

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/311107797>

Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naïve Bayes (Studi Kasus: Facebook Page BerryBenka)

Conference Paper · September 2016

CITATIONS

8

READS

11,407

3 authors, including:



[Muhammad Ihsan Zul](#)

Politeknik Caltex Riau

33 PUBLICATIONS 128 CITATIONS

SEE PROFILE

Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naïve Bayes (Studi Kasus: Facebook Page BerryBenka)

Syahmia Gusriani¹⁾, Kartina Diah Kusuma Wardhani²⁾, Muhammad Ihsan Zul³⁾

¹⁾Sistem Informasi, Politeknik Caltex Riau, Pekanbaru 28265, email: syahmia12si@mahasiswa.pcr.ac.id

²⁾Teknik Informatika, Politeknik Caltex Riau, Pekanbaru 28265, email: diah@pcr.ac.id

³⁾Teknik Informatika, Politeknik Caltex Riau, Pekanbaru 28265, email: ihsan@pcr.ac.id

Abstrak – Maraknya penggunaan jejaring sosial seperti Facebook mendorong munculnya data tekstual yang tidak terbatas, sehingga muncul kebutuhan penyajian informasi tanpa mengurangi nilai dari informasi tersebut. Salah satu pemanfaatan data ini adalah untuk mengetahui opini atau sentimen publik terhadap pelayanan dan produk suatu toko *online*. Metodologi yang digunakan untuk melakukan analisis sentimen dimulai dari *collecting data*, *preprocessing*, *feature selection*, klasifikasi dan pengukuran akurasi. *Naive Bayes* merupakan metode klasifikasi yang digunakan untuk memperoleh hasil prediksi klasifikasi sentimen. Hasil analisis pengujian menunjukkan kestabilan akurasi setelah diuji dengan *k-Fold Cross Validation* dan *Confusion Matrix* dengan akurasi 93.7% dimana *minimum support* untuk *Frequent Itemset Mining* adalah 0.014.

Kata Kunci : Opini, Sentimen, *Preprocessing*, *Feature Selection*, *Naïve Bayes*, *Frequent Itemset Mining*.

Abstract - The widespread use of social networks such as Facebook encourage the emergence of textual data that is infinite, so it appears the need of presenting information without reducing the value of the information. One of the **utilization** of this data is to find out the opinion or public sentiment to services and products an online store. The methodology used to perform sentiment analysis is done by employed several steps such as: *collecting data*, *preprocessing*, *feature selection*, *classification* and *measurement accuracy*. *Naive Bayes* is a classification method used to obtain the results of predicted sentiment classification. Results of the analysis showed the stability of tested accuracy. Accuracy was performed by using *k-Fold Cross Validation* and *Confusion Matrix*. This research reach 93.7% accuracy and 0.014 for minimum support of *Frequent Itemset Mining*.

Keywords: Opinion, Sentiment, *Preprocessing*, *Feature Selection*, *Naïve Bayes*, *Frequent Itemset Mining*.

1. PENDAHULUAN

Belanja *online* merupakan proses membeli barang dan jasa dari pedagang yang dijual di internet. Konsumen dapat mengunjungi toko *online* dari rumah secara nyaman sambil duduk di depan komputer. Berdasarkan survey pembelian *online* secara global termasuk Indonesia, sebanyak 71% konsumen melakukan peninjauan terhadap toko *online* sebelum membeli produk. Sebanyak 43% setuju bahwa media sosial menjadi alat bantu untuk memenuhi kebutuhan pengetahuan berupa *review* produk dan ulasan forum, guna membantu membuat keputusan pembelian^[1]. *Review* produk maupun ulasan forum disampaikan melalui komentar di sosial media yang berisi keluhan, pujian atau pandangan terhadap produk atau jasa dari suatu toko *online*. Komentar tersebut mendeskripsikan tanggapan yang berbeda-beda dari setiap konsumen.

