

Analisis Ulasan Pembelian Produk Beras di Marketplace Shopee dengan Menggunakan Topic Modelling

Dedy Sugiarto, Reyhan Dwi Putra, Wahyu Hidayat, Ema Utami, Ainul Yaqin

Program Studi PJJ S2 Teknik Informatika

Universitas Amikom Jogyakarta

Jogyakarta, Indonesia

dedy_sugiarto@students.amikom.ac.id, reyhan.p@students.amikom.ac.id,

20.55.1384_wahyuhidayat@students.amikom.ac.id, ema.u@amikom.ac.id, ainulyaqin@amikom.ac.id

Abstract- This research is motivated by the motivation to capture customer experience in purchasing rice products in the Shopee marketplace. The method used is Latent Dirichlet Allocation and analyzed based on several rating categories, namely low rating (negative sentiment), medium rating (neutral sentiment) and high rating (positive sentiment). The results of this study indicate that customer reviews of purchasing rice products in the shopee marketplace have a dominant positive sentiment, followed by negative and neutral sentiment. Negative sentiment consists of the words with the most frequency namely "quality", "delivery", "price" as well as negative words such as ticks and leaks. Positive sentiment consists of the words that appear the most, namely "price", "buy", "quality", "send", "good". The selection results based on the coherence value are visualized with a total of 3 topics, namely those related to the topics of rice quality and price, taste and quality, price and delivery.

Keywords: Product Review, Rice, Shopee, Sentiment, Topic Modelling

Abstrak- Penelitian ini dilatarbelakangi oleh keinginan untuk menangkap pengalaman pelanggan dalam melakukan pembelian produk beras pada marketplace Shopee. Metode yang digunakan adalah Latent Dirichlet Allocation dan dianalisis berdasarkan beberapa kategori rating yaitu rating rendah (sentiment negatif), rating sedang (sentiment netral) dan rating tinggi (sentiment positif). Hasil penelitian ini menunjukkan bahwa ulasan pelanggan pada pembelian produk beras di marketplace shopee memiliki dominansi pada sentiment positif disusul oleh sentiment negative dan netral. Sentimen negative berisi dengan kata dengan frekuensi terbesar yaitu "kualitas", "kirim", "harga" dan juga kata negatif seperti kutu dan lubang. Sentimen positif terdiri atas kata yang muncul paling banyak yaitu "beras", "rasa", "kirim", "bagus" dan "baik". Hasil pemilihan berdasarkan nilai koherensi divisualisasi dengan jumlah topik sebanyak 3 yaitu terkait dengan topik kualitas beras dan harga, rasa dan kualitas serta harga dan pengiriman.

Kata Kunci: Beras, Pemodelan Topik, Sentiment, Shopee, Ulasan Produk

1. Pendahuluan

Saat ini perkembangan bisnis tidak lagi bisa dipisahkan adanya inovasi di bidang teknologi informasi. Salah satu inovasinya adalah teknologi *internet*. Teknologi internet telah memungkinkan peningkatan transaksi perdagangan secara online [1]. Terlebih lagi saat ini dunia masih belum sepenuhnya terbebas dari pandemi Covid-19 yang berakibat pada penurunan mobilitas manusia dan mengalihkan transaksi jual beli melalui *e-commerce* [2]. Keputusan terhadap pilihan produk yang akan dibeli atau dipesan dalam sistem *e-commerce* semakin dipengaruhi oleh umpan balik pelanggan seperti dalam kasus pemesanan hotel [3] maupun belanja produk secara umum [4]. Penelitian ini dilatarbelakangi oleh keinginan untuk menangkap pengalaman pelanggan dalam melakukan

pembelian produk beras pada *marketplace* Shopee. Pilihan terhadap produk beras dikarenakan secara umum produk pertanian memiliki sifat mudah rusak dan tidak tahan lama [5] sehingga menarik untuk dapat diketahui bagaimana ulasan dari para pelanggan tersebut serta tingkat kepuasan selama pembelian dan topik yang dibicarakan. Keluaran yang diharapkan dapat bermanfaat bagi para pelanggan yang akan membeli produk beras secara online serta pelaku bisnis perberasan. *Topic Modelling* adalah salah satu teknik analisis teks untuk menemukan gagasan atau topik yang terjadi dalam kumpulan dokumen dan *Latent Dirichlet Allocation* (LDA) adalah salah satu model yang dapat digunakan [6],[7]. Beberapa penelitian telah melakukan analisis topik pada

