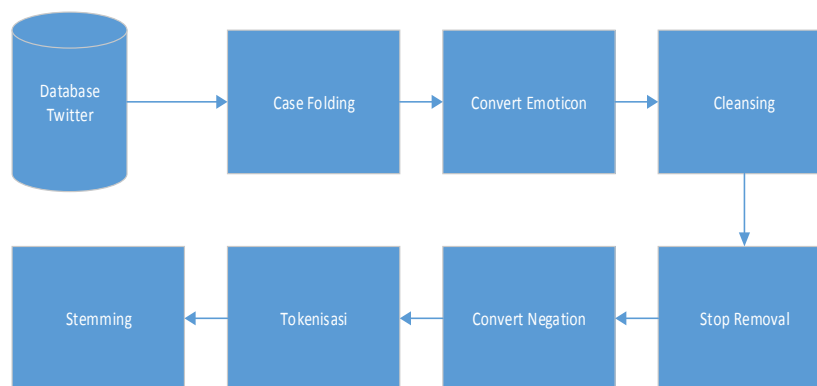


Gambar 1. Proses Pengumpulan Data

Pada tahap pengumpulan data pertama dilakukan pencarian brand mobil menggunakan aplikasi yang terhubung dengan API Twitter dengan pengaturan twitter yang berbahasa Indonesia. Langkah selanjutnya data diinput ke dalam *data base local*.

2.2 Preprocessing

Tujuan dilakukannya preprocessing dokumen adalah untuk menghilangkan *noise*, menyeragamkan bentuk kata dan mengurangi volume kata. Tahapan yang dilakukan dari dokumen preprocessing dapat dilihat pada Gambar 2.



Gambar 2. Proses Preprocessing Dokumen

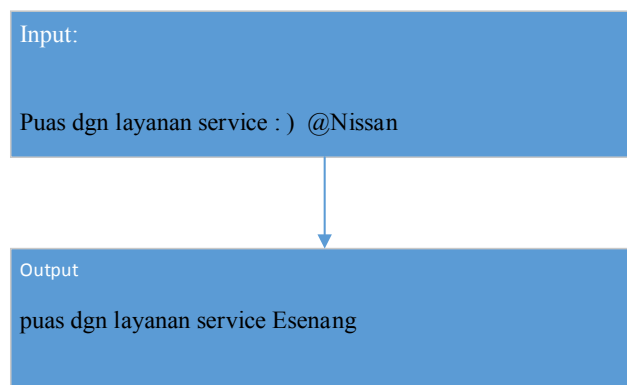
Pada tahap preprocessing terdiri dari proses Download data dari twitter, *case folding*, *convert emoticon*, *cleansing*, *stemming*, *tokenisasi*, *convert negation* dan *stop removal*.

1) *Case Folding*

Pada tahap ini, semua huruf akan diubah menjadi lowercase atau huruf kecil. Berikut merupakan langkah-langkah *case folding* dalam contoh salah satu contoh *tweets*: “Puas dgn layanan service @Nissan” dirubah menjadi “puas dgn layanan service @nissan”.

2) *Convert Emoticon*

Pada tahap ini, kumpulan tweets yang terdapat *emoticon* (*emotion icon*) akan dikonversikan ke dalam string yang bersesuaian. Namun, tidak semua akan diimplementasikan, karena tidak semua *emoticon* sering digunakan oleh pengguna Twitter.

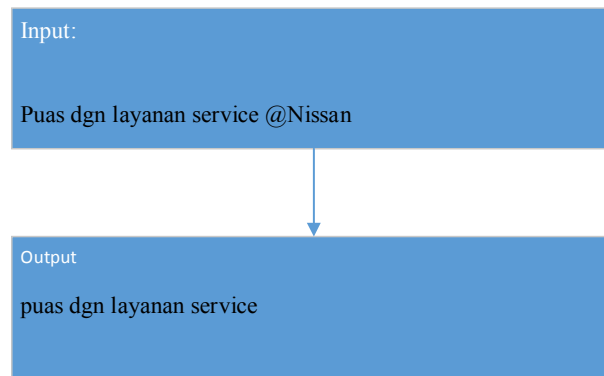


Gambar 3. Contoh Convert Emoticon

3) *Cleaning*

Ada beberapa komponen khas yang biasa ada di tweet yakni, username, URL (*Uniform Resource Locator*), dan “RT” (tanda retweet). Karena *username*, URL, dan “RT” tidak memiliki pengaruh apapun terhadap nilai sentimen, maka ketiga komponen di atas akan dibuang. Komponen *username* diidentifikasi dengan kemunculan karakter '@'. Selain *username*, karakter '@' biasa juga digunakan untuk pemanggilan suatu tempat seperti @FloatingMarket. Namun nama tempat tersebut tidak memiliki pengaruh pada analisis sentimen sehingga nama tempat pun harus dihapus. Pada komponen URL dikenali melalui ekspresi regular (seperti http, www). Berikut langkah-langkah pada tahap normalisasi fitur:

- a. Kata yang digunakan hasil dari *case folding*.
- b. Hasil dari *case folding* akan diperiksa apakah terdapat username, URL, dan RT.
- c. Jika terdapat username, URL, dan RT maka akan dihilangkan



Gambar 4. Contoh Cleansing

4) *Stopword Removal*

Pada tahap ini, kumpulan tweets yang telah melewati tahap cleaning dihilangkan karakter, tanda baca, serta kata-kata umum yang tidak memiliki makna atau informasi yang dibutuhkan.

5) *Convert Negation*

Pada tahap ini, setiap tweets yang mengandung kata-kata yang bersifat negasi akan diubah nilai sentimennya. Kata-kata yang bersifat negasi seperti “bukan”, “tidak”, “enggak”, “ga”, “jangan”, “nggak”, “tak”, dan “gak”. *Convert negation* dilakukan jika terdapat kata negasi sebelum kata yang bernilai positif, maka kata tersebut akan diubah nilainya menjadi negatif dan begitupun sebaliknya. Langkah-langkah pada tahap convert negation adalah sebagai berikut:

- Kata yang digunakan adalah hasil dari *stopword removal*.
- Setiap kata akan diperiksa dan dibandingkan dengan kumpulan katakata yang bersifat negasi pada database.
- Jika setelah kata yang bersifat negasi terdapat kata yang termasuk sentimen positif, maka sentimen tersebut akan diubah menjadi negatif.
- Jika setelah kata yang bersifat negasi terdapat kata yang termasuk sentimen negatif, maka sentimen tersebut akan diubah menjadi positif.

6) *Tokenisasi*

Pada tahap ini akan dilakukan pengecekan tweets dari karakter pertama sampai karakter terakhir. Apabila karakter ke-i bukan tanda pemisah kata seperti titik(.), koma(,), spasi dan tanda pemisah lainnya, maka akan digabungkan dengan karakter selanjutnya. Langkah-langkah pada tahap tokenizing adalah sebagai berikut:

- Kata yang digunakan adalah hasil dari convert negation.
- Memotong setiap kata dalam teks berdasarkan pemisah kata seperti titik(.), koma(,), dan spasi.
- Bagian yang termasuk dalam daftar emoticon tidak dibuang.

7) *Stemming*

Pada tahap ini mengubah token yang berimbuhan menjadi kata dasar, dengan menghilangkan semua imbuhan yang ada pada token tersebut.

2.3 Metode Klasifikasi Naïve Bayes Classifier (NBC)

Dalam penelitian ini label klas yang di gunakan yaitu positif, negative dan netral. Setiap tweet yang sudah melewati hasil preprocessing di cek arti kata yang mengandung kata negative, positif atau netral dengan prioritas menjadi 0.33 pada setiap kelas. Kelas yang memiliki skor terbanyak atau kemunculan terbanyak pada setiap tweet maka tweet tersebut dianggap menjadi kelas tersebut.

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Metode Naïve Bayes Classifier

Sebagai data uji adalah sejumlah data yang ada pada file data dipotong dari record paling akhir untuk masing-masing data opini positif dan opini negatif. Data latih diperoleh dari penggabungan file data opini positif dan file data opini negatif sisa dari penggunaan untuk data uji. Dibentuk enam variasi data masing-masing untuk file data positif dan negatif dengan menukar susunan record, dengan demikian akan terdapat enam variasi data pengujian untuk tiap proporsi data uji.

