



**ANALISIS SENTIMEN TERHADAP REVIEW E-COMMERCE
DENGAN METODE *STOCHASTIC GRADIENT DESCENT***

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Statistika

JURNAL ILMIAH

Oleh

**Nur Khotimah
B2A218022**

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS MUHAMMADIYAH SEMARANG
2019**

ANALISIS SENTIMEN TERHADAP REVIEW E-COMMERCE DENGAN METODE STOCHASTIC GRADIENT DESCENT

Nur Khotimah¹, Moh. Yamin Darsyah², Indah Manfaati Nur³

123Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

e-mail : nurkhotimah111@gmail.com

ABSTRAK

Bukalapak dan Shopee adalah e-commerce yang sudah memiliki aplikasi mobile phone dan website, sehingga dapat memudahkan konsumen untuk memberi review. Banyaknya review memerlukan sebuah metode atau teknik khusus yang mampu mengkategorikan review tersebut secara otomatis, apakah termasuk review positif atau negatif. Sentiment Analisis adalah cara mengidentifikasi dan mengelompokkan polaritas teks yang diberikan ditingkat dokumen, kalimat dan frasa. *Stochastic Gradient Descent* adalah metode klasifikasi text mining yang mampu diterapkan pada data dengan jumlah besar. Hasil yang didapatkan adalah hampir sebagian konsumen memberikan penilaian positif kepada Shopee dan Bukalapak. Namun bukhalapak memiliki penilaian positif terbanyak yakni 91.37% walaupun komentar yang diberikan juga cukup banyak. Tidak hanya itu, bukhalapak juga mendapatkan rating paling banyak dibandingkan shopee yakni 83%. Kata-kata yang paling banyak ditulis pada review negatif antara lain aplikasi lambat, perlu di update, ribet, mudah error, dan aplikasi berat. Sehingga dapat menjadi pertimbangan bagi bukhalapak dan shopee.

Kata Kunci : Analisis Sentimen, E-Commerce, Klasifikasi, *Stochastic Gradient Descent*

ABSTRACT

Bukalapak and Shopee are e-commerce sites that already have mobile phone applications and websites, so they can make it easier for consumers to give reviews. The number of reviews requires a special method or technique that is able to categorize the review automatically, including positive or negative reviews. Sentiment Analysis is a way to identify and classify the polarity of a given text at the level of documents, sentences and phrases. Stochastic Gradient Descent is a text mining classification method that can be applied to large amounts of data. The results obtained are that most consumers give positive ratings to Shopee and Bukalapak. However, Bukalapak has the most positive ratings, which is 91.37%, although there are also quite a lot of comments. Not only that, Bukalapak also got the most ratings compared to Shopee which is 83%. The words that are mostly written on negative reviews include slow applications, need to be updated, complicated, easy errors, and heavy applications. So it can be a consideration for bukhalapak and shopee

Keywords : *Analysis of Sentiment, E-Commerce, Classification, Stochastic Gradient Descent*

PENDAHULUAN

Teknologi menjadi alat yang membantu kebutuhan manusia, dengan teknologi, apapun dapat dilakukan dengan lebih mudah. Begitu pentingnya peran teknologi inilah yang mulai membawa peradaban memasuki

ke era digital. Banyak perusahaan dengan tergesa-gesa menggunakan teknologi informasi dan komunikasi termasuk perusahaan yang bergerak dibidang perdagangan (Karami, 2014). Pemanfaatan teknologi informasi dan komunikasi di

bidang perdagangan berupa penjualan dan pembelian barang secara online. Sebuah survei yang diselenggarakan Asosiasi Penyelenggaraan Jasa Internet Indonesia (APJII) menunjukkan bahwa jumlah pengguna internet di Indonesia meningkat tiap tahunnya hingga akhir tahun 2018 mencapai 171,17 juta orang, dari total penduduk sebanyak 264,16 juta orang. Perkembangan tersebut berdampak pada berbagai bidang. Salah satunya ialah maraknya kegiatan berbelanja melalui media internet. Menurut survei yang dilakukan oleh APJII 47,6% dari pengguna internet di Indonesia mengetahui internet sebagai tempat jual beli barang dan jasa, dan 41,2% pernah melakukan transaksi secara online. Istilah yang digunakan dalam transaksi jual beli secara online adalah e-commerce (APJII, 2018).

Indonesia telah menjadi pasar terbesar e-commerce di Asia Tenggara, penjualan online Indonesia mencapai US\$ 1,1 miliar, lebih tinggi dari Thailand dan Singapura. APJII pada 2018 juga melakukan survei dua konten internet komersial teratas yang sering digunakan untuk membeli barang atau jasa online yakni Shopee 11,2% dan Bukalapak 8,4%. Bukalapak dan Shopee merupakan model e-commerce Customer-To Customer (C2C). Pasar Customer-To Customer (C2C) saat ini masih dominan di pasar ritel online Indonesia. Bukalapak dan Shopee sudah memiliki aplikasi mobile phone dan website, yang memberikan akses kepada pengguna untuk memberikan ulasan terkait dua e-commerce tersebut (Iprice, 2019).