Vol.13 no.2 | Desember 2022

EXPLORE : ISSN: 2087-2062, Online ISSN: 2686-181X / DOI: <http://dx.doi.org/10.36448/jsit.v13i2.2348>



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

data ulasan pelanggan ketika berbelanja melalui marketplace dengan cara membaginya berdasarkan ulasan dengan sentimen bernada positif, netral dan negative [8],[9],[10]. Namun demikian ketiga penelitian tersebut mengambil obyek kasus yang bukan produk pertanian.

Oleh karena itu penelitian ini bertujuan untuk melakukan analisis ulasan produk pembelian beras di *Marketplace* Shopee dengan menggunakan teknik LDA dan membagi analisis berdasarkan tiga sentiment yaitu positif, negative dan netral.

2. Metodologi

A. Data

Data penelitian didapatkan melalui *web scraping* dari situs belanja di *marketplace* shopee menggunakan *library selenium* pada bahasa pemrograman python. Kata kunci yang digunakan adalah kata “beras” dengan jumlah yang didapatkan sebanyak 3531 ulasan. Setelah dilakukan pembersihan ulasan yang bersifat duplikat maka tersisa sebanyak 464. Data terdiri atas empat 4 kolom yaitu *url*, *title*, *comment* dan *rating*. Contoh ulasan produk dapat dilihat dilihat pada Tabel 1 dimana hanya ditampilkan dua kolom terakhir. Khusus pada kolom terakhir, data yang

akan diambil adalah adalah nilai angka dari ratingnya saja yang berkisar dari angka 1 sampai dengan 5.

B. Tahapan Penelitian

Tahapan penelitian dapat dilihat pada Gambar 1. Data hasil *scrapping* perlu dilakukan pembersihan data melalui tahapan *preprocessing*. Langkah pertama adalah membuang ulasan yang duplikat, kemudian menghapus kolom *url*, menghapus tanda baca, mengubah ke huruf kecil, tokenisasi, *stemming* sastrawi serta menghilangkan kata yang tidak penting (*stopword*).

Tabel 1. Sampel ulasan produk

Index	Comment	Rating
0	Terimakasih atas paket produknya sudah sampai dengan baik dan benar	5 Bintang (44)
1	Bubble wrapnya tebal banget, beli ini untuk kebutuhan jual nasi goreng	5 Bintang (9)
2	mau beli yg 500gr buat coba2 dulu, tapi dikirim yg 1kg.. jadinya mau ga mau harus dibayar daripada ribet dikembalikan.. kurang cocok buat mpasi, kurang wangi dari yg sebelumnya dibeli	3 Bintang (4)
3	Rasa: blm tau Harga: oke Kualitas: oke Baru mau nyoba makan beras hitam ini semoga enak dan bisa order lagi. Trimakasih	5 Bintang (2,6RB)
4	Rasa: blm dicoba Harga: agak sedikit mahal Kualitas: blm dicoba Packing rapih dan pengiriman lumayan cepat...produk blm dicoba jadi blm tau rasa dan kualitas nya	4 Bintang (174)
5	Pengiriman sih cepat tapi beras nya kurang bagus ya, sama dengan harga yg 19ribu. Kurang suka.....	3 Bintang (17)