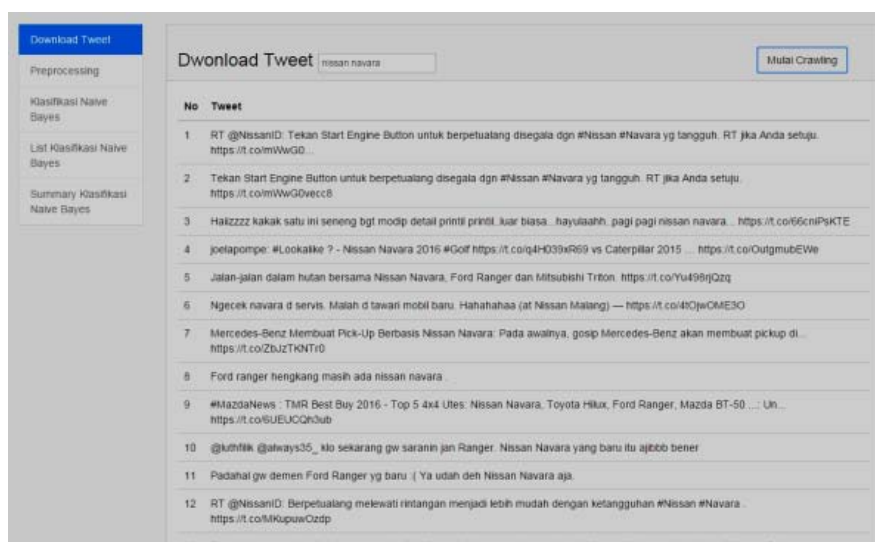
3.2 Rancangan Sistem

Pada penelitian ini penulis membuat aplikasi untuk membuat analisis sentiment secara otomatis, dimulai pengambilan data dari twitter, preprocessing data tweet melalui langkah case folding, convert emoticon, cleansing, stopword removal, convert negation, tokenisasi dan yang terakhir adalah proses stemming. Sedangkan pada tahap selanjut nya yaitu tahap klasifikasi dengan metode naive bayes. Hasil dari klasifikasi di tampilkan pada halaman list klasifikasi dan *summary* klasifikasi. Berikut merupakan *prototype* atau rancangan tampilan aplikasinya.

Implementasi merupakan tahap dimana sistem siap dioperasikan pada tahap yang sebenarnya, sehingga akan diketahui apakah sistem yang telah dibuat benar-benar sesuai dengan yang direncanakan. Pada implementasi perangkat lunak ini akan dijelaskan bagaimana program sistem ini bekerja, dengan memberikan tampilan sistem atau aplikasi yang dibuat.

3.2.1 Tampilan Download Tweet

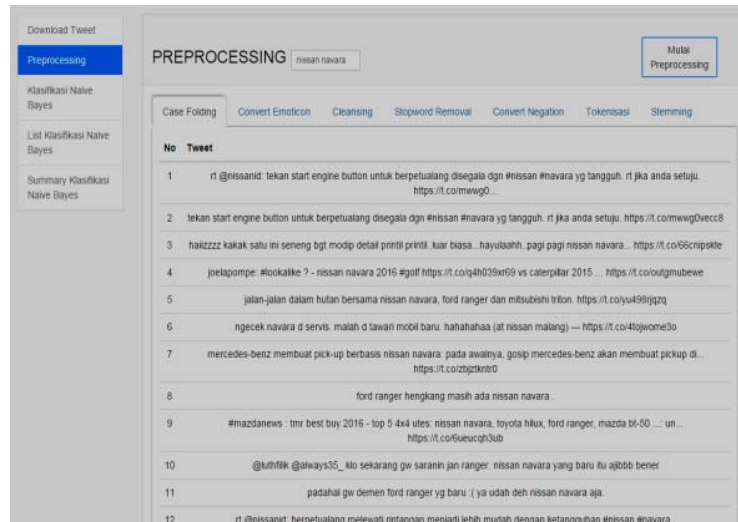
Tampilan ini digunakan untuk mengunduh data tweet dari tweeter melalui API, pada tools ini user tinggal mengetik keyword yang akan dijadikan analisis, pada penelitian ini keyword yang dicari adalah merk mobil. Dan program ini sudah terkoneksi langsung dengan data twitter melalui API Twitter.



Gambar 5. Tampilan Antarmuka Download Tweet

3.2.2 Tampilan Preprocessing

Tampilan preprocessing adalah fitur yang digunakan untuk membersihkan data *tweet* yang masih kotor. Pada tampilan ini pengguna memilih keyword yang sudah diunduh, setelah itu mengklik tombol “Mulai Preprocessing” maka tools akan memproses. Tahapan proses didalam halaman ini terdapat 7 proses dan di tampilkan didalam tab. Tahap yang pertama adalah “case folding” yang merupakan proses merubah semua huruf menjadi huruf kecil. Tab selanjutnya adalah “Convert Emoticon” yaitu proses mengubah symbol emoticon menjadi teks, tab *Cleansing* adalah proses menghapus symbol yang khas dari twitter seperti username, URL (*Uniform Resource Locator*), dan “RT” (tanda retweet). Tab “Stop Removal” adalah tampilan hasil dari proses menghilangkan karakter, tanda baca, serta kata-kata umum yang tidak memiliki makna atau informasi yang dibutuhkan. Dan untuk tab “Convert Negation”, “Tokenisasi” dan “Stemming” adalah hasil dari proses masih fungsi.