Komentar-komentar berupa teks tersebut dapat dikumpulkan dan diolah dengan analisis sentimen. Analisis sentimen mempelajari opini yang mengungkapkan atau mengekspresikan pandangan

positif atau negatif. Analisis sentimen disebut juga *opinion mining*, adalah bidang ilmu yang menganalisa pendapat, sentimen, evaluasi, penilaian, sikap dan emosi publik terhadap entitas seperti produk, jasa, organisasi, individu, masalah, peristiwa, topik, dan atribut mereka^[2]. Opini yang dibutuhkan untuk melakukan analisis berasal dari komentar halaman *fanpage* toko *online* di Facebook. Facebook dipilih karena penggunanya saling berinteraksi secara massif, dimana total pengguna Facebook sebanyak 1,44 miliar dengan pengguna harian sebanyak 936 juta (CNBC, 22 April 2015).

Pada penelitian analisis sentimen ini dibahas bagaimana mengimplementasikan analisis sentimen untuk menentukan kecenderungan padangan publik terhadap toko *online* menggunakan metode klasifikasi *naïve bayes*. *Naïve bayes* merupakan metode pembelajaran mesin yang memiliki model dalam membentuk probabilitas dan peluang. Maka dari itu, *naïve bayes* akan menghitung probabilitas kemunculan kata yang mempresentasikan komentar berdasarkan kelas positif maupun negatif. Berdasarkan hasil

klasifikasi dilakukan pengujian untuk mendapatkan nilai akurasi dari hasil prediksi analisis sentimen dengan metode *naïve bayes*.

2. LANDASAN TEORI

2.1 Penelitian Terdahulu

Analisis sentimen dilakukan oleh Ismail Sunni, dengan meneliti data twitter terkait opini terhadap tokoh publik. Topik pembicaraan mengenai tokoh politik menjadi domain penelitian. Metode yang digunakan adalah F3 (F3 is *Factor Finder*) untuk menangani model bahasa. Sedangkan *Naïve Bayes Classifier* digunakan untuk melakukan analisis sentimen dengan menggunakan *frequency based selection* untuk mengurangi kesalahan eja dari banyaknya *word*. Hasil analisis menyebutkan bahwa *casefolding* justru menurunkan hasil akurasi karena program menyamakan kata yang muncul sebagai kabar berita dan opini hasil dari pengguna Twitter.

Penelitian lain mengenai analisis sentimen dilakukan oleh Nur (2011). Objek yang diteliti adalah dokumen-dokumen berbahasa Indonesia terkait merk telepon seluler. Penelitian ini bertujuan untuk mengetahui akurasi perbandingan dua klasifikasi yang digunakan dalam analisis. Penelitian analisis sentimen menggunakan pendekatan *machine learning* yang dikenal dengan *Support Vector Machine* (SVM) dan dikhususkan pada dokumen teks berbahasa Indonesia.

Pada tahun 2014, Ahmad Fathan juga melakukan penelitian analisis sentimen. Objek penelitiannya adalah *tweet* yang membicarakan tokoh publik dengan tujuan mengetahui suatu topik terkait tokoh publik atau kejadian yang menyebabkan sentimen publik turun. Nama tokoh publik yang dianalisis adalah tokoh publik dengan hasil survei tertinggi berdasarkan survei lembaga-lembaga kompeten di Indonesia. *Naïve Bayes Classifier* digunakan untuk memperoleh hasil klasifikasi dari sentimen terkait tokoh publik.

Penelitian ini melakukan analisis sentimen terhadap data komentar yang ada pada *fanpage* facebook terkait layanan dan produk yang dijual toko online BerryBenka. *Frequent Itemset Mining* digunakan sebagai metode pemilihan atribut yang paling banyak muncul berdasarkan nilai *minimum support*. Klasifikasi menggunakan metode *naïve bayes*. Hasil dari penelitian ini ialah hasil analisis sentimen dan visualisasi kelompok kata sentimen yang paling banyak muncul.

2.2. Analisis Sentimen

Analisis sentimen disebut juga *opinion mining*, adalah bidang ilmu yang menganalisa pendapat, sentimen, evaluasi, penilaian, sikap dan emosi publik terhadap entitas seperti produk, jasa, organisasi, individu,

masalah, peristiwa, topik, dan atribut mereka^[2]. Analisis sentimen berfokus pada opini-opini yang mengekspresikan atau mengungkapkan sentimen positif atau negatif.