Analisis eksploratori dilakukan menampilkan diagram lingkaran yang menunjukkan persentase rating dari 1 hingga 5 serta dengan cara membuat *word cloud* yang menampilkan visualisasi kata dengan ukuran yang sesuai dengan frekuensi penggunaan kata. Analisis *word cloud* ini juga dibagi berdasarkan nilai rating dibagi menjadi 3 yaitu yang memiliki rating rendah atau sentiment negatif (1 atau 2), rating sedang atau sentiment netral (3), dan rating tinggi atau sentiment positif (4 dan 5). Dalam persiapan analisis LDA dilakukan pembentukan *dictionary* dan *corpus* yang mengkonversi dokumen ke dalam format *bag-of-words* menggunakan paket gensim. Setelah *corpus* dan *dictionary* terbentuk maka digunakan untuk membuat model LDA yang terdiri atas 4 model yaitu *lda_model_general* (seluruh sentiment), *lda_model_bottomTwo* (sentiment negative), *lda_model_3* (sentiment netral), dan *lda_model_topTwo* (sentiment positif). Visualisasi model menggunakan *multidimensional scaling* dari *pyLDAvis.gensim_models library*.

C. Metode Latent Dirichlet Allocation (LDA)

LDA adalah bagian dari pemodelan probabilistik dimana variabel yang diamati adalah kata-kata yang terkandung dalam dokumen serta variabel tersembunyi adalah struktur topik. Masalah komputasi untuk menyimpulkan

struktur topik tersembunyi dari dokumen adalah masalah menghitung distribusi *posterior*, yaitu distribusi bersyarat dari variabel tersembunyi yang diberikan dokumen [11]. Formula yang digunakan dalam LDA dapat dilihat pada persamaan (1) dimana β_k melambangkan distribusi topik k sepanjang kamus kata (*vocabulary*), $\theta_{d,k}$ melambangkan proporsi topik k pada dokumen d , $Z_{d,n}$ adalah penugasan topik untuk kata ke- n pada dokumen d , $W_{d,n}$ adalah kata ke- n pada dokumen d .

$$p(\beta_{1:k}, \theta_{1:D}, Z_{1:D}, W_{1:D}) = \prod_{i=1}^K p(\beta_i) \prod_{d=1}^D p(\theta_d) \left[\prod_{n=1}^N p(Z_{d,n} | \theta_d) p(W_{d,n} | \beta_{1:k}, Z_{d,n}) \right] \dots \dots \dots (1)$$

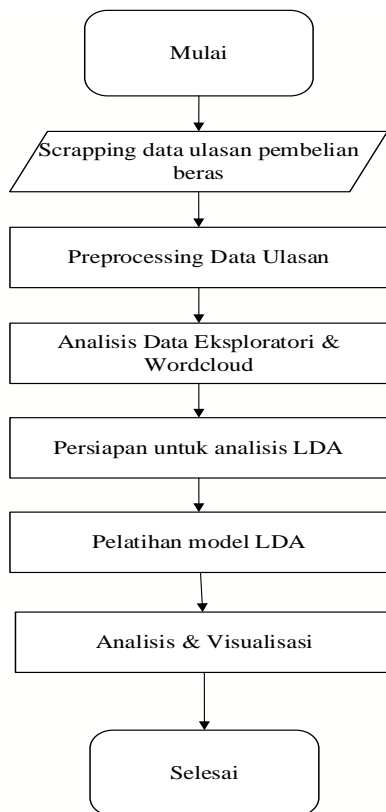
Koherensi topik adalah metrik yang menilai kesamaan semantik antara istilah dan skor yang tinggi di dalam sebuah topik [12] dengan formula seperti dapat dilihat pada persamaan (2). Metrik ini membantu membedakan masalah yang dapat dilihat secara semantik dan masalah yang bersifat inferensi statistik. Jika pernyataan atau fakta tersebut saling mendukung, maka dikatakan koheren. Produk yang koheren dapat menjangkau dalam konteks yang mencakup semua atau hampir semua fakta.

$$C(t; V^{(t)}) = \sum_{m=2}^M \sum_{l=1}^{m-1} \log \frac{D(v_m^{(t)}, v_l^{(t)}) + 1}{D(v_l^{(t)})} \dots \dots \dots (2)$$



. $D(v)$ adalah frekuensi dokumen dari tipe kata v (yaitu, jumlah dokumen dengan setidaknya satu token tipe v) dan $D(v, v')$ adalah frekuensi dokumen bersama dengan tipe kata v dan v' (yaitu, jumlah dokumen yang berisi satu atau lebih token tipe v dan setidaknya satu token tipe v').