Gambar 6. Tampilan Antarmuka Preprocessing Text

3.2.3 Tampilan Klasifikasi Naïve Bayes

Merupakan tampilan hasil klasifikasi menggunakan naïve bayes didalam gambar 13 terlihat program memberikan label secara otomatis yang berisi lebal positif dan netral.

No	Tweet Kotor	Tweet Bersih	Positif	Negatif	Netral	Kategori
1	#Berita: Nissan Mulai Buka Setelubng Generasi Penerus Sportcar Terbarunya https://t.co/HoYbEYXSKP	nissan buka setelubng generasi terus sportcar baru	0.5	0.25	0.25	Positif
2	Nissan Mulai Buka Setelubng Generasi Penerus Sportcar Terbarunya https://t.co/ABeXQvGDBt #STMkerasboss #MentionsFor #DoraemonPic	nissan buka setelubng generasi terus sportcar baru	0.5	0.25	0.25	Positif
3	Nissan Mulai Buka Setelubng Generasi Penerus Sportcar Terbarunya (okezone) https://t.co/RbTGNOFHk	nissan buka setelubng generasi terus sportcar baru okezone	0.5	0.25	0.25	Positif
4	RT @CarsandCubs: Nissan GTR https://t.co/BSQzSCZEX6	nissan gtr	0.3333333333333333	0.3333333333333333	0.3333333333333333	Netral
5	https://t.co/19q5r1w4Gq: Nissan Mulai Buka Setelubng Generasi Penerus Sportcar Terbarunya: Generasi... https://t.co/78FHHGTOIF	nissan buka setelubng generasi terus sportcar baru generasi	0.5	0.25	0.25	Positif
6	Nissan Mulai Buka Setelubng Generasi Penerus Sportcar Terbarunya	nissan buka setelubng generasi terus sportcar baru	0.5	0.25	0.25	Positif