Menurut Sung, bahwa penelitian baru-baru ini menemukan hampir 50% dari pengguna internet bergantung pada rekomendasi word-of-mouth (opini) sebelum menggunakan suatu produk, karena review dari pengguna lain dapat menyediakan informasi terbaru dari produk tersebut berdasarkan perspektif pengguna-pengguna lain yang sudah menggunakan produk

tersebut (Sung & Sung, 2014). Opini yang dimuat di media sosial jumlahnya terlalu banyak untuk diproses secara manual. Oleh sebab itulah, diperlukan sebuah metode atau teknik khusus yang mampu mengkategorikan review-review tersebut secara otomatis, apakah termasuk positif atau negatif (Potdar, Patil, Bagla, Pandey, & Jadhav, 2016).

Sentiment Analysis (SA) adalah cara mengidentifikasi dan mengelompokkan polaritas teks yang diberikan ditingkat dokumen, kalimat dan frasa (Kang, Yoo, & Han, 2012). Teknik ini digunakan di banyak bidang seperti e-commerce, perawatan kesehatan, hiburan dan politik, untuk beberapa nama. Sebagai contoh, Analisis Sentimen berguna bagi perusahaan untuk memantau opini konsumen mengenai produk mereka, dan bagi konsumen untuk memilih produk terbaik berdasarkan opini publik. Tantangan utama analisis sentimen adalah ulasan umumnya ditulis dalam bahasa informal, pendek pesan menunjukkan isyarat terbatas tentang sentimen serta akronim dan singkatan banyak digunakan. Secara umum, teknik klasifikasi dasar seperti Naïve Bayes (Multinomial and Bernoulli), Logistic Regression, SGD (*Stochastic Gradient Descent*), Linear SVM (Support Vector Machine) dan Random Forest (Mohaiminul & Sultana, 2018)

Penelitian yang dilakukan oleh Saraswati (2011) mengklasifikasikan opini dengan menggunakan metode NBC dan SVM. Dari hasil penelitiannya, Saraswati menyatakan bahwa metode SVM memiliki tingkat akurasi yang lebih tinggi daripada metode NBC untuk pengujian data opini positif, sedangkan metode NBC menunjukkan hasil yang lebih baik jika digunakan dalam pengujian data opini negatif.

Putranti dan Winarko (2014) tentang analisis sentimen dalam penelitian ini proses klasifikasi dokumen tekstual dibagi kedalam

dua kelas, yaitu kelas sentimen positif dan negatif bertujuan untuk menentukan sentimen publik terhadap objek tertentu yang disampaikan di Twitter dalam bahasa Indonesia.

Penelitian yang dilakukan oleh Alamsyah dkk (2017) analisis sentimen untuk mengetahui bagaimana perbandingan kepuasan pelanggan terhadap tiga situs e-commerce yang sering dikunjungi di Indonesia yaitu Bukalapak, Tokopedia dan Elevenia. Data yang didapatkan dari media sosial twitter, dalam penelitiannya Naïve Bayes Classifier digunakan sebagai teknik klasifikasi dengan pembobotan TF-IDF, sedangkan untuk melakukan validasi dan evaluasi pada klasifikasi teks Naïve Bayes dilakukan menggunakan K-fold cross validation dan confusion matrix. Hasil penelitian menunjukkan bahwa sentimen negatif terhadap ketiga situs e-commerce tersebut lebih mendominasi di media sosial, dan situs e-commerce dengan sentimen negatif paling tinggi ialah Bukalapak kemudian Tokopedia dan Elevenia (Alamsyah & Saviera, 2017).

Mohaiminul Islam dan Sultana tahun 2018 juga melakukan komparasi beberapa metode untuk klasifikasi analisis sentimen. Metode yang digunakan antara lain Multinomial Naïve Bayes, Bernouli Naïve Bayes, Logistic Regression, Linier SVM, *Stochastic Gradient Descent* (SGD) dan Random Forest. Berdasarkan hasil yang didapat untuk kedua dataset Linier SVM yang memiliki tingkat akurasi paling tinggi, yang selanjutnya *Stochastic Gradient Descent* memiliki tingkat akurasi tertinggi kedua (Mohaiminul & Sultana, 2018).

