

**ANALISA KONTEN MEDIA SOSIAL E-COMMERCE PADA INSTAGRAM
MENGUNAKAN METODE SENTIMEN ANALYSIS DAN LDA-BASED TOPIC
MODELING (STUDI KASUS: SHOPEE INDONESIA)**

**ANALYSIS OF CONTENT SOCIAL MEDIA E-COMMERCE IN INSTAGRAM USING
SENTIMENT ANALYSIS AND LDA BASED TOPIC MODELING (STUDY CASE :
SHOPEE INDONESIA)**

Irawan Noor Kabiru¹, Puspita Kencana Sari, S.kom, MTI²

^{1,2}Prodi S1 Manajemen Bisnis Telekomunikasi dan Informatika,
Fakultas Ekonomi dan Bisnis, Universitas Telkom
arbikabiru@student.telkomuniversity.ac.id¹, puspitakencana@telkomuniversity.ac.id²

Abstrak

Perkembangan teknologi informasi pada era digital ini memberikan sebuah dampak yang sangat besar bagi dunia. Perkembangan teknologi informasi ini membuat dampak pada sistem jual beli, dimana berkembang sistem model bisnis jual beli yang disebut *e-commerce*. Shopee merupakan salah satu perusahaan yang sukses yang menggunakan model bisnis *e-commerce* ini.

Dalam melakukan bisnis pada era digital, memahami konsumen merupakan salah satu hal yang perlu diperhatikan. Memahami pelanggan merupakan cara agar perusahaan dapat terus melihat *feedback* yang diberikan konsumen terhadap produk maupun jasa yang diberikan kepada perusahaan. Media sosial merupakan platform yang digunakan konsumen dalam bercerita maupun menulis yang berhubungan dengan sebuah produk atau jasa.

Metode yang digunakan pada sentiment analysis adalah menggunakan *Naïve Bayes Model*, serta pembuatan *Topic Modeling* dengan menggunakan *LDA-Based Topic Modeling*. Sentiment yang terbentuk dalam konten komentar Instagram Shopee pada *Flashsale 11:11* adalah negatif.

Saran untuk Shopee dapat meningkatkan strategi mereka dalam melakukan acara seperti ini seperti mengundang artis seperti yang dilakukan sebelumnya dan meningkatkan layanan agar konsumen dapat berhubungan dengan produk dalam shopee maupun sesama pengguna shopee, hal ini terkait dengan *tribal marketing*.

Shopee harus lebih responsif terhadap keadaan yang terjadi di lapangan, pada kendala yang terjadi pada aplikasi yang mengganggu pada kepuasan pelanggan.

Kata Kunci : Sentiment Analysis, Topic Modeling, Naive Bayes, Tribal Marketing.

Abstract

The development of information technology in this digital era has had a huge impact on the world. The development of information technology has made an impact on the buying and selling system, which has developed a system of buying and selling business models called e-commerce. Shopee is one of the successful companies that uses this e-commerce business model.

Understanding consumers is one of the things that needs attention. Understanding customers is a way for companies to continue to see feedback given by consumers to products and services provided to companies. Social media is a platform used by consumers in telling stories and writing related to a product or service.

The method used in sentiment analysis is using the Naïve Bayes Model, as well as making Topic Modeling using LDA-Based Topic Modeling. The sentiment that formed in Shopee's Instagram was negative.

Suggestions for Shopee, can improve their strategy in conducting events such as inviting artists like they did before and improving services so that consumers can connect with products in shopee or fellow shopee users, this is related to tribal marketing.

Shopee must be more responsive to the conditions that occur in the field, on the obstacles that occur in applications that interfere with customer satisfaction.

