

ANALISIS SENTIMEN MASYARAKAT TERHADAP DATA TWEET LAZADA MENGGUNAKAN *TEXT MINING* DAN ALGORITMA *NAIVE BAYES CLASSIFIER*

Dodi setian¹, iin seprina²

Fakultas Teknik Ilmu Komputer, Universitas Bina Darma
Email: dodisetian92@gmail.com¹, iin_seprina@binadarma.ac.id²

ABSTRACT

Lazada is an online shopping center that offers various types of products guaranteeing consumer comfort when consumers are browsing the product being sought and also guarantees secure payment options. Twitter is one of the most popular social media in Indonesia. Most Twitter users also use this social media to find business people, where they will become followers (followers) and interact with business people. Collection of tweet data from Twitter can be done by integrating Twitter API and Rapid miner. Data analysis method used to analyze public sentiment in Lazada's tweets using the Naïve Bayes classifier algorithm. The results of the analysis of public sentiment on Indonesian lazada tweet data using text mining with the naïve-bayes classifier algorithm if it is proven that shows significant results on sentiment analysis with an accuracy rate of 98.29%.

Keywords: *Lazada Indonesia, Twitter, Text Mining, Naïve Bayes Algorithm*

ABSTRACT

Lazada adalah pusat belanja online yang menawarkan berbagai macam jenis produk menjamin kenyamanan konsumen ketika konsumen browsing produk yang sedang dicari dan juga menjamin opsi pembayaran yang aman. *Twitter* merupakan salah satu media sosial paling populer di Indonesia. Sebagian besar pengguna *Twitter* juga memanfaatkan media sosial ini untuk menemukan pelaku bisnis, dimana mereka akan menjadi pengikut (*followers*) dan berinteraksi dengan pelaku bisnis. Pengumpulan data *tweet* dari *Twitter* dapat dilakukan dengan mengintegrasikan *Twitter API* dan *Rapid miner*. Metode analisis data yang dipakai untuk melakukan menganalisis sentimen masyarakat dalam *tweet* Lazada dengan menggunakan algoritma *Naive bayes classifier*. Hasilnya analisis sentimen masyarakat terhadap data *tweet lazada indonesia* menggunakan *text mining* dengan algoritma *naive-bayes classifier* jika Terbukti bahwa menunjukkan hasil yang signifikan terhadap analisis sentimen dengan tingkat akurasi sebesar 98,29%.

Kata kunci: *Lazada Indonesia, Twitter, Text Mining, Algoritma Naïve Bayes*

1. PENDAHULUAN

Advertising telah berkembang menjadi system komunikasi yang sangat penting antara pelaku bisnis dan konsumen. Pelaku bisnis sangat mengandalkan *advertising* untuk membantu memasarkan produk atau layanan yang mereka miliki agar dapat

Dikenal oleh konsumen. Kelebihan tersebut menjadikan *advertising* memiliki peranan utama dalam program pemasaran.

Lazada menjamin kenyamanan konsumen ketika konsumen browsing produk yang sedang dicari dan juga menjamin opsi pembayaran yang aman. Pilihan pembayaran Lazada termasuk kartu kredit, cash on delivery, Bank transfer, Mobile banking dan bahkan melalui layanan pembayaran online seperti halnya HelloPay. Bahkan, pada aplikasinya Lazada memiliki promo sendiri yang berbeda dari Lazada versi desktop.

Twitter merupakan salah satu media sosial paling populer di Indonesia, Negara, E. S. (2017). Sebagian besar pengguna *Twitter* juga memanfaatkan media sosial ini untuk menemukan pelaku bisnis, dimana mereka akan menjadi pengikut (*followers*) dan berinteraksi dengan pelaku bisnis, Antoni, D., Negara, E. S., & Suweno, S. (2015), Negara, E. S., Andryani, R., & Saksono, P. H. (2016). Pengumpulan data *tweet* dari *Twitter* dapat dilakukan dengan mengintegrasikan *Twitter API* dan *Rapid miner*. Untuk mempermudah mengetahui jenis konten dari sejumlah data *tweet*, maka perlu dilakukan proses *Text mining* terhadap data *tweet* tersebut dengan menerapkan teknik pembobotan kata, Andryani, R., Negara, E. S., & Triadi, D. (2019). Pada *Text Mining*. Pembobotan tersebut hanya berpedoman pada frekuensi munculnya term dalam suatu dokumen dan tidak melihat frekuensi kemunculan term tersebut di dalam dokumen lainnya.

