

Analisis pada dataset tokopedia *product reviews category food and drink* menggunakan metode Sentiment Analysis dan Top-IC Modelling dengan LDA

Herjuna Ardi Prakosa¹

¹ Program Studi Informatika, Fakultas Sain dan Teknik, Universitas Perwira Purbalingga, Indonesia

Penulis Korespondensi : Herjuna Ardi Prakosa (herjuna@unperba.ac.id)

ABSTRAK

Tokopedia sebagai salah satu online shop terbesar di Indonesia tentu saja menyediakan banyak barang termasuk makanan dan minuman. Berbagai pendapat yang disampaikan oleh para konsumen berupa review merekomendasikan, puas, dan kecewa. Review dengan volume yang sangat banyak perlu dianalisis secara akurat agar dapat dipisahkan antara sentiment positif dan negatif. Tujuan penelitian ini adalah untuk mengetahui reaksi senti-ment para konsumen Tokopedia terhadap barang mereka yang berupa Food and Drink melalui data set dari website *Kaggle* yang diambil dari hasil review product pada kolom re-view comment barang pada Tokopedia itu sendiri. Data yang diperoleh dari data set tersebut berjumlah 5427 data dan akan diolah menggunakan metode *Sentiment Analysis* dan *Topic Modeling*.

KATA KUNCI: Tokopedia; *Food and Drink*; *Sentiment Analysis*; *Topic Modeling*

1. PENDAHULUAN

Perkembangan teknologi dalam bentuk tersedianya koneksi internet telah merubah cara konsumen untuk membeli makanan dan minuman. Proses adopsi teknologi dalam bisnis makanan telah memfasilitasi munculnya system pemesanan dan pengantaran makanan secara daring (*online*) melalui berbagai pasar daring (*marketplace*), aplikasi aggregator makanan, pengantaran restoran, dan media social[1]. menurut data Badan Pusat Statistik 2019, jasa pesan antar makanan secara daring tersebut setiap tahunnya diperkirakan tumbuh hingga mencapai 11,5% (2020-2024). Lebih lanjut, kategori makanan menyumbang proporsi terbesar dari semua kategori *e-commerce* dengan kontribusi sebesar 27,85% [2]

Belanja online atau *E-Commerce* adalah sebuah proses transaksi yang dilakukan melalui media atau perantara yaitu berupa situs-situs jual beli online ataupun jejaring sosial yang menyediakan barang atau jasa yang diperjual belikan[3]. Kini belanja online telah menjadi sebuah kebiasaan bagi Sebagian orang, dikarenakan kemudahan yang diberikan, orang-orang banyak beranggapan bahwa belanja online adalah salah satu sarana untuk mencari barang-barang yang diperlukan seperti kebutuhan sehari-hari, hobi, dan sebagainya. Belanja online juga dapat diartikan sebagai keinginan konsumen untuk membelanjakan uangnya

untuk mendapatkan sesuatu yang diinginkan di toko online. Proses tersebut dapat dilakukan dengan cara memesan barang yang diinginkan melalui vendor atau produsen serta *reseller* dengan menggunakan internet. Selanjutnya melakukan pembayaran dengan cara mentransfer via bank, e-bank, ataupun COD (*Cash on Delivery*)[4]. Perilaku belanja online di Indonesia sangat beragam. Perilaku setiap konsumen dalam memutuskan membeli suatu produk menjadi kajian khusus setiap perusahaan sebelum melepaskan produknya ke pasar. Perkembangan era digital semakin tak terhindarkan yang harus diikuti setiap perusahaan menyesuaikan strategi pemasarannya dengan memasuki sistem online untuk menjual produknya[5]. Belanja online menjadi suatu kebiasaan bagi sebagian orang karena kemudahan yang diberikan, orang-orang banyak beranggapan bahwa belanja online adalah salah satu sarana untuk mencari barang-barang yang diperlukannya. Metode penelitian yang dipakai adalah membandingkan hasil-hasil penelitian dan jurnal yang meneliti tentang belanja online di Indonesia. Kemudian ditelaah dan dikaji teori-teori perilaku konsumen yang telah ada sehingga dapat disimpulkan pertimbangan konsumen berbelanja online di sebuah *online shop*. Hasil temuan dari penelitian sebelumnya menunjukkan banyak faktor yang mempengaruhinya. Hasil kajian dari beberapa penelitian sebelumnya, bisa menjadi acuan dan