Keywords : Sentiment Analysis, Topic Modeling, Naive Bayes, Tribal Marketing

1. Pendahuluan

Perkembangan teknologi memfasilitasi kegiatan masyarakat dalam bisnis jual beli dan memfasilitasi kegiatan lainnya. Teknologi yang berkembang dalam bisnis jual beli ini salah satunya adalah e-commerce. E-commerce adalah sarana transaksi komersial antara organisasi dan individu yang dilakukan secara digital [1]. Orang Indonesia memanfaatkan perkembangan teknologi e-commerce sebagai sarana jual beli. Nilai transaksi e-commerce di Indonesia meningkat, di mana pada tahun 2014 transaksi e-commerce di Indonesia bernilai Rp25 triliun, 2016 sebesar Rp69,8 triliun dan diproyeksikan meningkat pada 2018 senilai Rp114 triliun [2].

Pertumbuhan dalam penggunaan media sosial juga merupakan salah satu efek dari peningkatan jumlah pembelian barang dengan cara digital, tidak datang untuk melihat toko fisik secara langsung. Dengan pertumbuhan media sosial yang tinggi, media sosial juga sering digunakan dalam melakukan kegiatan bisnis, Instagram, di mana media sosial memudahkan bisnis untuk membuat katalog dan menampilkan produk yang mereka jual, sehingga konsumen dapat melihat dan membeli produk tersebut.

Shopee adalah e-commerce paling banyak yang memiliki pengikut terbanyak di media sosial Instagram di Indonesia, Shopee memiliki 1.101.070 pengikut, selain menjadi e-commerce yang memiliki jumlah pengikut paling banyak, Shopee adalah rangking teratas aplikasi belanja di iOS dan Android [3].

Berdasarkan hasil survei APJII pada tahun 2018, Indonesia mengalami peningkatan pengguna internet menjadi 143,26 juta orang, setara dengan 54,7% dari total populasi Indonesia. Ini membuktikan bahwa perkembangan teknologi di Indonesia cukup tinggi dan memiliki potensi untuk dikembangkan dalam berbagai aspek [4]

Dalam bisnis, ada strategi yang dilakukan oleh perusahaan sehingga produk yang ditawarkan menarik dan banyak yang membeli produk ini, salah satu strategi untuk meningkatkan penjualan adalah dengan memiliki strategi pemasaran. Strategi yang sering digunakan oleh e-commerce adalah membuat penjualan yang berbeda dari yang lain, yaitu penjualan dilakukan dari 25 Oktober 2018 hingga 11 November 2018. Penjualan tersebut, yang disebut 11:11 oleh shopee, adalah suatu usaha yang dapat meningkatkan penjualan suatu produk karena melihat penjualan dilakukan dalam jumlah waktu tertentu

Data yang ada di media sosial ada di media sosial disebut User Generated Content (UGC). Di UGC ada bagian dalam bentuk opini publik atau opini publik. Data di media sosial yang diposting oleh pengguna memiliki berbagai karakteristik, dan tersedia untuk pengguna lain, dalam data tersebut dapat berisi sejumlah upaya dan pendapat kreatif, dan dibuat oleh komunitas di luar kegiatan profesional mereka [6].

Salah satu media sosial yang memiliki UGC adalah Instagram, Instagram adalah foto dan video berbasis media sosial yang dapat dibagikan oleh pengguna ke pengguna lain. Di UGC ada banyak sentimen yang ditulis dan diungkapkan oleh publik, ada ekspresi dalam bentuk negatif dan dalam bentuk positif. Data dapat diolah dan digunakan oleh perusahaan, sehingga perusahaan dapat memahami sentimen apa yang mendominasi pemikiran konsumen tentang perusahaan atau produk mereka.

Analisis sentimen adalah salah satu bagian ilmiah dalam mengklasifikasikan teks dan mengklasifikasikan teks berdasarkan pendapat dalam teks. Analisis sentimen adalah suatu proses dalam mendeteksi polaritas suatu teks secara kontekstual, ini menentukan suatu teks yang memiliki makna positif, negatif atau netral. Analisis sentimen juga disebut penambangan opini, karena hasil analisis berasal dari pendapat seseorang atau dari sikap seseorang yang bersangkutan [7].