$V^{(t)} = (V_1^{(t)}, \dots, V_M^{(t)})$ adalah daftar M kata-kata yang paling memungkinkan ada dalam topik t . Nilai koherensi yang tertinggi akan dijadikan rujukan pembagian jumlah topik terbaik[13].

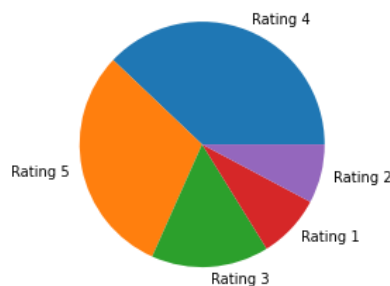


Gambar 1. Tahapan Penelitian

3. Hasil dan Pembahasan

Berdasarkan hasil pembersihan data duplikat dan *pre-processing* didapatkan ulasan sebanyak 464 ulasan dengan sampel hasil seperti dapat dilihat pada Gambar 2. Terlihat bahwa seluruh teks dalam dokumen sudah dalam bentuk huruf kecil, kemudian beberapa kata yang tidak penting seperti kata “dan”, “atas”, “sudah” pada dokumen pertama sudah dihilangkan melalui fungsi *remove stopwords* yang dibangun. Terlihat pula pada dokumen ulasan produk ke-2 pada Gambar 3 kata “kebutuhan” berubah menjadi “butuh” serta pada ulangan produk ke-6

kata “pengiriman” sudah berubah menjadi “kirim” menggunakan proses *stemming* dari sastrawi. Analisis eksploratori menampilkan diagram lingkaran untuk seluruh ulasan maupun kategori sentiment. Ulasan memiliki rating bintang 1 sebanyak 39, rating 2 sebanyak 36, bintang 3 sebanyak 72, bintang 4 sebanyak 141, dan bintang 5 sebanyak 176 ulasan seperti dapat dilihat pada Gambar 2. Terlihat bahwa ulasan didominasi oleh sentiment positif, diikuti oleh sentiment negatif dan terkecil adalah sentiment netral.



Gambar 2. Diagram lingkaran rating ulasan

Vol.13 no.2 | Desember 2022

EXPLORE : ISSN: 2087-2062, Online ISSN: 2686-181X / DOI: <http://dx.doi.org/10.36448/jsit.v13i2.2348>



1 to 25 of 463 entries

comment	ratings	comment_processed
Terimakasih atas paket produknya sudah sampai dengan baik dan benar	5	terimakasih paket produk sampai baik benar
Bubble wrapnya tebal banget beli ini untuk kebutuhan jual nasi goreng	5	bubble wrapnya tebal banget beli butuh jual nasi goreng
mau beli yg 500gr buat coba2 dulu tapi dikirim yg 1kg jadinya mau ga mau harus dibayar daripada ribet dikembalikan kurang cocok buat mpasi kurang wangi dari yg sebelumnya dibeli	3	beli coba kirim ga bayar ribet kurang cocok mpasi kurang wangi beli
Rasa blm tau Harga oke Kualitas oke Baru mau nyoba makan beras hitam ini semoga enak dan bisa order lagi Trimakasih	5	rasa blm harga oke kualitas oke nyoba makan beras hitam moga enak order trimakasih
Rasa blm dicoba Harga agak sedikit mahal Kualitas blm dicoba Packing rapih dan pengiriman lumayan cepat produk blm dicoba jadi blm tau rasa dan kualitas nya	4	rasa blm coba harga mahal kualitas blm coba packing rapih kirim lumayan cepat produk blm coba blm rasa kualitas
Pengiriman sih cepat tapi beras nya kurang bagus ya sama dengan harga ya 19ribu Kurang suka	3	kirim sih cepat beras kurang bagus harga ribu kurang suka

Gambar 3. Tangkapan layar perbandingan data awal dan hasil pemrosesan awal.