Gambar 7. Tampilan Antarmuka Klasifikasi Naïve Bayes

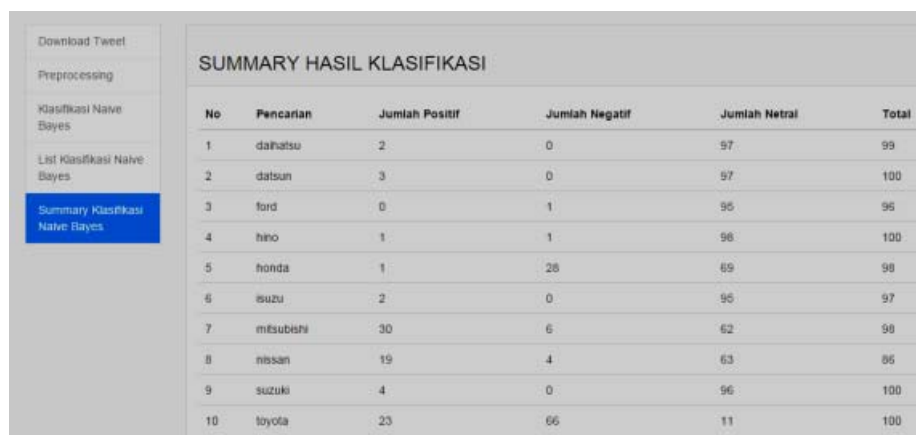
3.2.4 Tampilan List Klasifikasi

No	Tweet Kotor	Tweet Bersih	Kategori	Pencarian
1	@trmbodel Tim Hino Sugawara Lewati Etape ke-4 dengan Mulus https://t.co/GEzYaed5E #akinet	tim hino sugawara lewat etape mulus	netral	hino
2	Ford GT vs Lamborghini Murcielago https://t.co/CYUQ2U3L1Yh https://t.co/IMOCFTQ2IA	ford gt vs lamborghini murcielago	netral	ford
3	kot yo pun nak gotak hino kek den, gotak sopan sikit, kau tu pompuan malah oi https://t.co/8hDomoDHm	kot yo nak gotak hino kek den gotak sopan sikit kau tu pompu mi o	netral	hino
4	Bulan Depan Suzuki Luncurkan Penantang Honda BR-V https://t.co/SH2wJHBTg3	suzuk luncur tantang honda br	netral	honda
5	RT @luxbucketlat: Nissan GT-R • https://t.co/HGogtFuQ6	nissan gt	netral	nissan
6	1997 Toyota Tarago ABS Module #C197 https://t.co/ZabFC3RHbz https://t.co/TKGMBGHp	toyota tarago abs module	positif	toyota
7	for Nissan Safari Patrol Civilian Bus 4 2L, TD42 HT 18 Turbo Turbocharger TPM https://t.co/3yBxUJH4t https://t.co/2CG2RG8t	for nissan safari patrol civilian bus 4l 18 turbo turbocharger tpm	netral	nissan
8	RT @RajaPulauTidung: Sewa mobil di Bali - Mobil Alphard, Isuzu eif, Toyota Hiace, Avanza, APV, Inova dll https://t.co/Q27eA6Z5H https://t.co/...	sewa mobil bal mobil alphard isuzu eif toyota hiace avanza apv inova dll	netral	isuzu
9	1000hp #Mitsubishi #EVO vs 1300hp #CORVETTE! #Video by https://t.co/mQ5gm9Bedm: (https://t.co/2TALeWegOxA) https://t.co/YH4EzY2Gh	hp vs hp ty	netral	mitsubishi
10	RT @infoikanOS: Mobil Murah & Berkualitas	mobil murah amp kualitas op jta	netral	datsun

Gambar 8. Tampilan List Klasifikasi

3.2.5 Tampilan Summary Klasifikasi

Setelah data diklasifikasi menjadi tiga kelas, langkah selanjutnya yaitu meringkas berdasarkan merk mobil.



No	Pencarian	Jumlah Positif	Jumlah Negatif	Jumlah Netral	Total
1	daihatsu	2	0	97	99
2	datsun	3	0	97	100
3	ford	0	1	95	96
4	hino	1	1	98	100
5	honda	1	28	69	98
6	isuzu	2	0	96	97
7	mitsubishi	30	6	62	98
8	nisson	19	4	63	86
9	suzuki	4	0	96	100
10	toyota	23	66	11	100

Gambar 9. Tampilan Summary Klasifikasi

4. KESIMPULAN

Berdasarkan analisis dan pengujian yang dilakukan pada bab sebelumnya, maka kesimpulan yang dapat diambil adalah sebagai berikut :

- 1) Aplikasi ini dapat mendownload data dari twitter melalui *api streaming* dengan pengaturan data Bahasa indonesia.
- 2) Proses klasifikasi semakin akurat jika data latih yang digunakan dalam pembelajaran berjumlah banyak, akan tetapi dapat juga mengurangi keakuratan jika kata-kata yang terdapat pada Tweet tersebut mengalami bias atau bermakna ganda.
- 3) Implementasi naïve bayes pada penelitian ini menghasilkan nilai akurasi 93%, serta persisi sentiment positif 90%, persisi sentiment negative 90% dan persisi sentiment netral sebesar 100%.

DAFTAR PUSTAKA

- [1] Hepburn, A, Infographic: Twitter Statistics, Facts & Figures, <http://www.digitalbuzzblog.com/infographic-twitter-statistics-factsfigures/> Di akses tanggal 12 November 2016, 2010.
- [2] Feldman, R & Sanger, J., 2007. The Text Mining Handbook : Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press : New York.
- [3] Berry, M.W. & Kogan, J., 2010. Text Mining Application and theory. WILEY : United Kingdom.
- [4] Weiss, S.M., et al., 2005. Text Mining : Predictive Methods fo Analyzing Unstructured Information. Springer : New York.
- [5] Liu, Bing. Sentiment Analysis And Opinion Mining. Chicago: Morgan & Claypool Publisher. [http://www.dcc.ufjf.br/~valeriab/DTMSentimentAnalysisAndOpinionMini ng-BingLiu.pdf](http://www.dcc.ufjf.br/~valeriab/DTMSentimentAnalysisAndOpinionMining-BingLiu.pdf). Diakses tanggal 20 Nopember 2016, 2012.
- [6] Sentiment Analysis, Hard But Worth It!. [Online] . Tersedia di: http://www.customerthink.com/blog/sentiment_analysis_hard_but_worth_it (diunduh 14 Nopember 2016).
- [7] Cahyo Ryan Dwi, et al. 2014. Deteksi dan Validasi Informasi Gempa Secara Real-Time Berbasis Social Sensor dengan Twitter. JURNAL TEKNIK POMITS Vol. 2, No. 1.
- [8] Gaikindo Domestic Auto Market by Brand (2013-2014): <http://www.gaikindo.or.id/domestic-auto-market-by-brand-2013-2014/> akses tanggal 23 November 2016.