

9.6%

PlagScan PRO Results of plagiarism
analysis from 01/25/2019, 9:19 AM

jurnalku (1).docx

Date: 01/25/2019, 9:17 AM



* All sources 17 | Internet sources 11 | Organization archive 2 | Plagiarism Prevention Pool 4

- ☒ [0]  https://www.researchgate.net/publication...ARITAS_TUJUAN_WISATA
2.1% 8 matches
- ☒ [1]  <https://www.idntimes.com/business/economy/yoshi/pertumbuhan-e-commerce-indonesia-1>
2.4% 5 matches
- ☒ [2]  <https://www.coursehero.com/file/p74chg2/...dibandingkan-dengan/>
1.4% 3 matches
- ☒ [3]  <https://jurnal-ppi.kominfo.go.id/index.php/jppi/article/view/070203/0>
0.5% 5 matches
- ☒ [4]  blog.stekom.ac.id/blogdata/view/e00da03b...n-20-Validasi-Silang
1.3% 1 matches
- ☒ [5]  <https://www.youtube.com/watch?v=uuqr9HWN9uo>
1.0% 3 matches
- ☒ [6]  www.academia.edu/35799751/ANALISA_SENTIM...PADA_APLIKASI_GO-JEK
0.9% 2 matches
- ☒ [7]  "13410100234-2017-Winned-JURNAL-fix3.docx" dated 2018-01-10
0.6% 4 matches
- ☒ [8]  "jurnal fix.docx" dated 2017-12-22
0.4% 3 matches
- ☒ [9]  https://www.researchgate.net/profile/Sup...n=publication_detail
0.5% 3 matches
- ☒ [10]  <https://id.123dok.com/document/z1drg0vz-...di-kasus-banjir.html>
0.2% 2 matches

-
- ✓ [11]  from a PlagScan document dated 2018-08-15 04:15
0.3% 1 matches
-
- ✓ [12]  <https://www.slideshare.net/sofiankevin/ssdsd-96384590>
0.4% 2 matches
-
- ✓ [15]  <https://docplayer.info/71976102-Medanbis...27-kopi-arabika.html>
0.2% 1 matches
-

9 pages, 3175 words

PlagLevel: 9.6% selected / 11.0% overall

28 matches from 17 sources, of which 11 are online sources.

Settings

Data policy: *Compare with web sources, Check against my documents, Check against my documents in the organization repository, Check against organization repository, Check against the Plagiarism Prevention Pool*

Sensitivity: *Medium*

Bibliography: *Consider text*

Citation detection: *Reduce PlagLevel*

Whitelist: --

ANALISIS SENTIMEN UNTUK PEMERINGKATAN POPULARITAS SITUS BELANJA ONLINE DI INDONESIA MENGUNAKAN METODE NAIVE BAYES (STUDI KASUS DATA SEKUNDER)

Firrari Kusuma Wardani ¹⁾ Valentinus Roby Hananto ²⁾ Vivine Nurcahyawati ³⁾

Program Studi/Jurusan Sistem Informasi
Institut Bisnis dan Informatika Stikom Surabaya

Jl. Raya Kedung Baruk 98 Surabaya, 60298

E-mail : 1)15410100048@stikom.edu, 2)valentinus@stikom.edu, 3)vivine@stikom.edu

Abstract: Lazada, Bukalapak and Blibli.com are companies involved in online markets and malls. To respond to opinions, criticisms and complaints, Lazada, Bukalapak and Blibli.com have a website platform that reviews the services and services offered by the three companies. The website platform is named priceprice.com. The type of comment is a large amount of unstructured text. This condition can cause online shopping companies to miss useful information from a collection of text documents. The setimen analysis is specified in this case study to rank the popularity of online shopping sites in Indonesia. In this study only discussed three online shops, namely Lazada, Bukalapak and Blibli.com. This case study uses the Naïve Bayes Classifier (NBC) method with k-fold cross validation testing to ensure data accuracy. The Naive Bayes method is used because it is more suitable in the operation of sentiment analysis in this study compared to other methods. In this study Lazada is the most popular online shop compared to Bukalapak and Blibli.com. Based on calculation accuracy naïve bayes uses k-fold cross validation for Lazada to have an average 96,9 %, Bukalapak 93,9% and Blibli 89,7%.