Gambar 4 menampilkan *word cloud* untuk seluruh sentiment. Terdapat kata “beras”, “rasa”, “kirim”, “bagus”, “baik” yang memiliki ukuran font lebih besar dan menandakan frekuensi kemunculan yang lebih tinggi dibandingkan kata lainnya serta bernada positif yang memang mendominasi kategori sentiment.. Kemudian *word cloud* dibagi berdasarkan kategori sentiment. Pada rating 1 dan 2 (negatif) seperti terlihat pada Gambar 5, kata kunci yang terbanyak adalah “beras”, “kirim”,

“kualitas”, “harga”, “produk”. “rasa” masih mendominasi. Terdapat pula kata “kutu” yang juga muncul dalam *word cloud* dimana menunjukkan bahwa kata itu berasosiasi dengan kualitas produk beras yang buruk dan pelanggan memberikan sentiment negatif. Kata “baik” dan “bagus” sudah tidak lagi terlihat, sebaliknya muncul kata “kualitas” secara dominan yang artinya pelanggan mempersoalkan kualitas dari beras yang dibeli.



Gambar 4. Wordcloud kata kunci beras untuk seluruh kategori sentimen



Gambar 5. Wordcloud kata dalam rating 1 & 2



Pada gambar 6 menunjukkan hasil wordcloud yang dipersempit menjadi rating 3 dan menunjukkan kata terbanyak yang didapatkan dari data yaitu diantaranya “kualitas”, “beras”, “harga”, “kirin”, “bagus” dan “cepat”. Pada visualisais ini juga masih muncul kata yang

berasosiasi negtif yaitu kata “kecewa” dan “kutu”. Gambar 7 menunjukkan kata-kata yang paling banyak dibicarakan dalam rating 4 dan 5 (positif). Kata-kata yang muncul paling banyak adalah “beras”, “rasa”, “kirin”, “bagus” dan “baik”.



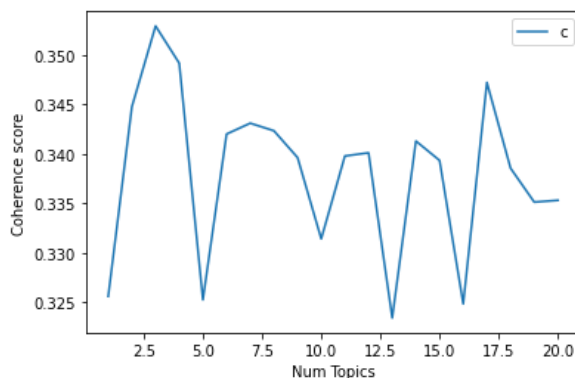
Gambar 6. Wordcloud kata dalam rating 3



Gambar 7. Wordcloud kata dalam rating 4&5

Setelah dilakukan pembagian Wordcloud berdasarkan rating, data scraping dilatih dengan menggunakan LDA menggunakan *library* Gensim. Jumlah topik didasarkan pada nilai skor koherensi tertinggi dengan grafik seperti

dapat dilihat pada Gambar 8. Terlihat bahwa untuk dapat seluruh kategori sentiment dapat diketahui bahwa koherensi tertinggi terdapat pada jumlah topik sebesar 3.



Gambar 8. Grafik skor koherensi untuk seluruh kategori sentiment

Persamaan matematis hasil dari model LDA yang menampilkan koefisien atau bobot pada setiap topik dapat dilihat pada Tabel 2. Angka yang tertera pada keluaran program adalah nilai probabilitas dari kata tersebut termasuk pada topik yang bersangkutan.