Berdasarkan fenomena yang diuraikan, pertanyaan penelitian yang akan diajukan adalah:

1. Apa sentimen pengguna Instagram tentang shopee, jika dilihat dari konten percakapan media sosial Instagram?
2. Topik apa yang dibicarakan pengguna tentang kolom komentar Shopee, tentang konten percakapan di Instagram secara keseluruhan?
3. Apa topik yang dibahas oleh pengguna pada setiap sentimen?

2. Dasar Teoritis

2.1. Big Data

Data besar adalah kumpulan data yang memiliki jumlah besar dan membutuhkan alat untuk mengambil data dan menganalisisnya. Informasi dan data adalah hal yang sangat penting dan menjadi salah satu poin keberhasilan suatu organisasi. Big data memiliki 3 elemen kunci utama yang membedakan data besar dari data biasa, juga dikenal sebagai 3 V Big Data, volume, kecepatan dan variasi [8]

2.2. Data Mining

Data Mining adalah pengambilan data mentah, biasanya data mining memiliki sejumlah besar data, untuk mendapatkan wawasan atau informasi baru dari data. Wawasan atau informasi memiliki pola tersembunyi karena kompleksitas dan ukuran data [9]

2.3. Text Mining

Text Mining adalah analisis data di mana ada bahasa alami dengan menggunakan teknik dan alat untuk merancang dan menemukan dan mengekstrak pengetahuan dari data yang tidak terstruktur [10]. Penambangan teks berfungsi dengan mengubah kata atau kalimat dalam data yang tidak terstruktur menjadi data dengan topik yang lebih terstruktur dan dapat lebih mudah dianalisis.

2.4. Sentiment Analysis)

Analisis sentimen berbasis teks dan mengklasifikasikan teks dalam opini dalam teks. Analisis sentimen adalah proses mendeteksi polaritas teks secara kontekstual, ini menentukan teks yang memiliki makna positif, negatif atau netral. Analisis sentimen juga disebut penambangan opini, karena hasil analisis berasal dari pendapat seseorang dari sikap orang yang bersangkutan [7].

2.5. Naïve Bayes

Klasifikasi naïve Bayes adalah klasifikasi sederhana dengan model probabilitas yang dijalankan berdampingan dengan asumsi yang kuat, asumsi ini tidak secara signifikan mempengaruhi keakuratan data dalam klasifikasi teks, tetapi ini membuat algoritma klasifikasi berjalan cepat dan dapat diterapkan untuk masalah [11].

2.6. Topic Modeling

Model topik fokus pada menemukan struktur laten dari topik tertentu dalam kumpulan data dokumen, yang berasal dari kata-kata yang muncul dalam dokumen [12]. Topic Modeling adalah teori dan metode untuk mengekstraksi dan mewakili konteks dalam dokumen yang digunakan dengan memanfaatkan statistik komputasi untuk sejumlah besar corpus dari teks.

2.7. Latend Dirichlet Allocation (LDA)

Latend-Dirichlet Allocation adalah model probabilistik generatif untuk kumpulan teks data. LDA adalah hierarki tiga tingkat model Naïve Bayes yang dapat disimpulkan sebagai kluster kata, yang disebut topik dari dokumen [13]

3. Metodologi Penelitian

Data yang digunakan dalam proyek ini adalah data komentar dari Shopee Instagram selama periode peristiwa 11:11 flashsale yang dimulai pada 25 Oktober 2018 hingga 11 November 2018. Data diambil menggunakan aplikasi webscrapping parsehub, dan kemudian data yang diambil oleh aplikasi dilabeli oleh sentimen positif atau negatif. Data yang telah dikumpulkan berjumlah 29.584 komentar dan dibagi menjadi sentimen positif dan negatif, data berubah menjadi 3.884.