Keywords: Lazada, Bukalapak, Blibli.com^[2], Naive Bayes

marketer. Menurut penelitian lembaga riset e-commerce, populasi pengguna internet mencapai 83,7 juta orang pada tahun 2014 (e-marketer, 2014).^[2] Angka yang berlaku untuk setiap orang yang mengakses internet setidaknya satu kali setiap bulan itu menjadikan Indonesia sebagai peringkat ke-6 terbesar di dunia dalam hal jumlah pengguna internet.^[5] Pada tahun 2017, e-marketer memperkirakan netter Indonesia bakal mencapai 112 juta orang, mengalahkan Jepang di peringkat ke-5 yang pertumbuhan jumlah pengguna internetnya lebih lambat. Secara keseluruhan, jumlah pengguna internet di seluruh dunia diproyeksikan bakal mencapai 3 miliar orang pada tahun 2015.^[5] Tiga tahun setelahnya, pada 2018, diperkirakan sebanyak 3,6 miliar manusia di bumi bakal mengakses internet setidaknya sekali tiap satu bulan.^[1]

Definisi *marketplace* berdasarkan ilmu One data adalah fasilitator pembelian online yang tidak memiliki investasi barangnya sendiri.^[1] Sedangkan *e-commerce*

menyebut suatu kegiatan perdagangan melalui sistem elektronik seperti internet ataupun televisi.^[1]

Dihitung dari jumlah *unique audience* pada akhir triwulan dua, didapatkan data dari 10 peringkat teratas yang terdiri dari 4 e-commerce dan 6 marketplace.^[1] Lazada memimpin seluruh e-commerce dengan 21,2 juta *unique audience*.^[1] Urutan ke-10 peringkat teratas tersebut antara lain Lazada, Blibli, Tokopedia, Elevenia, Matahari Mall, Shopee, Bukalapak, Zalora, Qoo10, dan Blanja. Penulis memutuskan untuk memilih Lazada, Bukalapak, dan Blibli untuk dijadikan objek penelitian dikarenakan banyak terdapat ulasan mengenai tiga situs tersebut pada platform website yang dituju.

Pada kasus ini penulis hanya menekankan pada tiga situs belanja online yakni Lazada, Bukalapak, dan Blibli.

Komentar pengguna akan digunakan oleh penulis sebagai sumber data yang akan diolah untuk dijadikan wawasan dalam penelitian. Komentar pengguna didapat dari *web id.priceprice.com* yang memuat komentar positif maupun negatif. *Priceprice.com* adalah suatu platform website yang unggul dan kredibel dalam melakukan *review e-commerce*. Oleh sebab itu, penulis menggunakan platform terbuka sehingga *review* tersebut sesuai dengan kondisi asli produk atau jasa tersebut.

Penelitian terdahulu sebagai bahan rujukan penulis terdapat tiga penelitian yang berkaitan dengan *sentiment Analysis*, antara lain adalah penelitiab (Gamalo, Pablo, & Fernandez Lanza, 2013) mengenai *strategy naïve bayes* untuk *tweets* berbahasa spanyol.^[10] Kemudian penelitian (Vinodhini and Chandrasekaran, 2012) tentang pengembangan system yang dapat mengidentifikasi dan mengklasifikasi *sentiment* masyarakat untuk memprediksi produk yang menarik dalam pemasaran. Yang terakhir adalah penelitian (Munawarman, Ardiles Sinaga, 2017) mengenai pemanfaatan analisi *sentiment* untuk pemeringkatan popularitas tujuan wisata.^[10]

Pada penelitian ini, *sentiment analysis* akan digunakan untuk menentukan peringkat popularitas situs belanja online berdasarkan data komentar atau opini yang terdapat pada forum diskusi yang spesifik membicarakan mengenai hal ini. Dengan harapan dapat mencapai tujuan penelitian dalam penentuan popularitas situs

of training web pages." (2009), mengatakan bahwa "*Naïve Bayes Classifier* memiliki tingkat akurasi yg lebih baik dibanding model *classifier* lainnya".