Semakin tinggi nilai probabilitasnya, maka semakin tinggi juga keterkaitan kata tersebut dengan topik yang bersangkutan, begitu juga sebaliknya. Hal ini dapat mempengaruhi dalam pembuatan judul topik untuk pengelompokkan kalimat ulasan. Jika dari sebuah kalimat



ulasan terdapat banyak kata yang terdapat pada suatu topik, maka asosiasi kalimat ulasan terhadap topik tersebut akan lebih tinggi dibandingkan topik yang lain. Berdasarkan susunan kata tersebut dapat diketahui bahwa topik pertama terkait dengan produk beras itu sendiri dari sisi kualitas, topik kedua terkait aspek rasa dan topik

ketiga terkait harga. Kemudian visualisasi *multidimensional scaling* dari LDA untuk seluruh kategori dilihat pada Gambar 10 dimana gambar lingkaran merah yang dipilih akan menampilkan susunan kata kunci yang berkontribusi pada penyusunan topik dengan kata paling dominan adalah “rasa”.

Tabel 2. Hasil pembentukan model LDA untuk seluruh kategori sentiment

Topik	Kata
0	0.059*"beras" + 0.027*"harga" + 0.022*"kualitas" + 0.019*"rasa" + 0.016*"bagus" + 0.014*"kirim" + 0.012*"kasih" + 0.012*"merah" + 0.011*"terima" + 0.010*"beli"
1	0.043*"rasa" + 0.043*"kualitas" + 0.038*"harga" + 0.026*"kirim" + 0.021*"cepat" + 0.019*"baik" + 0.018*"enak" + 0.018*"beras" + 0.018*"bagus" + 0.017*"ok"
2	0.041*"harga" + 0.039*"kualitas" + 0.031*"kirim" + 0.027*"rasa" + 0.025*"baik" + 0.021*"produk" + 0.019*"sesuai" + 0.018*"cepat" + 0.016*"terima" + 0.013*"kasih"

Analisis jumlah topik kemudian diperdalam untuk kategori sentiment yang memang menjadi prioritas untuk perbaikan. Gambar 9 memperlihatkan grafik koherensi dari sentiment negative dimana nilai tertinggi terdapat pada jumlah topik sebanyak 13. Namun dalam penelitian

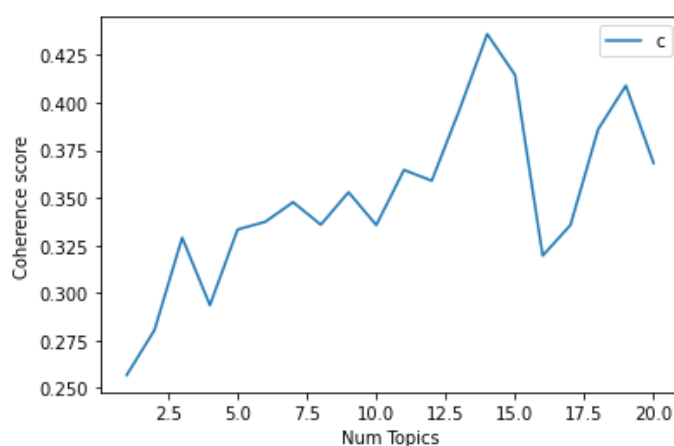
ini dipilih maksimum lokal sebanyak 3 topik dimana grafik menaik kemudian menurun pada 4 topik pertama. Hasil pembentukan model LDA pada kategori sentiment negative dapat dilihat pada Tabel 3.