Gusriani, dkk (2016) melakukan analisis sentimen berdasarkan komentar publik terhadap toko online seperti Zalora dan Berry Benka pada media sosial facebook. Metode klasifikasi Naive Bayes, K-NN dan Decision Tree. Hasil analisis pengujian

menunjukkan Naive Bayes, memiliki kestabilan akurasi setelah diuji dengan beberapa nilai Frequent Itemset. Naive Bayes memiliki rata-rata akurasi 90,3%.

Moh dkk pada tahun 2015 juga melakukan penelitian mengenai multi-tier classification untuk analisis sentimen dengan menggunakan 150.000 review film yang menggunakan beberapa metode antara lain Naïve Bayes, SVM (Support Vector Machine), Random Forest, dan *Stochastic Gradient Descent*. Berdasarkan hasil yang didapat *Stochastic Gradient Descent* memiliki akurasi paling tinggi saat dilakukan multi-tier sentimen analisis (Moh, Gajjala, Gangireddy, & Moh, 2016).

Stochastic Gradient Descent merupakan salah satu tipe dari Gradient Descent. Penerapan SGD dalam proses klasifikasi memiliki kelebihan mampu mengoptimasi dalam data yang begitu besar sebab *Stochastic Gradient Descent* menemukan gradien fungsi biaya dari satu contoh pada setiap iterasi, bukan jumlah gradien fungsi biaya dari semua contoh, selain itu SGD mampu meningkatkan kinerja generalisasi masalah berskala besar, Mampu diterapkan pada data dengan jumlah besar dan memiliki missing value dan waktu komputasi relatif lebih singkat meskipun datanya besar dan dari beberapa penelitian memiliki tingkat akurasi yang lebih tinggi dari algoritma klasifikasi pada analisis sentimen lainnya (Sharma, 2018). Oleh karena itu pada penelitian ini penulis akan melakukan Analisis Sentimen Terhadap Review E-Commerce dengan Metode *Stochastic Gradient Descent* dengan studi kasus Shopee dan Bukalapak.

TINJAUAN PUSTAKA

1. *Stochastic Gradient Descent*

Stochastic Gradient Descent (SGD) merupakan metode *Gradient Descent* yang melakukan update parameter untuk setiap

data pelatihan $x^{(i)}$ serta label $y^{(i)}$ dan memiliki persamaan dasar sebagai berikut.

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta; x^{(i)} y^{(i)})$$

Model Stochastic Gradient Descent atau SGD merupakan model dalam deep learning yang mengoptimasi fungsi dengan mengikuti gradient yang memiliki nosy dengan ukuran langkah yang menurun (Dong & Chen, 2009). Stochastic Gradient Descent menggunakan gradien stochastic yang meminimalkan fungsi kerugian yang dipilih dengan fungsi linear. Algoritma ini mendekati gradien yang benar dengan mempertimbangkan satu sampel pada suatu waktu, dan secara bersamaan memperbarui model berdasarkan gradien fungsi kerugian.

2. Text Mining

Text mining atau text analytics adalah istilah yang mendeskripsikan sebuah teknologi yang mampu menganalisis data teks semi-terstruktur maupun tidak terstruktur, hal seperti itu yang membedakannya dengan data mining dimana data mining mengolah data yang sifatnya terstruktur (Jamil, 2017).

Perbedaan antara text mining dengan data mining terletak pada sumber data yang digunakan. Dalam text mining pola-pola yang diekstrak dari data tekstual yang tidak terstruktur bukan berasal dari suatu database. Beberapa kesamaannya adalah data yang digunakan merupakan data besar dan data berdimensi tinggi dengan struktur yang terus berubah.

3. Analisis Sentimen

Analisis sentimen dapat digunakan dalam berbagai kemungkinan domain, dari produk konsumen, jasa kesehatan, jasa keuangan, peristiwa sosial dan politik pada pemilu. Kecendrungan penelitian tentang analisis sentimen berfokus pada pendapat yang menyatakan atau menyiratkan suatu sentimen positif atau negatif. Pendapat mewakili hampir semua aktivitas manusia, karena pendapat dapat mempengaruhi terhadap perilaku seseorang. Setiap kali kita perlu membuat keputusan, kita ingin tahu pendapat orang lain. Dalam dunia nyata, bisnis dan organisasi selalu ingin melihat opini publik tentang suatu produk

atau jasa (Liu, 2012). Dengan analisis sentimen, suatu bisnis dapat melacak produk-produk, merek dan orang-orang misalnya dan menentukan apakah dilihat positif atau negatif di web. Hal ini memungkinkan bisnis untuk mengetahui komentar buruk, persepsi produk baru dan persepsi terhadap suatu merek tertentu.