Metode *Naive Bayes* digunakan karena lebih cocok dalam pengoperasian analisis sentimen pada penelitian ini dibandingkan dengan metode lainnya seperti *K-Nearest Neighbour* atau *learning machine*. *K-nearest neighbors* atau knn adalah algoritma yang berfungsi untuk melakukan klasifikasi suatu data berdasarkan data pembelajaran (*train data sets*), yang diambil dari k tetangga terdekatnya (*nearest neighbors*). Dengan menggunakan data, pembelajaran mesin memungkinkan komputer menemukan wawasan tersembunyi tanpa diprogram secara eksplisit saat mencarinya. Dengan adanya email baru, algoritma tersebut kemudian akan menghasilkan prediksi apakah email baru itu spam atau tidak.

Proses klasifikasi *Naïve Bayes Classifier* terhadap dokumen yaitu dengan mempresentasikan setiap dokumen dengan atribut " $X_1, X_2, X_3, \dots, X_n$ " yang mempunyai makna bahwa X_1 untuk kata pertama, X_2 adalah kata kedua, dan seterusnya. Untuk himpunan kategori komentar dipresentasikan dengan V . Saat melakukan proses klasifikasi dokumen, *Naïve Bayes Classifier* akan mencari nilai probabilitas tertinggi dari [4]:

belanja online di Indonesia. Manfaat yang dapat diambil pada penelitian ini, yakni pengguna aplikasi belanja online dapat lebih bijak dalam menentukan situs mana yang aman dalam proses transaksinya. Selain itu, pihak pemilik situs juga dapat meningkatkan kinerja layanannya dengan adanya analisis ini.

Pada penelitian ini, penulis memutuskan untuk menggunakan metode Naive Bayes (NB) karena merupakan metode yang sederhana dan intuitif dan kinerjanya mirip sekali dengan pendekatan lain. Selain itu, berdasarkan penelitian terdahulu, NB menggunakan efisiensi yang cukup optimal dan akurat.

Naive Bayes Classifier

Naive Bayes Classifier bekerja sangat baik dibanding dengan model *classifier* lainnya. Hal ini dibuktikan pada jurnal Xhemali, Daniela, Chris J. Hinde, and Roger G. Stone. “*Naive Bayes vs. decision trees vs. neural networks in the classification*

$$V_{MAP} = \underset{V_j \in V}{\operatorname{argmax}} \frac{P(x_1, x_2, x_3, \dots, x_n | V_j) P(V_j)}{P(x_1, x_2, x_3, \dots, x_n)}$$

Keterangan:

V_{MAP} : semua kategori yang diujikan

V_j : Kategori komentar, dengan:

J_1 : Sentimen positif

J_2 : Sentimen Negatif

J_3 : Sentimen Netral

$P(X_i | V_j)$: probabilitas X_i pada kategori V_j

$P(V_j)$: probabilitas dari V_j

K-Fold Cross-Validation

Model yang didapatkan dari naïve bayes kemudian dilakukan pengujian menggunakan *k-fold cross validation*. *Cross-validation* adalah bentuk sederhana dari teknik statistik. Jumlah fold standar untuk memprediksi tingkat error dari data adalah dengan menggunakan *10-fold cross validation* (Witten, et al, 2011: 153). Data yang

digunakan dibagi secara acak ke dalam k subset yaitu $D_1, D_2, \dots, D_k$ dengan ukuran yang sama. Dataset akan dibagi menjadi data *training* dan data *testing*. Proses training dan testing dilakukan sebanyak k kali secara berulang-ulang. Pada iterasi ke-i, partisi D_i disajikan sebagai data testing dan partisi sisanya digunakan secara bersamaan dan berurutan sebagai data *training*.

Rapid Miner

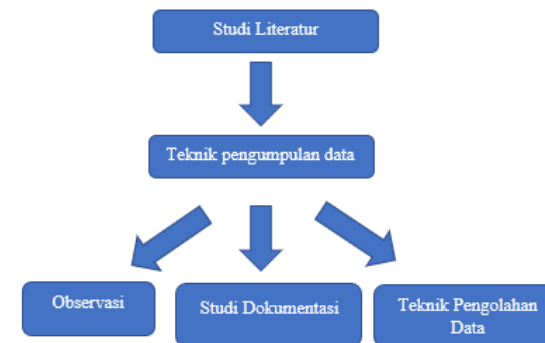
Aplikasi yang digunakan pada penelitian ini adalah Rapidminer yakni aplikasi data mining yang tidak perlu ditanyakan lagi dan berbasis *opensource*. Tersedia sebagai aplikasi mandiri yang digunakan untuk analisis data dan sebagai mesin data mining untuk integrasi ke dalam produk sendiri.