2.3. Text Mining

Text mining juga dikenal dengan *text data mining* atau pencarian pengetahuan di basis data tekstual adalah proses yang semi otomatis melakukan ekstraksi dari pola data^[3]. Tipe pekerjaan *text mining* meliputi kategorisasi, *text clustering*, ekstraksi konsep/entitas, analisis sentimen, *document summarization*, dan *entity-relation modeling* (yaitu, hubungan pembelajaran antara entitas)^[4]. Sumber data yang digunakan pada *text mining* adalah kumpulan teks yang memiliki format yang tidak terstruktur atau minimal semi terstruktur. Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen.

2.4. Preprocessing

Preprocessing dilakukan untuk menghindari data yang kurang sempurna, gangguan pada data, dan data-data yang tidak konsisten^[5]. Tahap *preprocessing* pada penelitian ini diantaranya:

Cleansing: tahap menghapus tanda baca dan angka yang terdapat pada teks komentar.

Case folding: yaitu penyeragaman bentuk huruf menjadi *lower case* atau *upper case* serta penghapusan angka dan tanda baca^[6].

Filtering: *Filtering* dilakukan dengan menggunakan algoritma *stopword removal*. *Stopword removal* digunakan untuk membuang kata-kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansinya dengan teks^[7].

Stemming: proses mencari akar kata dan menghilangkan imbuhan pada kata^[8]. *Stemming* bertujuan mengurangi variasi kata yang memiliki kata dasar sama.

Convert negation: proses mengganti negasi yang terdapat dalam komentar. Negasi adalah sesuatu yang dikenal dalam semua bahasa dan biasanya negasi digunakan untuk mengubah polaritas dari suatu pernyataan^[9]. Contoh kata-kata yang bersifat negasi adalah “kurang”, “tidak”, “enggak” dan “tak”.

2.5. Frequent Itemset Mining

Frequent item set mining (FIM) merupakan proses yang melakukan ekstraksi informasi dari *database* berdasarkan seberapa sering suatu *event* terjadi, yaitu suatu peristiwa atau suatu set peristiwa^[10]. *Frequent itemset* merupakan tahapan yang dilakukan untuk memperoleh kata-kata yang paling sering muncul diantara himpunan item (*event*).

2.6. Naïve Bayes

Naïve Bayes Classifier membuat asumsi yang sangat kuat (naif) akan independensi dari masing-masing kelas kejadian yang diberikan label^[4]. *Naïve Bayes Classifier* digunakan untuk klasifikasi sentimen dari

data yang telah diperoleh. Persamaan *naïve bayes* yang digunakan untuk menentukan kelas dari komentar ditunjukkan dalam persamaan berikut.

$$V_{MAP} = \underset{v_j}{\operatorname{argmax}} \, \epsilon v P(v_j) \prod_{i=1}^n P(a_i | v_j)$$

Keterangan :

a_i = atribut atau fitur ke- i

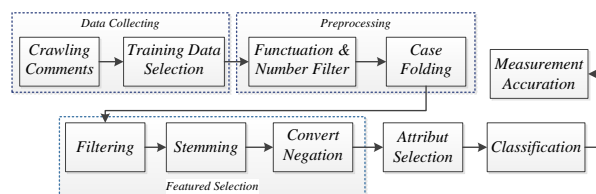
v_j = kelas ke- j (positif atau negatif)

V = himpunan kelas target

V_{MAP} = kelas sentimen suatu komentar

3. METODOLOGI

3.1. Alur Perancangan



Gambar 1 Alur Perancangan

Metodologi perancangan dimulai dari pengumpulan data komentar, *preprocessing*, *featured selection*, pemilihan atribut, klasifikasi, dan pengujian akurasi.

Data collecting merupakan tahapan pengumpulan *training data* menyaring *keywords* yang berhubungan dengan pelayanan/produk *fashion* dikombinasikan dengan kata-kata sentimen dan atau *emoticon*. Berikut ketentuan yang digunakan dalam memilih data *training* diantara ribuan data komentar yang hasil *crawling* :

1. Menentukan target/entitas. Target entitas diperoleh setelah dilakukan frequent itemset terhadap data komentar yang masih belum melalui *preprocessing*.
 - a. Pelayanan, yaitu terkait pelayanan pengiriman dan respon komplain pelanggan. *Keyword* terkait pelayanan diantaranya; kirim, order, pesan, paket, klarifikasi, komplain, balas, respon, jawab, retur, barang, konfirmasi, jawab dan retur.
 - b. Produk *fashion*, yaitu terkait kualitas produk yang dikirim apakah bentuk, ukuran, jenis, motif ataupun warnanya sesuai dengan yang dipesan. *Keyword* yang digunakan ialah baju, shirt, kaos, kasu, *dress*, ukuran, warna, *size*, sepatu, jam, *watch*, celana, sandal, sendal, *heels*, *wedges*, tas, koko, kemeja.