4. Klasifikasi

Teknik klasifikasi adalah salah satu dari teknik data mining yang termasuk supervised learning. Supervised learning artinya proses pembentukan sebuah korespondensi menggunakan sebuah training dataset. Tujuannya adalah untuk memprediksi target dari beberapa atribut (Zaki & Meira, 2014). Terdapat pada dua pekerjaan utama pada klasifikasi yaitu melakukan training untuk disimpan sebagai prediksi dan melakukan testing untuk proses klasifikasi agar diketahui di label mana objek data tersebut (Liu, Loh, & Sun, 2009).

5. e-Commerce

Electronic commerce (disingkat e-commerce) sebagai sarana berbisnis menggunakan jaringan komputer, sebenarnya adalah dikenal sejak 20 tahun lalu sejak akhir tahun 70-an dan awal tahun 80-an. Generasi pertama e-commerce dilakukan hanya antar perusahaan berupa transaksi jual beli yang difasilitasi oleh Electronic Data Intechange (EDI) dalam transaksi jual beli elektronik ini banyak aspek-aspek yang bersentuhan langsung maupun tidak langsung (Firdaus, 2015). E-commerce adalah proses pembelian dan penjualan antara dua belah pihak di dalam suatu perusahaan dengan adanya pertukaran barang, jasa, atau informasi melalui media internet (Indrajit, 2001). Onno (2000) memberikan pengertian tentang e-commerce yaitu asset dinamis teknologi, aplikasi, dan proses bisnis yang menghubungkan perusahaan, konsumen dan komunitas melalui elektronik dan perdagangan barang, pelayanan dan informasi yang dilakukan secara elektronik. Sedangkan menurut Berkatulloh dan Prasetyo (2005) menjelaskan bahwa e-

commerce adalah kegiatan-kegiatan bisnis yang menyangkut konsumen (consumers), manufaktur (manufactures), service providers dan pedagang perantara (intermediaries), dengan menggunakan jaringan-jaringan computer (computer networks) yaitu internet.

6. Shopee

Shopee adalah platform e-commerce terkemuka di Asia Tenggara dan Taiwan. Ini adalah platform yang dirancang untuk kawasan ini, memberikan pelanggan pengalaman belanja online yang mudah, aman, dan cepat melalui pembayaran dan dukungan logistik yang kuat.

Shopee bertujuan untuk terus meningkatkan platformnya dan menjadi tujuan pilihan e-commerce kawasan. Shopee memiliki berbagai pilihan kategori produk. Tujuan dari shopee adalah menyediakan platform untuk menghubungkan pembeli dan penjual dalam satu komunitas. Karena berbelanja di perangkat seluler menjadi kebiasaan baru, Shopee bertujuan untuk terus meningkatkan platformnya untuk menghadirkan pengalaman belanja yang mulus dan menyenangkan bagi semua pengguna dan menjadi platform pilihan e-commerce di era ini (Shopee, 2019).

7. Bukalapak

Bukalapak merupakan salah satu online marketplace terkemuka di Indonesia yang dimiliki dan dijalankan oleh PT. Bukalapak dengan model bisnis Customer-To-Customer (C2C).

Bukalapak menyediakan sarana penjualan dari konsumen ke konsumen di mana pun. Siapapun bisa membuka toko online untuk kemudian melayani calon pembeli dari seluruh Indonesia baik satuan ataupun dalam jumlah banyak. Pengguna perorangan ataupun perusahaan dapat membeli dan menjual produk, baik baru maupun bekas, aksesoris gawai, komputer, sabak (tablet), perlengkapan rumah tangga, busana, elektronik, dan lain-lain.

8. Google Play Store

Google Play Store adalah layanan konten digital milik Google yang melingkupi toko online untuk produk-

produk seperti musik/lagu, buku, aplikasi, permainan, ataupun pemutar media berbasis cloud. Layanan ini dapat diakses baik melalui web, aplikasi android (Play Store) dan Google TV. Google Play Store mulai dikenalkan pada bulan Maret 2012 sebagai pengganti dari Google Play Store dan layanan musik Google.

9. Review

Menurut Kotler dan Armstrong (1996) ada empat metode untuk mengukur kepuasan pelanggan yang salah satunya adalah sistem keluhan dan saran melalui berbagai media seperti kartu komentar dan juga internet seperti online review yang ada dalam situs-situs e-commerce. Menurut Mudambi dan Schuff dalam jurnal yang berjudul "A Study of Customer Review on Amazon.com, "Online customer reviews can be defined as peer-generated product evaluations posted on company or third party websites". yaitu online review dari pelanggan diartikan sebagai hasil evaluasi produk yang diunduh dalam situs perusahaan atau situs pihak ketiga.

METODE PENELITIAN

1. Sumber Penelitian

Data yang digunakan adalah data sekunder. Data diambil dengan menggunakan proses *Scraping* pada kolom komentar Shopee dan Bukalapak disitus Google Play Store. Populasi pada penelitian ini adalah seluruh Review mengenai Shopee dan Bukalapak di Google Play Store. Sampel yang digunakan adalah data diambil sebanyak Satu bulan yaitu Juni 2019.