3.1 Teknik Analisis Data

1. Pengumpulan Data: Data dikumpulkan dengan merangkak menggunakan aplikasi Parsehub. Pengumpulan data dilakukan oleh semua komentar dari pengguna di Instagram Shopee dari periode waktu.
2. Data Preprocessing: adalah proses pembersihan data dimana dalam proses ini data yang diperoleh dalam bentuk teks dilakukan dengan membersihkan dengan mengubah huruf kapital, tokenize, stopword removal dan stemming.
3. Analisis sentimen: setelah preprocessing data, analisis sentimen dilakukan dengan metode Naïve Bayes untuk menentukan sentimen dalam teks
4. Pemodelan topik: setelah mengetahui sentimen yang terkandung dalam teks, menemukan topik yang sering dibahas dalam teks.

4. Hasil dan Pembahasan

4.1. Karakteristik Data

Berbagi data dilakukan dengan membagi data menjadi data pelatihan dan data pengujian, dengan proporsi 30% data pelatihan, dan 70% data pengujian, ini dilakukan agar pembelajaran mesin dapat mempelajari dan mengevaluasi pengujian data.

Tabel 4. 1 Klasifikasi berdasarkan Sentimen

| Comment | Sentiment |
|--|-----------|
| Sama dari kemarin sampai pagi ini memang muter terus biar tidak rugi kali jadi di error in | Negative |
| Hai shopee semakin hari semakin sukses ya | Positive |

dan Dalam proses pengolahan data pelatihan dan pengujian data yang telah dilakukan bahwa mesin dapat memproses dengan baik karena memiliki tingkat akurasi 88,73%, tingkat klasifikasi kesalahan 11,27 nilai kappa 0,764

Tabel 4. 2 Performansi Naive Bayes

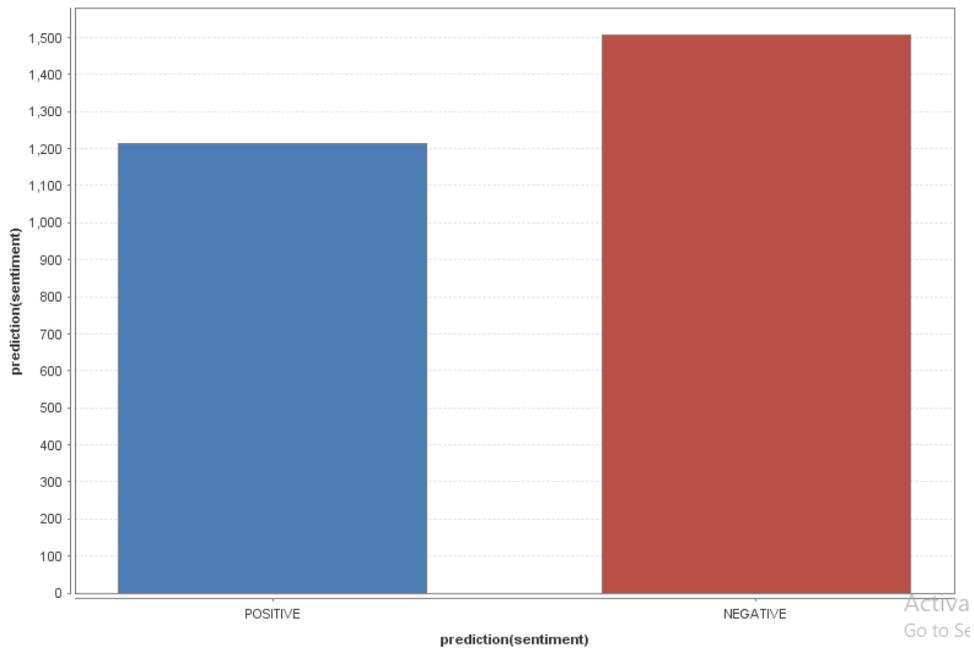
| Accuracy | Classification Error | Kappa |
|----------|----------------------|-------|
| 88,73 | 11,27 | 0,764 |

4.2. Hasil Pengolahan Data

4.2.1 Hasil Pengolahan data Sentimen

Hasil pengolahan data yang telah dilakukan menunjukkan tingkat akurasi yang baik, dan data yang telah diolah memiliki sentimen negative, data berjumlah 3.884