2. Menentukan *keyword* sentimen positif/negatif dan atau *emoticon* yang mengiringi suatu kalimat komentar.

- a. Pelayanan-sentimen-positif ; sampai, sampe, cepat, cepet, sudah, datang, mudah, thx, dan terima kasih.
- b. Pelayanan-sentimen-negatif ; lama, kecewa, lambat, telat, susah, sulit, parah, bohong, salah, batal dan maaf.
- c. Produk-sentimen-positif ; bagus, percaya, keren, cantik, cakep, puas, oke dan sesuai.
- d. Produk-sentimen-negatif ; kecewa, jelek, rusak, parah, komplain, salah, ganti, tukar, retur, *refund*, rongsok dan bohong.

3. Melakukan *filtering* kata target/entitas dan sentimen ada dalam satu kalimat komentar. Jika salah satu tidak ada, maka tidak layak menjadi data *training*.

Tahap *preprocessing* diperlukan untuk membersihkan data dari *noise*, menyeragamkan bentuk kata dan mengurangi volume kata. Agar pada tahap masuk ke dalam metode klasifikasi lebih optimal dalam perhitungannya. Tahap *feature selection* merupakan tahap lanjut dari pengurangan dimensi pada proses transformasi teks.

Sebelum hasil *preprocessing* diproses dengan algoritma klasifikasi, data *training* akan dilakukan proses penentuan atribut (*attribut selection*). Tahap ini akan menyaring setiap kata yang muncul berdasarkan jumlah kemunculan kata.

Proses klasifikasi merupakan tahapan utama untuk mengetahui pola sentimen dari komentar-komentar yang telah *dipreprocessing*. Klasifikasi dilakukan dengan membandingkan hasil akurasi dari tiga metode yakni, *k-NN*, *Naïve Bayes*, dan *Decision Tree*. Perhitungan akurasi menggunakan *k-Fold Cross Validation* dan *Confusion Matrix*.

4. HASIL DAN PEMBAHASAN

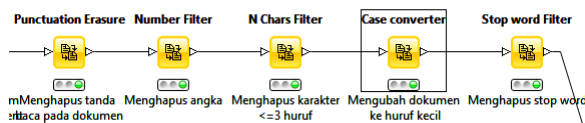
4.1. Collecting Data Result

Sumber data teks yang akan diolah pada penelitian ini berasal dari halaman *fanpage* facebook toko *online* Berrybenka. Data teks berasal dari komentar-komentar terkait produk dan pelayanan dari pengunjung *fanpage* yang diambil dengan melakukan *crawling text*. Jumlah data yang berhasil *dicrawling* sebanyak 755 data komentar selama satu tahun (Agustus 2014-Agustus 2015). Jumlah data ini akan menyusut ketika melalui proses pemilihan data latih.

Row3	aku sudah 3 kali loh belanja disini....tadinya takut2 karena OnlineShop...ta
Row4	aku suka gaya bgt' cucok.....
Row5	aku pengen baju ya bagus ! tp tidak tau nii tempat ya
Row6	bagaimana cara menjadi member?? produknya bagus2,,, via desktop?
Row7	bagus2 bangeetzzzz... but only perfect body yach ??? for big zise body l
Row8	Bajunya bagus , aku suka !
Row9	belanja disini pokoknya suka respon cepat dan terpercaya,suka suka suka
Row10	Belanja di berrybenka sangat memuaskan bgt.. aku senang bgt pketku dh c
Row11	Berry benka,pesanan ku ud sampai td siang.. oke bgt brgny,dipacking dlm c
Row12	Berrybenka barangnya sudah sampai tadi siang,dress nya bagus,sesuai sai
Row13	Berrybenka barangnya sudah sampai.makasih ya. bagus sepatunya. And
Row14	Berrybenka is the best. Saya baru pertama kali berbelanja di berrybenka di
Row15	Berrybenka katanya mw ada diskon ya sehubungan dg 12.12.12. Jam 12 s
Row16	Berrybenka.com / hijabenka.com... Jazakillah khairannnnnnn paketan aku s
Row17	BIG terimakasih 4 BERRYBENKA suka suka suka Sebrnnya sy sudah orde
Row18	boleh bli saputnya tidak ??? bagus bagus soale :3