2. Variabel Penelitian

Variabel yang digunakan sebanyak tiga atribut dan satu label. Atribut yang digunakan antara lain tanggal Review, komentar, dan rating. Label yang digunakan adalah data hasil pelabelan dari analisis sentimen. Jenis data label adalah kategorik yakni positif dan negatif.

3. Langkah-langkah penelitian

Adapun langkah-langkah penelitian ini adalah:

1. Mengambil data dengan teknik *scraping*

2. Melakukan *preprocessing* data yakni Spelling Normalization, Case Folding, Tokenizing, Stemming dan Filtering
3. Analisis Deskriptif
4. Melakukan pelabelan data
5. Membagi data dengan X-Cross Fold Validation
6. Melakukan klasifikasi dengan *stochastic gradient descent*.
7. Tahap Visualisasi

HASIL DAN PEMBAHASAN

1. Scraping Data

Hasil Scraping data review komentar Shopee dan Bukalapak disitus Google Play Store pada bulan Juni 2019 didapatkan data review Shopee sebanyak 2197 record dan Bukalapak sebanyak 4975 record.

2. Preprocessing data

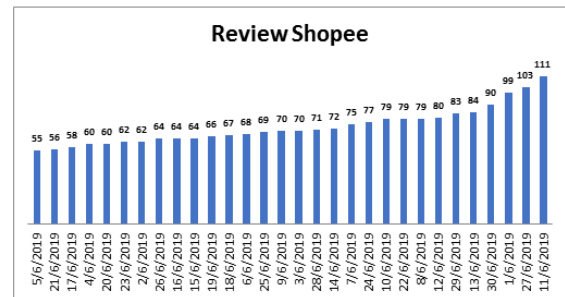
Tabel 1 Hasil Preprocessing Data Review

Hasil Case Folding dan Tokenisasi	
Recomended,proses cepat,aman terpercaya,cuman kadang ada juga yg melakukan pen!puan dengan kedok cuci gudang jd mohon hal seperti cepat di atasi	terekomendasi proses cepat aman terpercaya cuma kadang ada juga yang melakukan penipuan dengan kedok cuci gudang jadi mohon hal seperti cepat diatasi
Hasil Stemming dan Filtering	
terekomendasi proses cepat aman terpercaya cuma kadang ada juga yang melakukan penipuan dengan kedok cuci gudang jadi mohon hal seperti cepat diatasi	terekomendasi proses cepat aman terpercaya kadang melakukan penipuan kedok cuci gudang jadi mohon cepat diatasi
Hasil Data Bersih	
terekomendasi proses cepat aman terpercaya kadang melakukan penipuan kedok cuci gudang mohon cepat diatasi	rekomen proses cepat aman percaya kadang lakukan tipu kedok cuci gudang mohon cepat atasi

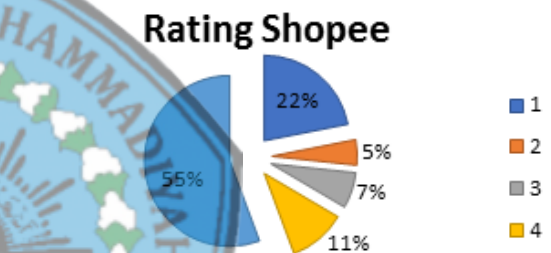
3. Analisis Deskriptif

Analisis deskriptif dilakukan untuk mengetahui gambaran umum mengenai e-commerce di laman google play. Informasi

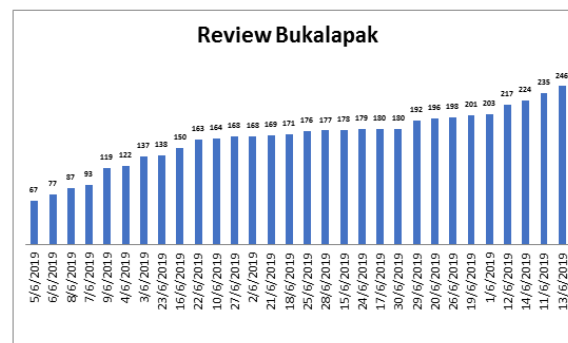
yang dapat diambil adalah bagaimana Jumlah review yang diberikan oleh konsumen perharinya serta mengetahui tingkat kepuasan konsumen berdasarkan pemberian bintang.