Lazada Indonesia merupakan salah satu pelaku bisnis di Indonesia yang bergerak di bidang e-commerce dan menggunakan media sosial *Twitter* sebagai sarana untuk melakukan *advertising*. Dari latar belakang tersebut, menuliskan “**ANALISIS SENTIMEN MASYARAKAT TERHADAP DATA TWEET LAZADA INDONESIA MENGGUNAKAN TEXT MINING DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER**”.

2. Metode Penelitian

2.1. Waktu Penelitian

Waktu penelitian dilakukan pada bulan Mei 2019 sampai dengan bulan September 2019. Objek penelitian yaitu Lazada Indonesia dengan database *tweet* yang diambil dari *Twitter*.

2.2. Metode Analisis Data

Metode analisis data yang dipakai untuk melakukan menganalisis sentimen masyarakat dalam *tweet* Lazada dengan menggunakan algoritma Naïve bayes classifier. Algoritma tersebut banyak digunakan dalam analisis khususnya analisis sentimen agar dapat menghasilkan data sentimen baik positif dan negative, Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018).

Naïve bayes classifier (NBC) merupakan sebuah metode klasifikasi pada teorema bayes. Metode klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan inggris Thomas bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari naïve bayes classifier adalah asumsi yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian. Dengan metode naïve bayes ini pengguna dapat mencapai probabilitas Bayesian dan bisa membuat pernyataan probabilitas. Naïve bayes dapat di formulasikan menjadi persamaan

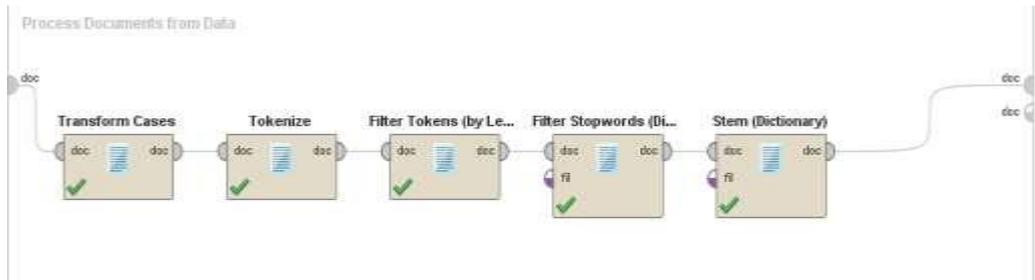
$$P(c|X) = \frac{P(x|C)P(c)}{P(x)} \quad (1)$$

Naïve bayes classifier ini menggunakan model classifier probabilistic yang juga termasuk salah satu algoritma yang dapat mengklarifikasi sebuah data.

3. HASIL DAN PEMBAHASAN

3.1 process Text

Sebelum melakukan analisis sentiment dan klasifikasi menggunakan algoritma naïve bayes, penerapan text mining untuk mengolah text dalam hal ini adalah preprocessing text. Text akan di pisahkan menjadi unit yang lebih kecil (token), dan akan di lakukan *filtering, normalization (stemming)* pada rapid miner



Gambar 4.5 Process Text Pada Rapid Miner

Berikut ini tahapan dalam process text sebelum text di analisa menggunakan algoritma naïve bayes di rapid miner

1. *Transform cases*

Langka pertama dari proses text adalah *transform cases* yaitu mengubah data *input*, ini berlaku transformasi huruf kecil secara default

- Lowercases* akan mengubah semua text menjadi huruf kecil (Ban Motor Kenda = ban motor kenda)
- uppercases* akan mengubah semua text menjadi huruf besar (Ban Motor Kenda = BAN MOTOR KENDA)