Gambar 2 Data Komentar BerryBenka

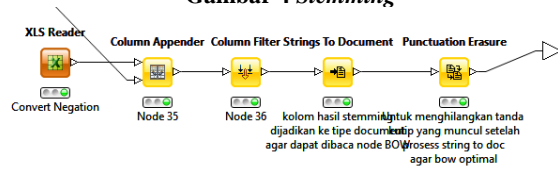
4.2. Preprocessing



Gambar 3 Preprocessing (1)



Gambar 4 Stemming



Gambar 5 Convert Negation

Preprocessing dilakukan menggunakan aplikasi miner yaitu KNIME. Preprocessing diawali dari proses *cleansing* hingga *stopword removal/filtering*. Dilanjutkan dengan *stemming* dan *convert negation*. Proses *stemming* Bahasa Indonesia harus dilakukan dengan menambahkan kode java pada Java Snippet. Hasilnya adalah data komentar dengan tingkat *noise* lebih rendah.

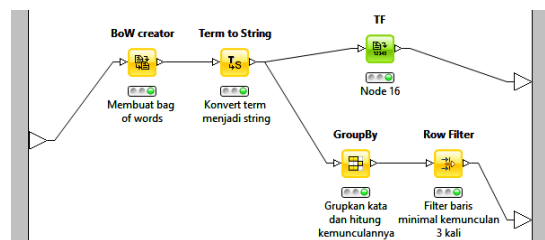
Row3	"belanja takut2 onlineshoptapi san perdana terima puas bgtjadi t...
Row4	"suka gaya cucok"
Row5	"pengen baju bagus tidak"
Row6	"memberproduk bagus2desktop"
Row7	"bagus2 bangeetzzzzz fect body yach zise body katidak muaaat...
Row8	"baju bagus suka"
Row9	"belanja pokok suka respon cepat cayasuka suka suka senang se...
Row10	"belanja rybenka puas senang pket hsl sesuai inginkandn psen lgi...
Row11	"berry nkapesen sampai siangbrgnydipacking hitam elegan imaka...
Row12	"berrybenka barang sampai siangdress bagusseuai foto cepat s...
Row13	"berrybenka barang sampaimakasih bagus patunya love"
Row14	"berrybenka best belanja rybenka takut barang tidaksesuai terim...
Row15	"berrybenka diskon hubung nang"
Row16	"berrybenacom hijabenacom jazakil khairannnnnnn paket sampai...

Gambar 6 Hasil Akhir Preprocessing

4.3. Frequent Itemset Mining

FIM atau pemilihan atribut dilakukan dengan menggunakan node *tf* dan dikombinasikan dengan node *GroupBy*. *Tf* digunakan untuk mendapatkan nilai *frequent term relative* atau *minimum support* dari

suatu kata sentimen. Nilai *minimum support* yang diperoleh adalah 0.014.



Gambar 7 Frequent Itemset Mining

4.4. Klasifikasi Naïve Bayes

Row ID	D saran	D ner	D tenggang	D formulir	D terimak...	S. Docum...
S191	0	0	0	0	0	negatif
S192	0	0	0	0	0	negatif
S193	0	1	0	0	0	negatif
S194	0	0	0	0	0	positif
S195	0	0	0	0	0	negatif
S196	0	0	0	0	0	positif
S197	0	0	0	0	0	positif
S198	0	0	0	0	0	negatif
S199	0	0	0	0	0	positif
S200	0	0	0	0	0	positif
S201	0	0	0	0	0	positif
S202	0	0	0	0	0	positif
S203	0	0	0	0	0	positif
S204	0	0	0	0	0	positif

Gambar 8 Data Komentar Sebelum Klasifikasi

Sebelum masuk ke proses klasifikasi, data atribut harus diberi label sentimen seperti pada Gambar 3.7. Klasifikasi dengan metode *naïve bayes* pada KNIME memerlukan dua node, yaitu *Naïve Bayes Learner* dan *Naïve Bayes Predictor*. *Naïve Bayes Learner* berfungsi sebagai node yang mempelajari pola klasifikasi data latih dan *Naïve Bayes Predictor* berfungsi memprediksi data testing berdasarkan pola yang diperoleh dari node *Predictor*.