Lebih dari 40 negara di dunia telah menggunakan ribuan aplikasi rapidminer yang memberikan keunggulan yang kompetitif. Solusi yang diberikan antara lain adalah : Analisis ETL, Data Analisis, Integrasi Data, dan pelaporan dalam satu suite tunggal.

Metodologi Penelitian

Dalam tahapan penelitian ini akan melalui tiga tahapan yaitu tahap awal penelitian, tahap analisis dan tahap akhir.

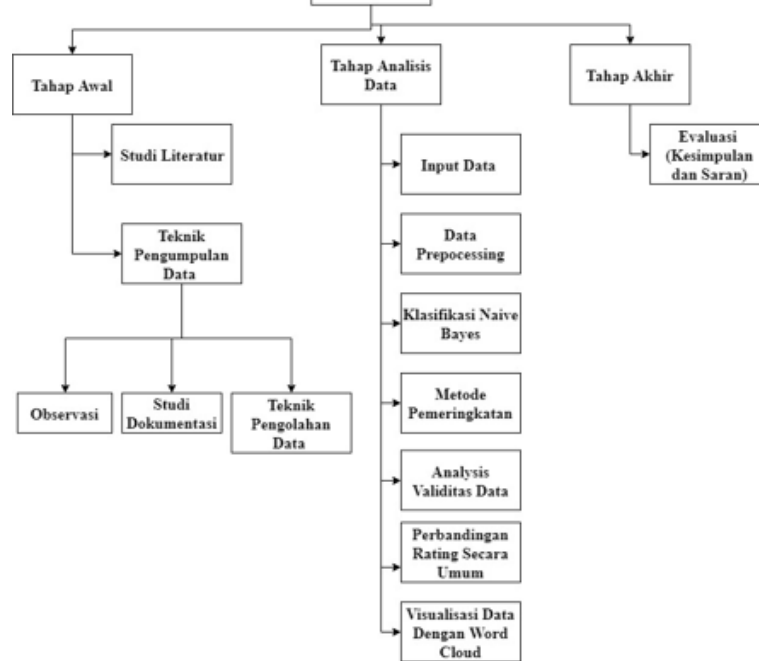
Hal yang dilakukan pada penelitian ini:



Gambar 2. Bagan Tahapan Awal (Wardani, 2019).

Studi literatur ini dilakukan untuk mengetahui tren penelitian yang terjadi pada bidang analisis sentimen, studi literatur ini dilakukan untuk mengulas jurnal-jurnal yang telah ada dan digunakan untuk menentukan studi kasus serta metode yang akan digunakan oleh penulis saat melakukan penelitian.

Pengumpulan data ini dilakukan untuk memperkuat alasan mengapa penelitian ini harus dilakukan, pengumpulan data diambil dari sumbernya langsung yaitu *website* resmi *id.priceprice.com* serta komentar pengguna aplikasi Lazada,



Gambar 1. Diagram metode penelitian (Wardani, 2019)

Bukalapak, dan Blibli.com. Observasi ini dilakukan untuk mengamati objek yang akan diteliti. Pada penelitian ini dilakukan dengan mengamati komentar pengguna situs belanja online, adanya banyak macam komentar yang diberikan oleh pengguna pada aplikasi ini yang membuat penulis melakukan penelitian dalam presentase klasifikasi komentar masyarakat yang ada pada *website id.priceprice.com*.

Pada studi dokumentasi ini penulis mengumpulkan data sebagai berikut :

- a. Kurva dari transaksi dan pengguna aplikasi Lazada, Bukalapak dan Blibli.com.
- b. Klasifikasi yang diberikan pengguna dalam melakukan perbandingan.

- c. Ranging aplikasi Lazada, Bukalapak dan Blibli.com.

Data yang dibutuhkan pada penelitian ini adalah link dari marketplace yang dijasikan sebagai objek penelitian. Data link yang dimaksud berbentuk excel (.xls). Data *excel* yang didapat akan diolah pada aplikasi *RapidMiner* untuk mengetahui hasil selanjutnya. Berikut dapat dilihat gambar dari data *excel* yang akan diolah pada aplikasi *RapidMiner*.^[9]

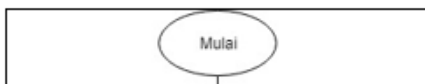
Data yang digunakan dalam penelitian ini dibagi menjadi data *testing* sebanyak 90% dan data *training* sebanyak 10% menggunakan *10-fold cross validation*.