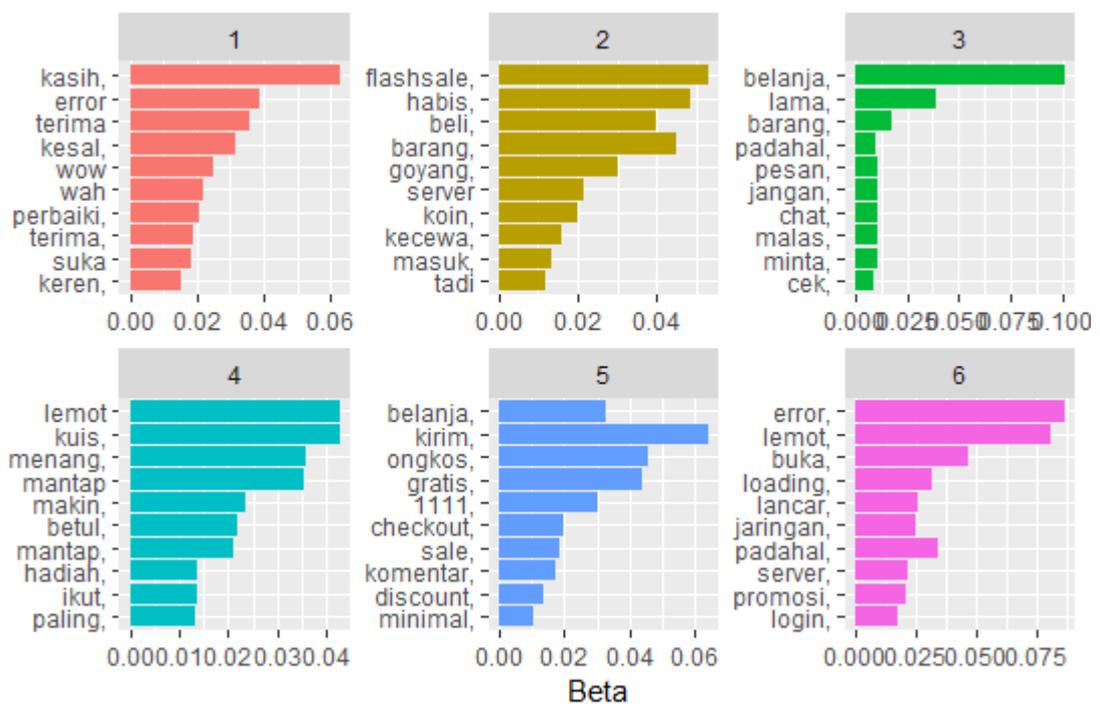
Gambar 4. 1 Hasil Sentimen Analysis



Berdasarkan hasil grafik di atas, dapat dilihat bahwa hasil keseluruhan sentimen di Shopee Instagram pada periode acara flashsale 11:11 adalah negatif, dengan jumlah sentimen 1505 sedangkan untuk sentimen positif adalah 1213.

4.2.2 Hasil LDA-Based Topic Modeling

Dengan menggunakan Pemodelan Topik Berbasis LDA data di proses Menggunakan R-Studio, data yang dimiliki akan dilihat sebagai keseluruhan topik yang dibahas oleh pengguna di kolom komentar Shopee Instagram, sebagai berikut:

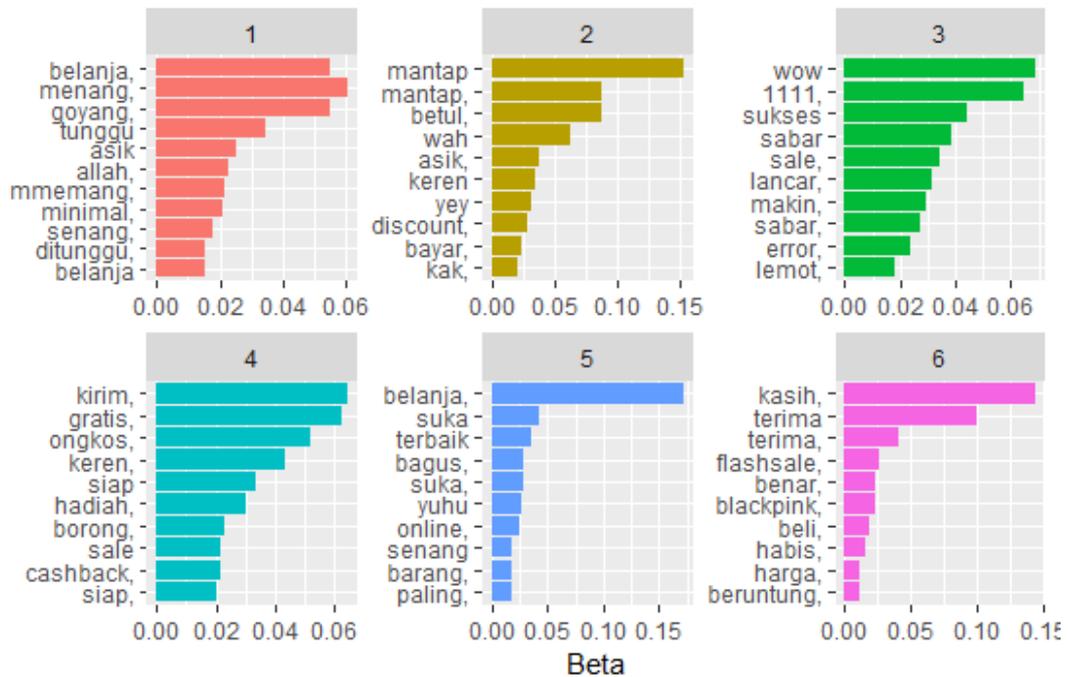


Gambar 4. 2 Hasil LDA- Baset Topic Modeling

n 6 topik yang dibicarakan secara keseluruhan pada Instagram Shopee periode 25 Oktober 2018 – 11 November 2018, topik yang terdapat pada Instagram Shopee sebagai berikut :

1. Aplikasi keren, namun error dan konsumen ingin perbaikan pada server
2. Kekecewaan pada flashsale ketika beli barang dan server
3. Belanja barang yang lama membuat konsumen malas
4. Ketika kuis aplikasi lemot serta keinginan konsumen untuk memenangkan kuis shopee agar mendapat hadiah
5. Konsumen yang berkomentar mengenai program cashback dan gratis ongkos kirim yang dinilai menipu
6. Aplikasi yang lemot selalu loading dan error padahal untuk membuka aplikasi lainnya lancar

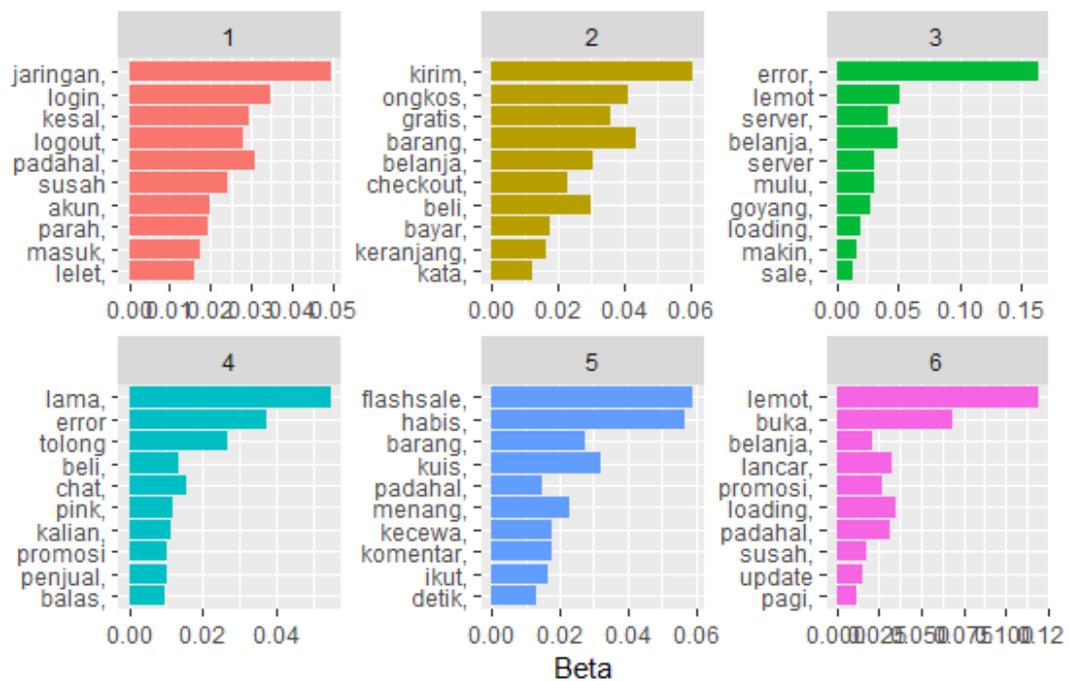
4.2.3 Hasil LDA-Based Topic Modeling sentimen Positif