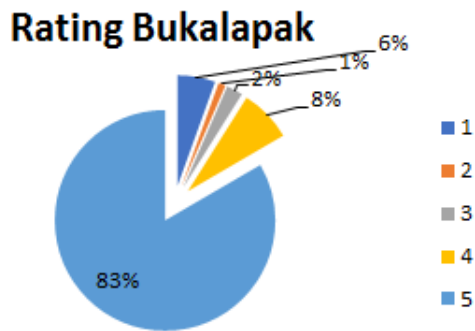
Gambar 2 Grafik Review Shopee Juni 2019 Berdasarkan hasil grafik dapat dilihat bahwa komentar terbanyak terjadi pada tanggal 11, 27 dan 1 Juni 2019 serta komentar paling sedikit terjadi pada tanggal 5 Juni 2019.



Gambar 2 Rating Review Shopee Juni 2019 Berdasarkan grafik dapat digambarkan bahwa sebesar 55% konsumen yang memberikan komentar memberikan rating 5 dan selanjutnya sebanyak 22% konsumen memberikan rating 1 kepada Shopee .



Gambar 3 Grafik Review Bukalapak Juni 2019 Berdasarkan hasil grafik dapat dilihat bahwa komentar terbanyak terjadi pada tanggal 13, 11 dan 14 Juni 2019 serta komentar paling sedikit terjadi pada tanggal 5 Juni 2019.



Gambar 4 Rating Review Bukalapak Juni 2019 Berdasarkan grafik 4.5 dapat digambarkan bahwa sebesar 83% konsumen yang memberikan komentar memberikan rating ke 5 dan selanjutnya hanya 6% konsumen memberikan rating 1 kepada Bukalapak.

4. Pelabelan Data

Pelabelan pada kelas sentimen hanya menggunakan kategori negatif dan kategori positif adapun hasil dari pelabelan untuk bukhalapak dan shopee dapat di tunjukan sebagai berikut:

Tabel 3 Jumlah pelabelan kelas sentimen

Sentimen	Review Shopee	Review Bukalapak
Positif	1731	4546
Negatif	466	429

5. Klasifikasi

Tabel 3 Hasil Klasifikasi dengan Menggunakan *Stochastic Gradient Descent*

Hasil	Shopee	Bukalapak
Accuracy	89.89 %	95.09 %
Sensitivitas	92,4%	96,9%
Spesifisitas	79,3%	73,9%

Hasil klasifikasi yang dilakukan dengan *Stochastic Gradient Descent* untuk bukhalapak juga lebih tinggi artinya kesalahan prediksi pada bukhalapak lebih tinggi daripada shopee yakni 95.09% sedangkan shopee 89.89%. Berdasarkan nilai sensitivitas ketepatan memprediksi kelas aktual positif lebih tinggi Bukalapak yakni sebesar 96,9%, namun berdasarkan nilai spesifisitas shopee memiliki nilai lebih tinggi artinya kelas false positive lebih sedikit

dibandingkan bukhalapak yakni sebesar 79,3%.

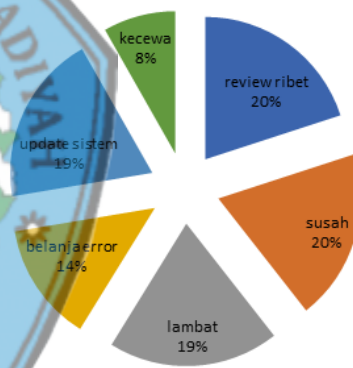
6. Visualisasi



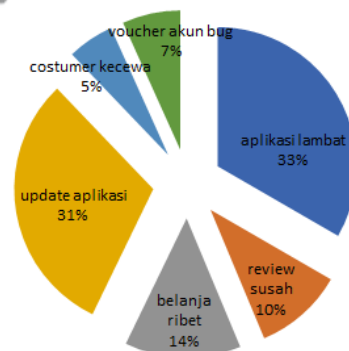
Visualisasi Shopee

Visualisasi Bukalapak

Berdasarkan gambar diatas visualisasi kategori review negatif shopee dan bukhalapak kata-kata yang memiliki ukuran lebih besar adalah kata-kata yang paling banyak muncul di review. Pada e-commerce shopee fokus konsumen terhadap aplikasi dan bukhalapak pada barang. “Lambat, Ribet, Error, Susah”. Sehingga peningkatan kualitas yang dapat di perbaiki adalah perbaikan aplikasi pada shopee dan proses pencarian barang untuk bukhalapak.