Variabel	Keterangan	Skala Data
Y	Sentiment (Positif/Negatif) 0 = Sentiment Positif 1 = Sentimen Negatif	Nominal
X_k	Kata Kunci (Frekuensi Kemunculan Kata ke-k pada komentar ke-n)	Rasio

Gambar 3. Variabel Penelitian (Wardani,2019).

Diagram Alir

Berikut adalah diagram alir dari penelitian ini :

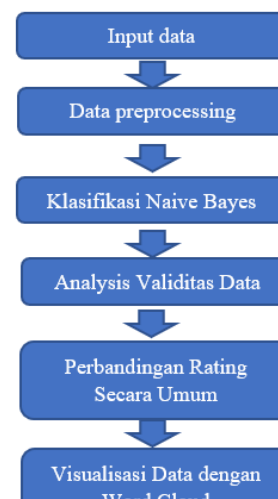


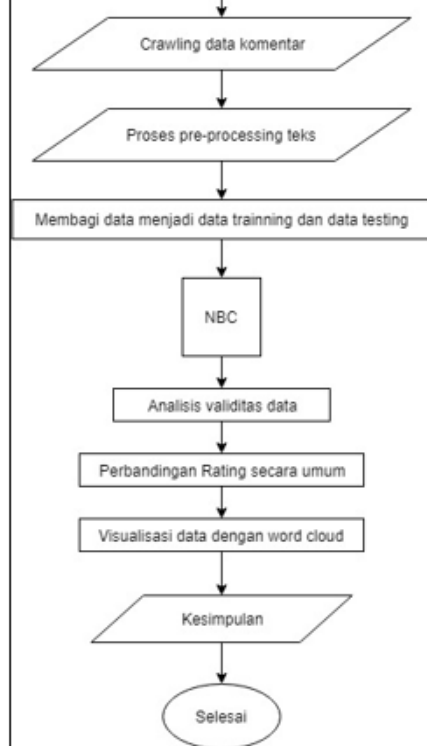
NBC (Naive Bayes Classification)



Gambar 5. Diagram Alir NBC.(Wardani, 2019)

Tahap Analisis Data





Gambar 4. Diagram Alir (Wardani,2019)

Gambar 6. Teknik Analisis Data (Wardani,2019).

Uji Validitas Data

Pada tahap analisis validitas data penujian dilakukan dengan *k-fold cross validation*. Berikut adalah tahapan yang akan dilakukan dalam penghitungan menggunakan cross validation pada Rapidminer :

Cross validation dilakukan dengan membagi contoh yang ditetapkan ke dalam bagian yang sama dan berputar memulai semua bagian, selalu menggunakan satu untuk pengujian dan yang lainnya untuk menghitung keakuratan dan harus menjadi pendekatan estimasi standar kita bila ada upaya penghitungan tambahan yang layak dilakukan.

Perbandingan Rating Secara Umum

Pada langkah berikut ini akan ditentukan rating dari seluruh kegiatan yang telah dilakukan mulai dari *crawling* data hingga testing yang telah dilakukan. Pada tahap ini akan dihitung presentase komentar negatif dan positif yang didapatkan dari aplikasi *Rapidminer*.

- Skala yang digunakan: 1 – 3.
- Masing-masing dihitung presentase nilai positif dari data yang ada.
- Rating dihitung berdasarkan jumlah seluruh presentase nilai positif.
- Lalu dibuat perankingannya.

Visualisasi Data dengan Word Cloud

Langkah berikutnya yang harus dilakukan adalah memvisualisasikan data dengan *wordcloud*. Dalam penelitian ini menggunakan aplikasi *Tagcrowd* sebagai visualisasi.



Tabel 1. Rating Secara Umum Komentar *Priceprice*.

Rank	Situs	Comment Count (cc)	Rating
1	Lazada	673	1
2	Bukalapak	298	2
3	Blibli.com	75	3

Sumber : Wardani, 2019.