2. *Tokenize*

Setelah melakukan *transformcases* langkah kedua dari process text adalah *tokenize* yaitu metode memecah text menjadi komponen yang lebih kecil (kata, kalimat, text, tweet)

3. *filters tokens*

Setelah melakukan *tokenize*, langkah ketiga dari process text adalah *filters tokens* yaitu memfilter token berdasarkan panjangnya (yaitu jumlah karakter yang dikandungnya)

4. *filter stopwords*

Setelah melakukan *filters tokens* langkah selanjutnya adalah *filter stopwords* dalam tahap ini *stopwords*, menghapus *stopwords* dari teks misalnya (ada, adalah, adanya, adapun.....) . pada aplikasi rapid miner tidak menyediakan *stopwords* yang dapat di unduh langsung menggunakan aplikasi rapid miner, akan tetapi *stopwords* dapat di unduh melalui browser (mozilla, google chrome). Berikut ini beberapa *stopwords* yang akan di hilangkan

Indonesia Stopwords
Yang, di, dan, itu, dengan, untuk, tidak, ini, dari, dalam, akan, pada, juga, saya, ke, karena, tersebut, bias, ada, mereka, lebih, kata, tahun, sudah, atau, saat, oleh, menjadi, orang, ia, telah, adalah, seperti, sebagai, bahwa, dapat, para, harus, namun, kita, dua, satu, masih, hari, hanya, mengatakan, kepada, kami, setelah, melakukan, lalu, belum, lain, dia, kalau, terjadi, banyak, menurut, anda, hingga, tak, baru, beberapa, ketika, saja, jalan, sekitar, secara, dilakukan, sementara, tapi, sangat, hal, sehingga, seorang, bagi, besar, lagi, selama, antara, waktu, sebuah, jika, sampai, jadi, terhadap, tiga, serta, pun, salah, merupakan, atas, sejak, membuat, baik, memiliki, kembali, selain, tetapi, pertama, kedua, memang, pernah, apa, mulai, sama, tentang, bukan, agar, semua, sedang, kali, kemudian, hasil, sejumlah, juta, persen, sendiri, katanya, demikian, masalah, mungkin, umum, setiap, bulan, bagian, bila, lainnya, terus, luar, cukup, termasuk, sebelumnya, bahkan, wib, tempat, perlu, menggunakan, memberikan, rabu, sedangkan, Kamis, langsung, apakah, pihak, melalui, diri, mencapai, minggu, aku, berada, tinggi, ingin, sebelum, tengah, kini,

5. *Stem(stemming)*

Proses terakhir pada preproses text adalah *stem(stemming)* yaitu merubah berbagai kata imbuhan menjadi kata dasarnya misalnya (membenarkan = betul, berpegangan = pegang)

3.2 Hasil Analisis Sentiment

Hasil analisis sentiment menggunakan algoritma naïve bayes pada penelitian ini berupa netral, positif dan negative yang dilihat dari hasil akhir berikut :

accuracy: 98.29%

	true neutral	true negative	true positive	class precision
pred. neutral	591	0	0	100.00%
pred. negative	1	15	0	93.75%
pred. positive	11	0	82	88.17%
class recall	98.01%	100.00%	100.00%	

Gambar 4.7 Hasil Analisis Sentimen Terhadap Tweet Lazada Indonesia

Mengacu pada hasil penelitian (handayani : 2018) maka dari hasil tabel di atas dapat di simpulkan bahwa respon suasana hati terhadap respon di media sosial bersifat negatif sebanyak 16 tweet dikarenakan penguanya cenderung tidak menggunakan pertimbangan logika dan perasaan terlebih dahulu saat mengambil suatu keputusan dan memiliki perasaan dan pikiran negative. Hasil respon positif sebanyak 93 tweet dikarenakan sipengguna cenderung lebih berfikir matang dan

mempertimbangkan resiko yang didapat terlebih dahulu sebelum berkomentar sehingga mengakibatkan si pengguna dapat berkomentar yang bersifat positif.