Gambar 5 Review Negatif Bukalapak



Gambar 6 Review Negatif Shopee

KESIMPULAN

1. Gambaran umum review bukhalapak dan shopee pada Bulan Juni 2019 yakni komentar yang diberikan perharinya

- fluktuatif dan mendapatkan review paling banyak pada tanggal 11 Juni 2019 yang bertepatan dengan hari promo. Serta mayoritas pengguna sudah memberikan rating positif kepada dua e-commerce tersebut.
- Hasil klasifikasi sentimen menggunakan metode Stochastic Gradient Descent dengan pembagian data menggunakan 10 Fold Cross Validation diperoleh tingkat akurasi sebesar 95.09 % untuk bukalapak dan 89.89 % untuk shopee. Selain itu sensitivitas sebesar 96,9% untuk bukalapak dan 92,4% untuk shopee serta spesifisitas sebesar 79,3% untuk bukalapak 79,3% untuk shopee. Hal tersebut menunjukkan kemampuan prediksi dengan menggunakan Stochastic Gradient Descent baik sehingga dapat dijadikan metode rujukan untuk mengolah data text.
 - Berdasarkan hasil klasifikasi dan asosiasi teks yang dilakukan, secara umum dapat diketahui bahwa pengguna aplikasi Bukalapak dan Shopee mayoritas membicarakan mengenai aplikasi dan transaksi karena selalu muncul pada kelas sentimen negative yang menunjukkan hasil ulasan negatif yang sering dikeluhkan diantaranya aplikasi lambat, perlu di update, ribet, mudah error, dan aplikasi berat.

DAFTAR PUSTAKA

- Alamsyah, A., & Saviera, F. 2017. A Comparison of Indonesia's E-Commerce Sentiment Analysis for Marketing Intelligence Effort (case study of Bukalapak, Tokopedia and Elevenia). *The 8 Th International Conference on Sustainable Collaboration in Business , Technology, Information and Innovation*. Retrieved from <http://scbtii.telkomuniversity.ac.id/>
- APJII. (2018). Penetrasi & Profil Perilaku Pengguna Internet Indonesia. *Apjii*. Retrieved from www.apjii.or.id
- Aprilia. 2017. *Analisis Faktor-Faktor yang Mempengaruhi Perilaku Pembelian Online Pada Mahasiswa Yogyakarta*. Skripsi. Program Studi Manajemen Fakultas Ekonomi UNY Yogyakarta.
- Berkatullah, A.,H., & Prasetyo, T. 2005. *Bisnis E-commerce Studi Sistem Keamanan dan Hukum di Indonesia*. Yogyakarta:Pustaka Pelajar.
- Bukalapak. 2016. *Tentang Bukalapak*. <https://www.bukalapak.com/about>. Akses 10 Juli 2019.
- Dong, X. M., & Chen, D. R. 2009. Learning rates of gradient descent algorithm for classification. *Journal of Computational and Applied Mathematics*, 224(1), 182–192. <https://doi.org/10.1016/j.cam.2008.04.022>
- Fanani, F. 2017. *Klasifikasi Review Software Pada Google Play Menggunakan Pendekatan Analisis Sentimen*. Skripsi. Program Studi Teknologi Informasi Fakultas Teknik UGM Yogyakarta.
- Feldman, R., & Sanger, J. 2007. *The Text Mining Handbook Advanced Approaches In Analyzing Unstructured Data*. New York : Cambridge University Press.
- Firdaus, M. 2015. *Analisis Faktor-Faktor Yang Mempengaruhi Pengguna Belanja Secara Online (E-Commerce)*. Skripsi. Program Studi Akuntansi Fakultas Ekonomi UII Yogyakarta.
- Gusriani, S., Wardhani, K.D.K., & Zul, I.M. 2016. *Analisis Sentimen Berdasarkan Komentar Publik Terhadap Toko Online Pada Media Sosial Facebook (Studi Kasus:Zalora dan BerryBenka)*. *Jurnal Aksara Komputer Terapan* Vol 5, No
- Han, J., & Kamber, M. 2006. *Data Mining :*