Tabel 1 menjelaskan bahwa secara *Comment Count* (cc) Lazada menempati peringkat 1 untuk *online shop* yang paling banyak disebut selama kurun waktu 2 tahun terakhir.

II. Berikut adalah rating untuk komentar positif paling banyak berdasarkan komentar pengguna pada *website priceprice.com* setelah proses uji validitas data :

Tabel 2. Rating secara umum berdasarkan komentar positif.

Rank	Situs	Positive Comments (pc)	Rating

1	Lazada	84%	1
2	Bukalapak	93%	2
3	Blibli.com	85%	3

Sumber : Wardani, 2019.

HASIL DAN PEMBAHASAN

Berikut adalah hasil dari setiap proses yang telah dilakukan dalam penelitian ini :

- I. Berikut adalah paparan rating secara umum bila dilihat berdasarkan komentar pengguna website priceprice.com ^[15] selama kurun waktu 2 tahun terakhir :

Tabel 2 menjelaskan bahwa secara *Positive Comments* (pc) Lazada menempati peringkat 1 untuk *online shop* yang paling banyak mendapatkan komentar positif oleh pengguna website priceprice.

- III. Berikut adalah rating untuk komentar negatif paling banyak berdasarkan komentar pengguna pada *website priceprice.com* setelah proses uji validitas data :

Tabel 3.^[10] **Rating Secara Umum Berdasarkan Komentar Negatif.**

Rank	Situs	Negative Comments (nc)	Rating
1	Lazada	6%	3
2	Bukalapak	7%	2
3	Blibli.com	15%	1

Sumber : Wardani, 2019.

Tabel 3 menjelaskan bahwa secara *Negative Comments* (nc) Blibli.com menempati peringkat 1 untuk *online shop* yang paling banyak mendapatkan komentar negatif oleh pengguna *website priceprice*.

- IV. Berikut adalah rating untuk situs belanja online yang paling populer dan paling banyak disebut pada kolom komentar *website priceprice.com* :

Tabel 4. Rating Secara Umum Berdasarkan Paling Populer.

Rank	Situs	Popular Online Shop (pos)	Rating
1	Lazada	92,9%	1
2	Bukalapak	92,3%	2

aman menurut komentar pengguna *website priceprice*.

Tabel 5. Rating secara umum berdasarkan paling aman.

Rank	Situs	Tingkat Keamanan (TK)	Rating
1	Lazada	79%	3
2	Bukalapak	93%	1
3	Blibli.com	92%	2

Sumber : Wardani, 2019.

- VI. Berikut adalah rating untuk situs belanja online yang paling memuaskan dan paling banyak disebut pada kolom komentar *website priceprice.com* :

Tabel 6. Rating secara umum berdasarkan kepuasan pelanggan.

Rank	Situs	Kepuasan Pelanggan (KP)	Rating
1	Lazada	91%	1
2	Bukalapak	69%	3
3	Blibli.com	58%	2

Sumber : Wardani, 2019.

Tabel 6 menjelaskan bahwa secara Kepuasan Pelanggan (KP) Lazada

Tabel 4 menjelaskan bahwa secara *Popular Online Shop* (pos) Lazada menempati peringkat 1 untuk *online shop* yang paling populer dibandingkan dengan bukalapak dan Blibli.com menurut komentar pengguna *website priceprice.com*.

- V. Berikut adalah rating untuk situs belanja online yang paling aman dan paling banyak disebut pada kolom komentar *website priceprice.com* :

Tabel 5. Rating Secara Umum Berdasarkan Tingkat Keamanan.

Tabel 5 menjelaskan bahwa secara Tingkat Keamanan (TK) Bukalapak menempati peringkat 1 untuk *online shop* yang paling

menempati peringkat 1 untuk *online shop* yang paling memuaskan berdasarkan komentar pengguna *website priceprice*.

Visualisasi data dengan wordcloud

Pendapat dari konsumen belanja online di Lazada, Bukalapak dan Blibli.com dapat divisualisasikan dalam bentuk *Wordcloud*. *Wordcloud* merupakan representasi grafis dari sebuah dokumen yang dilakukan dengan plotting kata-kata yang sering muncul pada sebuah dokumen pada ruang dua dimensi. Frekuensi dari kata yang sering muncul ditunjukkan melalui ukuran huruf kata tersebut. Semakin besar ukuran kata menunjukkan semakin besar frekuensi kata tersebut muncul dalam dokumen. Berikut adalah tampilan visualisasi data dengan *Wordcloud* berdasarkan hasil uji validitas.