Sedangkan 591 tweet yang bersifat netral yang didapat dikarenakan sipengguna sering berpikir matang terlebih dahulu dan menggunakan pertimbangan logika dan perasaan saat mengambil suatu keputusan untuk mengurangi resiko, cenderung memiliki perasaan dan pikiran negatif saat mengalami kegagalan namun dapat berperilaku positif sehingga komentar yang dihasilkan berbentuk tweet bersifat netral.

Berikut contoh tweet sentiment masyarakat yang di prediksi bersentimen positif, negative, dan netral :

1. Prediksi Positif

@ErieRizaNugraha “Siapa yang suka diskon? Hayo ngaku!”
@Mutiara Goenawan “@LazadaID kamu baaaeekk deh. Rikues ku d approved.
Dududududu”

2. Prediksi Negative

@Khaanggoro “Ini hape saya yg error atau kamunya yg error @LazadaID
@ari257 “@LazadaID min, sy pesan barang dr tanggal 7 Agustus, smp hari ini kok belum sampe2 yah? Gak bisa tracking, dan pas nanya sellernya katanya suruh tunggu aja?”

3. Prediksi Netral

@Lazada.co.id “@nisability Hi Kak Nisa, diabaikan saja ya infonya. Apabila kamu memiliki pertanyaan seputar Lazada bisa langsung ke akun @LazadaIDCare ya agar dibantu lebih lanjut.”

@Lazada.co.id “@anisaraish Hi Kak @anisaraish, apabila kamu memiliki pertanyaan mengenai pemesanan di Lazada bisa langsung DM ke akun @LazadaIDCare ya agar nanti dapat dibantu lebih lanjut”

Data diatas diproses menggunakan 700 data keseluruhan dari data tweet dan retweet pada akun twitter lazada Indonesia Berdasarkan data di atas terlihat bahwa class recall dan class precision. Pada class recall netral terdapat total keseluruhan data 603 tweet 591 di prediksi sebagai sentiment netral 1 sentiment negative dan 11 sentimen positive maka persentase class recall pada true netral sebanyak 98,01%

dari keseluruhan true netral. dan Pada class recall negative terdapat total keseluruhan data 15 tweet dan ke 15 data tersebut benar – benar di prediksi sebagai sentiment negative maka persentase class recall pada true negative sebanyak 100 %. Sedangkan Pada class recall positive terdapat total keseluruhan data 82 tweet dan ke 82 data tersebut benar – benar di prediksi sebagai sentiment positive maka persentase class recall pada true positive sebanyak 100 %

Pada class precision prediksi sentiment netral terdapat 591 tweet benar-benar netral, 16 tweet yang dimana 1 di prediksi sentiment negative namun mengandung unsur sentiment netral dan 15 tweet yang diprediksikan benar-benar merupakan sentiment negative, dan class precision terdapat 93 data yang dimana 11 diprediksikan sebagai sentiment positif namun masih mengandung unsure netral dan 82 benar-benar di prediksi sebagai sentiment positif.

Berikut merupakan data akurasi dari proses algoritma naïve bayes menggunakan rapid miner

:

```
PerformanceVector:  
accuracy: 98.29%  
ConfusionMatrix:  
True:  neutral negative      positive  
neutral:      591      0      0  
negative:      1      15      0  
positive:     11      0      82  
classification_error: 1.71%
```

Berdasarkan data di atas di jelaskan bahwa dari 700 data keseluruhan berdasarkan sentiment masyarakat melalui tweet lazada terdapat 684 data akurat yang merupakan true sentiment 591 true netral 15 true negative 82 true positive dengan tingkat akurasi sebanyak 98,29%, dan 12 data error yang merupakan prediksi di mana 1 data di prediksi negative namun masih memiliki unsure netral sedangkan 11 prediksi positive namun mengandung unsure netral dengan tingkat error 1,71%.