Gambar 9 Hasil Prediksi Naïve Bayes

Hasil klasifikasi *naïve bayes* langsung menampilkan hasil prediksi dari semua data. Pada proses *naïve bayes* ini menggunakan *minimum support* 0.014 pada data BerryBenka Layanan dan Produk.

4.5. Pengujian

Pengujian pertama dilakukan dengan *k-Fold Cross Validation* yang berfungsi menentukan data *training* dan data *testing* terbaik. Pada data komentar BerryBenka Layanan dan Produk fold dengan nilai

error paling kecil adalah *fold* #0. Kesalahan klasifikasi terjadi pada 2 data diantara 68 baris data.

Row ID	selamat	D	terimakasih	S	Hasil Pr...	fold #
Row614	1	0	postif	postif	0	
Row631	0	0	negatif	negatif	0	
Row636	0	0	negatif	negatif	0	
Row659	0	0	negatif	negatif	0	
Row665	0	0	negatif	negatif	0	
Row1	0	0	negatif	negatif	1	
Row4	0	0	negatif	negatif	1	
Row15	0	0	negatif	negatif	1	
Row17	0	0	negatif	negatif	1	
Row30	0	0	postif	postif	1	
Row37	0	0	negatif	negatif	1	
Row70	0	0	postif	postif	1	
Row74	0	0	negatif	negatif	1	
Row81	0	0	postif	negatif	1	
Row85	0	0	negatif	negatif	1	
Row87	0	0	negatif	negatif	1	
Row115	0	0	negatif	negatif	1	

Gambar 10 Hasil Pengujian Prediksi 10-Fold
Hasil akurasi dari *confusion matrix* adalah 93.7% dengan kecenderungan sentimen adalah negatif. Dimana sentimen *true negative* adalah sebanyak 530 baris dan *true positive* adalah 105 baris.

Row ID	D	Error in %	Size of ...	Error C...
fold 0	2.941	68	2	
fold 1	8.824	68	6	
fold 2	7.353	68	5	
fold 3	5.882	68	4	
fold 4	8.824	68	6	
fold 5	5.882	68	4	
fold 6	10.294	68	7	
fold 7	2.985	67	2	
fold 8	5.97	67	4	
fold 9	2.985	67	2	

Gambar 11 Hasil Perhitungan Error Rate k-Fold
Hasil akurasi dari *confusion matrix* (Gambar 3.11) adalah 93.7% dengan kecenderungan sentimen adalah negatif. Dimana sentimen *true negative* adalah sebanyak 530 baris dan *true positive* adalah 105 baris.

Document ...	positif	negatif
positif	105	35
negatif	7	530

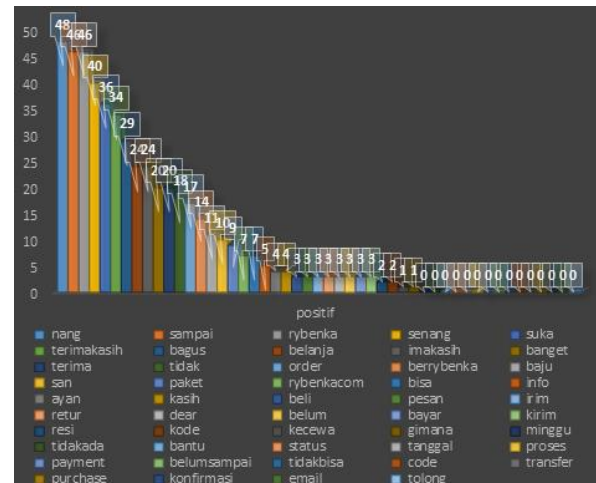
Correct classified: 635 Wrong classified: 42
Accuracy: 93.796 % Error: 6.204 %
Cohen's kappa (κ) 0.796