Gambar diatas menunjukkan 5 topik yang dibicarakan pada sentimen positif di Instagram Shopee periode 25 Oktober 2018 – 11 November 2018, topik yang terdapat pada Instagram Shopee sebagai berikut :

1. Konsumen merasa waktu goyang shopee ditunggu dan belanja di shopee adalah hal yang asik
2. Discount yang diberikan oleh shopee dinilai mantap oleh konsumen
3. Walaupun flashsale 11:11 lemot dan lambat namun konsumen tetap mendoakan shopee agar sukses
4. Konsumen menganggap cashback dan gratis ongkos kirim shopee adalah event yang keren dan akan membeli banyak barang dengan hal itu
5. Shopee merupakan marketplace terbaik dan ter bagus
6. Terima kasih konsumen terhadap shopee

4.2.4 Hasil LDA-Based Topic Modeling Negatif



Gambar diatas menunjukkan 6 topik yang dibicarakan pada sentimen negatif di Instagram Shopee periode 25 Oktober 2018 – 11 November 2018, topik yang terdapat pada Instagram Shopee sebagai berikut :

1. Konsumen merasa kesal karena jaringan tidak bisa dimasuki karena lemot sedangkan sudah dicoba untuk logout dan login
2. Konsumen merasa promosi, cashback hanya sebuah pemberian harapan palsu oleh shopee dan mereka ingin cepat di respon oleh shopee
3. Server yang selalu error dan loading membuat susah belanja
4. Keluhan karena aplikasi error namun admin shopee di berikan pesan tidak pernah membalas pesan tersebut dan lama untuk respon
5. Kecewa karena flashsale dan kuis yang diselenggarakan karena selalu ikut kuis namun tidak pernah menang
6. Promosi membuat aplikasi loading terus menerus dan susah belanja, sedangkan aplikasi sudah di update

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil dari sentimen analysis dan topic modeling yang dilakukan pada data komentar Instagram Shopee pada kurun waktu event 11:11, berdasarkan tahapan penelitian yang telah dilakukan dapat disimpulkan bahwa :

1. Sentimen yang terbentuk dalam konten komentar Instagram Shopee pada waktu event 11:11 adalah negatif, dengan menggunakan metode Naïve Bayes, didapat bahwa terdapat jumlah sentimen negatif
2. Topik yang dibahas dalam konten keseluruhan pada Instagram Shopee pada waktu event 11:11 adalah ungkapan kekecewaan terhadap aplikasi maupun promosi yang dilakukan shopee karena server error dan lemot.
3. Topic yang dibahas pada sentimen positif adalah Secara keseluruhan sentimen positif yang terdapat pada instagram shopee pada event 11:11 ini adalah konsumen merasa event ini bagus karena ada goyang shopee, cashback, gratis ongkos kirim dan senang belanja di aplikasi tersebut walaupun lemot dan suka error.
4. Topic yang dibahas pada sentimen negatif adalah Secara keseluruhan sentimen negatif yang terdapat pada instagram shopee pada event 11:11 ini adalah konsumen merasa kecewa karena aplikasi serta promosi yang dilakukan oleh shopee dinilai menipu dan memberikan harapan palsu serta aplikasi yang selalu error dan lemot ketika ada flashsale 11:11 tersebut.