A. Lazada



Gambar 8. Visualisasi Wordcloud Lazada
(Wardani,2019)

Berdasarkan Gambar 8 dapat diketahui bahwa tiga kata yang sering muncul pada sentimen positif akun Lazada adalah Lazada, barang, dan belanja. Peningkatan pelayanan dan kecepatan kepada pelanggan dari *online shop* Lazada perlu dipertahankan dan diberi apresiasi. Selain respon positif, respon negatif juga perlu diketahui oleh Lazada untuk meningkatkan pelayanan.

B. Bukalapak

C.



D. Blibli.com



Gambar 10. Visualisasi Wordcloud
Blibli(Wardani,2019)

Berdasarkan Gambar 10 dapat diketahui bahwa tiga kata yang sering muncul pada sentimen positif akun Blibli.com adalah blibli, barang, dan aman. Keamanan dalam bertransaksi Blibli.com perlu dipertahankan dan diberi apresiasi. Selain respon positif, respon negatif juga perlu diketahui oleh Blibli.com untuk meningkatkan kepuasan pelanggan dalam berbelanja.^[8]

KESIMPULAN

Berdasarkan hasil analisis yang telah diperoleh, kesimpulan yang didapatkan dari penelitian ini adalah sebagai berikut.

1. Konsumen yang memberikan tanggapan, pendapat, kritik, saran dan masalah complain lebih banyak ditujukan pada *online shop* Blibli.com dengan presentase

Gambar 9. Visualisasi Wordcloud
Bukalapak(Wardani,2019)

Berdasarkan Gambar 4.12 dapat diketahui bahwa tiga kata yang sering muncul pada sentimen positif akun Bukalapak adalah bukalapak, belanja, dan barang. Peningkatan pelayanan tentang bantuan dan memberikan informasi kepada masyarakat dari *online shop* Bukalapak perlu dipertahankan dan diberi apresiasi. Selain respon positif, respon negatif juga perlu diketahui oleh Bukalapak untuk meningkatkan keamanan pelanggan dalam bertransaksi.

komentar negatif 15% disusul dengan Bukalapak 7% dan Lazada sebanyak 6%. Jumlah sentimen data pada ketiga *online shop* bersifat imbalanced, sehingga ketepatan klasifikasi perlu dilihat berdasarkan uji validitas data.

2. Berdasarkan visualisasi *wordcloud*, penulis menganbil 3 sentimen positif teratas yang banyak dibicarakan oleh konsumen pada *online shop* :
 - a) Lazada : baik, murah, cepat.
 - b) Bukalapak : terima kasih, barang, belanja.
 - c) Blibli.com : aman, puas, mudah.
3. *Online Shop* yang paling populer menurut penelitian ini adalah Lazada dengan rata-rata penghitungan manual, sentimen positif 95%, disusul dengan Bukalapak dengan rata-rata sentimen positif 94%,

kemudian di posisi paling akhir adalah Blibli.com dengan rata-rata sentimen 88%.

4. Berdasarkan perhitungan *naive bayes* didapatkan persentase untuk situs belanja *online* Lazada memiliki 93% kalimat positif dan 7% kalimat negatif, Bukalapak memiliki 92% kalimat positif dan 8% kalimat negatif, sedangkan Blibli memiliki 82% kalimat positif dan 18% kalimat negatif.
5. Berdasarkan pemeringkatan yang dilakukan, didapatkan bahwa *online shop* Lazada memimpin sebagai *online shop* paling populer diantara Bukalapak dan Blibli.com menurut data dari *platform website* priceprice.com dan metode perhitungan *naive bayes*.

[7]

SARAN

Berdasarkan hasil analisis yang telah diperoleh, saran yang didapatkan dari penelitian ini adalah sebagai berikut.

1. Pada penelitian klasifikasi teks, membutuhkan data sentimen yang cukup besar untuk mengurangi resiko dari imbalanced data. Pengklasifikasian sentimen awal sebaiknya dilakukan dengan banyak

peringkat kedua. Perlu diperhatikan juga untuk keamanan dalam transaksi.