Keakurasian klasifikasi algoritma naïve bayes di lihat dari besarnya performance Vector dari data akurat dan error melalui rapid miner, hasil analisis menunjukkan bahwa data yang akurat memiliki persentase 98,29% yang lebih tinggi di dibandingkan dengan jumlah persentase error yaitu 1,17%

Algoritma naïve bayes adalah algoritma yang di gunakan untuk mengklasifikasikan data , dan juga dapat di gunakan untuk mengambil sebuah keputusan (berdasarkan keakurasian data)sesuai dengan rumusan masalah yang telah dikemukakan Bagaimana melakukan analisis sentiment masyarakat terhadap data *tweet lazada* Indonesia menggunakan *text mining* dengan algoritma *naïve-bayes classifier*? Terbukti bahwa analisis sentiment terhadap data tweet lazada Indonesia menggunakan text mining dengan algoritma naïve bayes menunjukkan hasil yang signifikan terhadap analisis sentiment dengan tingkat akurasi sebesar 98,29%.

4. Kesimpulan

Berdasarkan hasil analisis dan pembahasan analisis sentimen masyarakat terhadap data *tweet lazada* indonesia menggunakan *text mining* dengan algoritma *naïve-bayes classifier*, dapat disimpulkan bahwa analisis sentimen masyarakat terhadap data *tweet lazada* indonesia menggunakan *text mining* dengan algoritma *naïve-bayes classifier* jika Terbukti bahwa menunjukkan hasil yang signifikan terhadap analisis sentimen dengan tingkat akurasi sebesar 98,29%.

DAFTAR PUSTAKA

- 1) Adawiyah, D.R., dan Waysima. 1994. Evaluasi Sensori Produk Pangan (edisi 1). Departemen Ilmu dan Teknologi Pangan, Fakultas Teknologi Pertanian, IPB.

- 2) Alexa.(2016). *PeringkatSitus Di Indonesia*. Diambil kembali dariAlexa: <http://www.alexacom/topsites/countries/ID>
- 3) Antoni, D., Negara, E. S., & Suweno, S. (2015). Ekstraksi Data Geo-Spatial Twitter (Studi Kasus: Badan Penyelenggara Jaminan Sosial Kesehatan).
- 4) Andryani, R., Negara, E. S., & Triadi, D. (2019). Social Media Analytics: Data Utilization of Social Media for Research. *Journal of Information Systems and Informatics*, 1(2), 193-205.
- 5) Buntoro,G.A.2017.Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter.*Integer Journal*, Vol 2, No 1, Maret 2017:32-41
- 6) Kesuma, H.W.A. 2016.Penerapan *Cosine Similarity* dalam Aplikasi Kitab Undang-Undang Hukum Dagang (*Wetboek Van Koop handle Voor Indonesia*). Semarang: Universitas Negeri Semarang
- 7) Kurniawan, Bambang. 2012. Klasifikasi Konten Berita Dengan Metode *Text Mining* . Medan: Universitas Sumatera Utara.
- 8) Negara, E. S., Andryani, R., & Saksono, P. H. (2016). Twitter data analytics: geospatial data extraction and analysis, *J. INKOM*, 10(1), 27-36.
- 9) Negara, E. S. (2017). Kajian Terhadap Tools dan Framework Social Media Analytics Untuk Pemanfaatan Data Social Media Dalam Penelitian Ilmu Sosial. *Jurnal Teknologi Technoscientia*, 9(2), 132-138.
- 10) Sanger. J.2007. *Text Mining Handbook:Advanced Approaches in AnalyzingUnstructured Data*. New York: Cambridge University Press.
- 11) *Sugiyono*. 2013. *Metode Penelitian Kuantitatif Kualitatif dan R & D*. Bandung :Alfabeta.
- 12) Sukmadinata.2006.*Metode Penelitian Pendidikan*, Bandung Remaja Rosda karya
- 13) Soemanagara. 2012. *Strategic Marketing Communication*. Bandung:Alfabeta
- 14) Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018). Improving Naïve Bayes in Sentiment Analysis For Hotel Industry in Indonesia. In 2018 Third International Conference on Informatics and Computing (ICIC) (pp. 1-6). IEEE.