5.2. Saran

Berdasarkan data yang telah diuji pada penelitian ini dan berdasarkan ilmu marketing dalam manajemen, maka penulis menyarankan untuk tetap melakukan acara seperti flashsale 11:11 ini dan mendengar kembali konsumen dan tetap menasar berbagai grup sosial yang terdapat pada konsumen. Shopee dapat meningkatkan strategi mereka dalam melakukan acara seperti ini seperti mengundang artis seperti yang dilakukan sebelumnya dan meningkatkan layanan agar konsumen dapat berhubungan dengan produk dalam shopee maupun sesama pengguna shopee, hal ini terkait dengan tribal marketing. Dapat disimpulkan bahwa Shopee harus lebih responsive terhadap keadaan yang terjadi di lapangan, pada kendala yang terjadi pada aplikasi yang mengganggu pada kepuasan pelanggan dalam menggunakan aplikasi maupun mengikuti acara tersebut.

DAFTAR PUSTAKA

- [1] [1] Laudon, K. C., & Traver (2014). *E-Commerce 2014*, 10th Edition. Pearson
- [2] (2018), Transaksi e-commerce Indonesia Akan Capai Rp 144 Triliun. Didapat dari : <https://www.liputan6.com/teknoread/3057134/2018-transaksi-e-commerce-indonesia-akan-capai-rp-144-triliun>
- [3] The Map of E-Commerce in Indonesia di dapat dari :
<https://iprice.co.id/insights/mapofecommerce/en/MuhammadNayazri>.
- [4] Buletin APJII edisi 22 Maret 2018 di dapat dari : <https://apjii.or.id/content/read/104/348/BULETIN-APJII-EDISI-22---Maret-2018>
- [5] Insights into the growth of Indonesian E-Commerce with payment methods, target audience, marketing, social media, economy and logistics all profiled. Didapat dari : <https://www.eshopworld.com/blog/indonesia-ecommerce-insights-2018/>
- [6] Marie-Francine Moens, Juanzi Li, and Tat-Seng Chua (2014). Mining User Generated Content. *Computer Science, Engineering & Technology*
- [7] Devika M. D, Sumitha C, and Amal Ganesh (2016). *Sentiment analysis : A Comparative Study on Different Approach*. Elsevier
- [8] Milan Kubina, Michael Varmus, and Irena Kubinova (2015). *Use of Big Data for Competitive Advantage of Company*. Elsevier
- [9] Reza Arghandeh, Daniel Bienstock, Pavel Etingov, Sean Murphy, Femi Omitaomu, Mathew Rhodes, Tom Rizy, Anurag K Srivastava, Kai Sun, and Xiangyang Zhou, (2019). *Data Mining Techniques and Tools for Synchrophasor Data*. NASPI EATT
- [10] Felderman and Sanger (20016) dalam Mahdi et al., (2016). Text mining stackoverflow an insight into challenges and subject-related difficulties face by computer science learners. *Journal of Enterprise Information Management*
- [11] Neha S. Joshi and Suhasini A. Itkat (2014). A Survey on Feature Level Sentiment Analysis. *International Journal of Computer Science and Information Technologies*
- [12] Derek O'Callaghan, Derek Greene, Joe Carthy, and Padaraig Cunningham (2015). Analysis of the Coherence of Descriptors in Topic Modeling. *Expert System with Application*
- [13] P. Anupriya and S. Karpagavalli (2015). LDA Based Topic Modeling of Journal Abstracts. *International Conference on Advanced Computing and Communication Systems*.