Gambar 12 Confusion Matrix

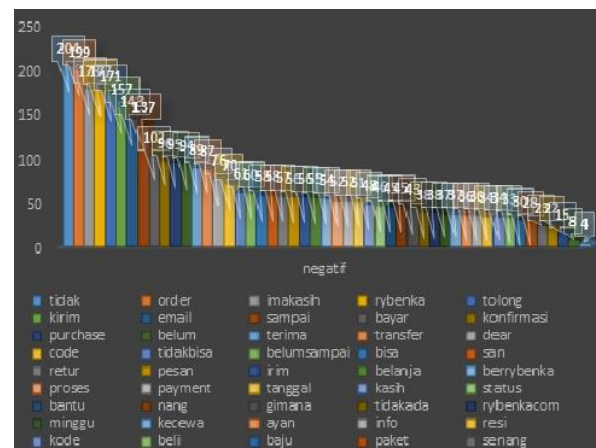
4.6. Visualisasi

Kata-kata komentar yang paling banyak muncul sebagai sentimen pada dokumen Berrybenka layanan dan produk divisualisasikan pada Gambar 4.13 dan Gambar 3.14. Sentimen positif terhadap pelayanan

dan produk cenderung mengungkapkan rasa senang dan terimakasih, produk yang dipesan sudah sampai dan kualitasnya bagus. Sedangkan sentimen negatif mengeluhkan barang yang dipesan belum sampai dan meminta konfirmasi pihak Berrybenka untuk memberikan penjelasan kapan pesannya akan sampai.



Gambar 13 Visualisasi Kata Sentimen Positif



Gambar 14 Visualisasi Kata Sentimen Negatif

5. KESIMPULAN

Kesimpulan yang diperoleh dari penelitian ini ialah *Naïve bayes* dapat dijadikan metode klasifikasi untuk analisis sentimen dengan keakuratan 93.7%. Hasil akurasi diperoleh setelah dilakukan pengujian terhadap berbagai nilai *Frequent Itemset* dengan *minimum support* 0.014.

Analisis sentimen ini sebaiknya diimplementasikan dengan meningkatkan kemampuan *crawling* data menggunakan *power query* dengan mempelajari *query* facebook untuk memperoleh lebih banyak data dalam jangka waktu lebih dari satu tahun. Sehingga, akan memungkinkan melihat sentimen publik dari tahun

ketahun. Penelitian ini dapat dikembangkan dengan membuat sistem analisis sentimen menggunakan metode klasifikasi *Naïve bayes*.

DAFTAR REFERENSI

- [1] Nielsen. (2014). *E-Commerce : Evolution or Revolution in the Fast-Moving Consumer Goods World?*. United States : New York.
- [2] Liu, Bing. (2012). *Sentiment Analysis and Opinion Mining*. USA : Morgan & Claypool Publishers.
- [3] Ma, Jian dkk. (2012). An Ontology-Based Text-Mining Method to Cluster Proposals for Research Project Selection. *IEEE Trans. Syst., Man, Cybern. Part A, Syst & Humans* (vol. 42).
- [4] Han, Jiawei., Kamber, Micheline., & Pei, Jian. (2012). *Data Mining Concepts and Techniques* (3rd ed.). USA : Simon Frase University.
- [5] Hemalatha, I., Varma, G. P Saradhi., & Govardhan, A. (2012). Preprocessing the Informal Text of efficient Sentiment Analysis. India : *International Journal of Emerging Trends & Technology in Computer Science*.
- [6] Putranti, Noviah Dwi., & Winarko, Edi. (2014). Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan *Maximum Entropy* dan *Support Vector Machine*. Yogyakarta : Universitas Gajah Mada.
- [7] Sentiaji, Aditia Rakhmat., & Bachtar, Adam Mukaharil. (2014). Analisis Sentimen Terhadap Acara Televisi Berdasarkan Opini Publik. Bandung : Universitas Komputer Indonesia.
- [8] Hidayatullah, Ahmad Fathan. (2014). Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik pada Twitter. Yogyakarta : Universitas Gajah Mada.
- [9] Blanco, Eduardo., & Moldovan, Dan. (2011). Some Issues on Detecting Negation from Text. *Journal of Human Language Technology Research Institute*. USA : The University of Texas at Dallas.
- [10] Moens, Sandy., Aksehirli, Emin., & Goethals, Bart. (2013). *Frequent Itemset Mining for Big Data*. Belgium : Universiteit Antwerpen.