- Concepts and Techniques Second Edition.* San Francisco : Morgan Kauffman.
- Ha, S.H., Bae, S., & Son, L.K. 2015. *Impact of Online Consumer Reviews on Product Sales: Quantitative Analysis of the Source Effect.* Journal of Applied Mathematics & Information Sciences. Vol 9. No 2: 373-387.
- Indrajit, R., E. 2011. *Manajemen Sistem Informasi dan Teknologi Informasi.* Jakarta: Elex Media Komputindo.
- Iprice. 2019. Ini e-Commerce Indonesia Paling Diminati pada Triwulan IV 2018, 2.Retrieved from <https://databoks.katadata.co.id/datapublish/2019/01/31/ini-e-commerce-indonesia-paling-diminati-pada-triwulan-iv-2018>
- Jamil, H.N. 2017. *Analisis Sentimen Pada Online Review Menggunakan Kombinasi Metode Lexicon Based dan Naïve Bayes Classifier.* Skripsi. Program Studi Statistika FMIPA UII Yogyakarta.
- Kaestner, C. 2013. *Support Vector Machines and Kernel Functions for Text Processing.* RITA Volume 20 Number 3 201.
- Kang, H., Yoo, S. J., & Han, D. 2012. Expert Systems with Applications Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems With Applications*, 39(5), 6000–6010. <https://doi.org/10.1016/j.eswa.2011.11.107>
- Karami, G., Kordlar, A.E., Amini, Y., & Hajipour, S. 2014. Earnings Management and Underpricing of Initial Public Offerings (IPO), Evidence from Iran. *International Business Research.*
- Kotler, P., & Armstrong, G. 1996. *Dasar-Dasar Pemasaran, Edisi V, jilid 2.* Jakarta: Intermedia
- Lee, L., & Pang, B. 2008. *Opinion Mining and Sentiment Analysis.* Foundation and Trends in Information Retrieval, 2(1-2): 1-135.
- Liu, B. 2012. *Sentiment Analysis and Subjectivity.* Synthesis Lectures on Human Language Technologies. USA: Morgan & Claypool Publishers.
- Liu, L., Chunfang & Zhouyang. 2013. *Analysis of Customer Satisfaction from Chinese Reviews using Opinion Mining.* 6th IEEE International Conference on Software Engineering and Service Science (ICSESS).
- Liu, Y., Loh, H. T., & Sun, A. 2009. Imbalanced text classification: A term weighting approach. *Expert Systems with Applications*, 36(1), 690–701. <https://doi.org/10.1016/j.eswa.2007.10.042>
- Moh, M., Gajjala, A., Gangireddy, S. C. R., & Moh, T. S. 2016. On multi-tier sentiment analysis using supervised machine learning. *Proceedings - 2015 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015*, 1, 341–344. <https://doi.org/10.1109/WI-IAT.2015.154>
- Mohaiminul, M., & Sultana, N. 2018. Comparative Study on Machine Learning Algorithms for Sentiment Classification. *International Journal of Computer Applications*, 182(21), 1–7. <https://doi.org/10.5120/ijca2018917961>
- Onno, P. 2000. *Mengenal E-commerce.* Jakarta: PT Elek Media Komputindo.

- Pang, B. 2002. "Thumbs up? Sentiment Classification Using Machine Learning Techniques." Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). 79 - 86.
- Potdar, A., Patil, P., Bagla, R., Pandey, R., & Jadhav, N. 2016. SAMIKSHA - Sentiment Based Product Review Analysis System. *Physics Procedia*, 78 (December 2015), 513–520. <https://doi.org/10.1016/j.procs.2016.02.096>
- Putranti, N. D., & Winarko, E. 2014. Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 8(1), 91–100. <https://doi.org/10.22146/ijccs.3499>
- Putri, D.U.K. 2016. *Implementasi Inferensi Fuzzy Mamdani Untuk Keperluan Sistem Rekomendasi Berita Berbasis Konten*. Skripsi. Program Studi Ilmu Komputer FMIPA UGM Yogyakarta.
- Ruder, S. 2016. An overview of gradient descent optimization algorithms, 1–14. Retrieved from <http://arxiv.org/abs/1609.04747>
- Sharma, A. 2018. Guided Stochastic Gradient Descent Algorithm for inconsistent datasets. *Applied Soft Computing Journal*, 73, 1068–1080. <https://doi.org/10.1016/j.asoc.2018.09.038>
- Sung, J., & Sung, E. 2014. Exploring the Usefulness of a Decision Tree in Predicting People 's Locations. *Procedia - Social and Behavioral Sciences*, 140, 447–451. <https://doi.org/10.1016/j.sbspro.2014.04.451>
- Saraswati, N.S. 2011. *Text Mining dengan Metode Naive Bayes Classifier dan Support Vector Machines untuk Sentiment Analysis*. Skripsi. Program Studi Teknologi Informasi Fakultas Teknik UGM Yogyakarta.
- Shopee. 2019. *Tentang Shopee*. <https://careers.shopee.sg/about/>. Akses 10 Juli 2019.
- Susanti, A.R. 2016. *Analisis Klasifikasi Sentimen Twitter Terhadap Kinerja Layanan Provider Telekomunikasi Menggunakan Varian Naive Bayes*. Tesis. Institut Pertanian Bogor.
- Tjiptono, F. 2006. *Manajemen Pelayanan Jasa*. Yogyakarta: Andi.
- Turban, E., & King, D. 2002. *Electronic Commerce 2002 – A Managerial Perspective (Second edition)*. New York: Prentice Hall.
- Zafikri, A. 2008. *Implementasi Metode Term Frequency Inverse Document Frequency (TF-IDF) Pada Sistem Temu Kembali Informasi*. Skripsi. Program Studi S-1 Ilmu Komputer FMIPA USU.
- Zaki, M. J., & Meira, W. J. 2014. *Data Mining and Analysis: Fundamental Concepts and Algorithms*, 562. Retrieved from <https://books.google.com/books?hl=en&lr=&id=Gh9GAAwAAQBAJ&pgis=1>