3. Blibli : Sejauh ini Blibli.com menempati peringkat 1 untuk *online shop* yang paling aman dalam transaksinya. Namun, masih jauh dibelakang dibanding dengan Lazada dan Bukalapak untuk tingkat popularitas, dan masih banyak ulasan-ulasan negatif mengenai Blibli.com. Diharapkan ada suatu pembenahan untuk meningkatkan kualitas layanan dan lebih mengutamakan kepuasan pelanggan.

RUJUKAN

- Alpaydin, Ethem. (2014). *Introduction to Machine Learning, Third Edition*. London: The MIT Press.
- E-marketer. (2014). *Pengguna Internet Indonesia Nomor Enam Dunia*. Tersedia di <http://tekno.kompas.com>. (diakses pada 24/07/2018).
- Ervinzein. (2016). *E-commerce-blibli-com*. Tersedia di <https://Ervinzeintugas.wordpress.com>. (diakses pada 24/07/2018).
- Gamallo, Pablo, & Fernandez-Lanza. (2013). *TASS: A Naive-Bayes strategy for*

- orang sehingga dapat mengurangi obyektifitas pada hasil sentimen.
2. Pada proses stemming, perlu menambahkan kata kunci khusus untuk Bahasa sehari-hari dikarenakan pada data komentar banyak menggunakan bahasa sehari-hari.

Kemudian, saran untuk *online shop* yang dibahas dalam penelitian ini adalah :

1. Lazada : Pertahankan pencapaian hingga kini. Karena Lazada menempati peringkat pertama untuk online shop yang paling populer seperti dibahas dalam penelitian ini. Namun, perlu diperhatikan untuk keamanan dalam transaksinya.
2. Bukalapak : Bukalapak menempati peringkat kedua dalam penelitian ini. Pertahankan dalam menjaga kepuasan pelanggan karena pada hal ini Bukalapak menempati

sentiment analysis on Spanish tweets. In In Workshop on Sentiment Analysis at SEPLN (TASS2013) (pp 126–132). Madrid, Spain.

L. Whitten, Jeffrey, & D. Bentley, Lonnie .(2007). *System Analysis & Design Methods Seventh Edition*. New York, USA : McGraw-Hill

Liu, *Sentiment Analysis and Opinion mining*, Morgan & Claypool Publisher, (2012).

Makinudin dan Hadiyanto (2006). *Analisis Sosial*. Bandung : Yayasan Akatiga.

McLeod, Pearson (2008). *Sistem Informasi Manajemen*. Salemba. Jakarta.

Ni Wayan Sumartini Saraswati (2011). *Text Mining Dengan Metode Naïve Bayes Classifier Dan Support Vector Machines Untuk Sentiment Analysis*. Tesis [Online]. Tersedia di <http://www.pps.unud.ac.id/thesis/pdf>

- _thesis/unud-209-236721286-tesis.pdf. (diunduh pada 13/09/2015).
- Pang, B. dan Lee, L., 2008, *Opinion Mining And Sentiment Analysis*, Foundations and Trends in Information Retrieval, 2, 1-2, 1-135.
- Rizqiyah, Putri. (2018). *Analisa Sentimen Menggunakan Naïve Bayes Untuk Mengetahui Presentase Komentar Pada Aplikasi Go-Jek*. Program Studi Informatika, Institut Teknologi Telkom Purwokerto
- Sanianto. (2014). *Profil Perusahaan Besar Lazada*. Tersedia di <http://irvanarifsanianto.blogspot.com/2014/11/perusahaan-lazada.html>
- Silviosilviando. (2016). *Analisis web e-commerce bukalapak*. Tersedia di <https://silviosilviando.wordpress.com>(diakses pada 23/5/2018).
- Triawati, C (2009). *Metode Pembobotan Statistical Concept Based Untuk Klastering dan Kategorisasi Dokumen Berbahasa Indonesia*. Institut Teknologi Telkom Bandung.
- Vinodhini, G. And RM. Chandrasekaran (2012). *"Sentiment Analysis and Opinion Mining: a Survey"*. International Journal of Advanced Research in Computer Science and Software Engineering, Vol.2, No.6, pp.283-192.
- Wong, Jony. (2010). *Internet Marketing for Beginners*. PT Elex Media Komputindo. Jakarta.