Tabel 3. Hasil pembentukan model LDA untuk kategori sentiment negatif

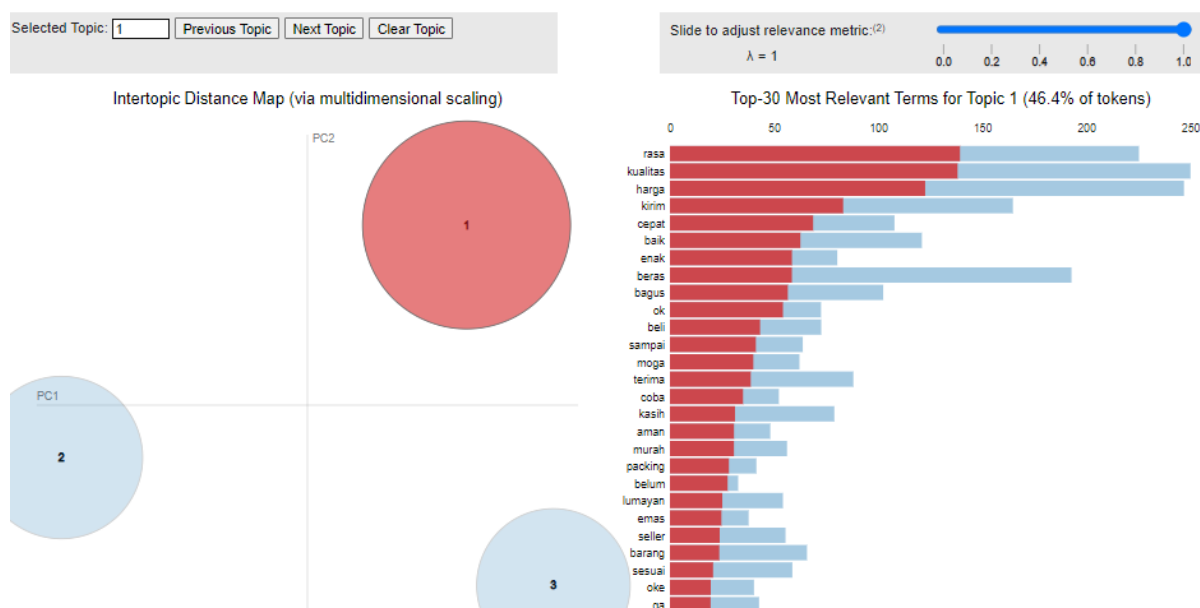
Topik	Kata
0	0.024*"beras" + 0.024*"kirim" + 0.019*"barang" + 0.017*"harga" + 0.016*"kualitas" + 0.015*"merah" + 0.015*"jual" + 0.013*"rasa" + 0.010*"kasih" + 0.008*"beli"
1	0.038*"beras" + 0.032*"kualitas" + 0.024*"harga" + 0.021*"kirim" + 0.019*"beli" + 0.018*"rasa" + 0.017*"produk" + 0.013*"kutu" + 0.012*"jual" + 0.011*"ga"
2	0.026*"kirim" + 0.016*"jual" + 0.014*"baik" + 0.013*"kecewa" + 0.012*"lubang" + 0.011*"beras" + 0.011*"kualitas" + 0.009*"tdk" + 0.009*"rasa" + 0.009*"harga"

Tabel 3 memperlihatkan 3 topik dengan kata kunci yang paling berkontribusi untuk sentiment negative dimana topik pertama dan kedua sama-sama didominasi oleh kata “beras” namun disusul oleh kata kunci yang berbeda yaitu “kirim” pada topik pertama dan kualitas pada topik

“kedua”. Kualitas ini juga terkait dengan kata “kutu”. Sedangkan topik ketiga didominasi oleh kata “kirim” dengan probabilitas paling tinggi dan kata “harga” dengan probabilitas paling kecil dan 10 kata kunci teratas.



Gambar 9. Grafik skor koherensi untuk kategori sentiment negative



Gambar 10. Visualisasi LDA untuk seluruh kategori sentiment

4. Kesimpulan

Hasil penelitian ini menunjukkan bahwa ulasan pelanggan pada pembelian produk beras di *marketplace* shopee memiliki dominansi pada sentiment positif disusul oleh sentiment negatif dan netral dengan jumlah relative sama. Terdapat beberapa kata yang muncul pada sentimen negatif pada *word cloud* yang menunjukkan aspek yang dikeluhkan yaitu “kualitas”, “kirim”, “harga” dan juga

kata negatif seperti kutu, lubang dan kecewa. Sedangkan sentimen positif terdiri atas kata yang muncul paling banyak yaitu “beras”, “rasa”, “kirim”, “bagus” dan “baik”. Hasil pemilihan jumlah topik untuk seluruh kategori sentiment divisualisasi dengan jumlah topik sebanyak 3 yaitu kualitas beras dan harga, rasa dan kualitas serta harga dan pengiriman.

5. Daftar Pustaka

- [1] A. S. Nisafani, A. Wibisono, and M. H. T. Revaldo, “Analyzing the Effectiveness of Public e-Marketplaces for Selling Apparel Products in Indonesia,” *Procedia Comput. Sci.*, vol. 124, no. April 2019, pp. 274–279, 2017, doi: 10.1016/j.procs.2017.12.156.
- [2] N. A. Rakhmawati, A. E. Permana, A. M. Reyhan, and H. Rafli, “Analisa Transaksi Belanja Online Pada Masa Pandemi Covid-19,” *J. Teknoinfo*, vol. 15, no. 1, p. 32, 2021, doi: 10.33365/jti.v15i1.868.
- [3] S. Molinillo, J. L. Ximénez-de-Sandoval, A. Fernández-Morales, and A. Coca-Stefaniak, “Hotel Assessment through Social Media: The case of TripAdvisor,” *Tour. Manag. Stud.*, vol. 12, no. 1, pp. 15–24, 2016, doi: 10.18089/tms.2016.12102.
- [4] G. Gesitera, “Pengaruh Online Customer Review terhadap Purchase Intention dengan Trust sebagai Intervening pada Toko Online Bukalapak di Kota Padang,” vol. 9, no. 2, pp. 173–194, 2020.
- [5] Arifin Rente, *Pengantar agroindustri*, no. August. 2018.
- [6] B. Liu, *Sentiment Analysis and Opinion Mining*, no. May. 2012.
- [7] D. Blei, L. Carin, and D. Dunson, “Probabilistic topic models,” *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 55–65, 2010, doi: 10.1109/MSP.2010.938079.
- [8] Y. Putranto, A. Djuraidah, and B. Sartono, “Topic Modelling and Hotel Rating Prediction based on Customer Review in Indonesia Topic modelling and hotel rating prediction based on customer review in Indonesia,” no. January, 2021, doi: 10.1504/IJMDM.2021.10036033.
- [9] K. Dheanis, A. Salsabila, and N. Trianasari, “Jurnal Teknologi dan Manajemen Informatika Analisis Persepsi Produk Kosmetik Menggunakan Metode Sentiment Analysis Dan Topic Modeling (Studi Kasus : Laneige Water Sleeping Mask),” vol. 7, no. 1, pp. 1–9, 2021.
- [10] M. Y. Febrianta and S. Widiyanesti, “Analisis Ulasan Indie Video Game Lokal Pada Steam Menggunakan Sentiment Analysis Dengan Algoritma Naïve Bayes Classifier Dan Lda-Based Topic Modeling Review Analysis Of Local Indie Video Games On Steam Using Sentiment Analysis With Naïve Bayes Classifier,” in *e-Proceeding of Management*, 2021, vol. 8,

- no. 4, pp. 3102–3109.
- [11] Z. Liu, Y. Lin, and M. Sun, *Representation Learning for Natural Language Processing*. 2020.
- [12] D. Mimno, H. M. Wallach, E. Talley, and M. Leenders, “Optimizing Semantic Coherence in Topic Models,” in *Conference on Empirical Methods in Natural Language Processing*, 2011, no. 2, pp. 262–272.
- [13] T. K. Kurniasari, W. Maharani, and J. H. Husen, “Analisis Media Sosial Twitter untuk Mengetahui Pengguna Berpengaruh pada Portal Berita di Indonesia dengan Metode TSIM (Topic-based Social Influence Measurment) Pendahuluan,” vol. 7, no. 3, pp. 10043–10060, 2020.