pertimbangan bagi online shop yang ada di Indonesia dalam menggaet dan mempertahankan pelanggannya agar tetap berbelanja di tokonya sehingga tokonya diminati dan disukai pembelinya [6].

Salah satu cara untuk mengetahui perilaku para konsumen adalah dengan melalui review product atau *Online Customer Reviews* (OCR), dengan mengumpulkan data review tersebut menjadi sebuah data dengan volume yang besar yang kemudian di cari apakah perilaku tersebut bersentiment positif atau bersentiment negatif[7].

Dari banyak nya product dan *marketplace* yang ada di Indonesia saya memilih Tokopedia pada kategori *food and drink* sebagai bahan untuk penelitian berupa *Sentiment Analysis* menggunakan metode *Sentiment Analysis dan Topic Modeling*. Data review product yang di ambil berjumlah 5427 data yang termasuk dalam kategori big data yang kemudian akan dilakukan pemrosesan data.

2. BAHAN DAN METODE

Penelitian ini menggunakan data yang bersumber dari kolom *review product e-commerce* Tokopedia yang di upload ke website Kaggle sebanyak 5427 comment review. Gambar 1. Menunjukkan *machine learning orange* 3. Teknik analisi yang di lakukan dalam penelitian ini adalah sebagai berikut:

1. Pengolahan data

Pada tahap pengolahan data ini ada beberapa hal yang di lakukan antara lain :

- Url Remove: merupakan proses penghilangan alamat alamat Url.
- Case Folding merupakan proses penyeragaman semua huruf menjadi huruf kecil dan penghilangan tanda baca.
- Tokenizing: merupakan proses pemisahan kalimat menjadi kata independent.
- Filtering: merupakan proses pengurangan kata yang tidak digunakan sesuai dengann stopwords
- Stemming: merupakan proses mentransformasi kata ke bentuk baku nya.

2. Sentiment Analysis

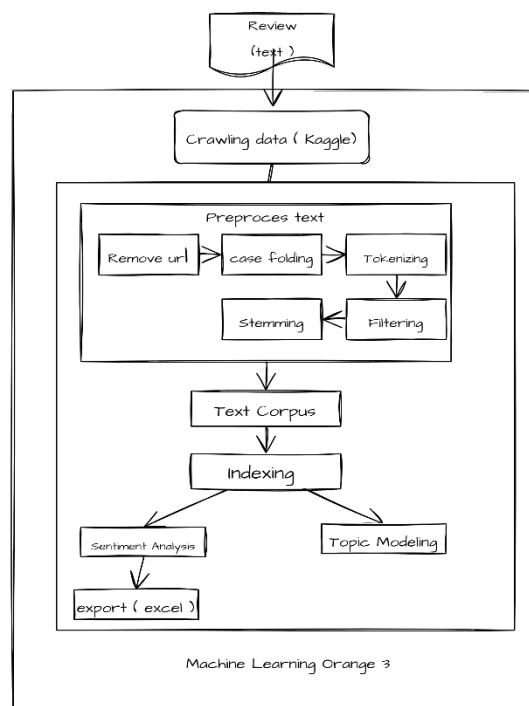
Proses menentukan sentiment dan meneglompokkan polaritas teks sehingga dapat di kategorikan antara sentiment positif atau sentiment negative[8].

3. Topic Modeling

Merupakan mengelompokkan unit dasar diskrit dari kata dalam sebuah dokumen

3. HASIL DAN PEMBAHASAN

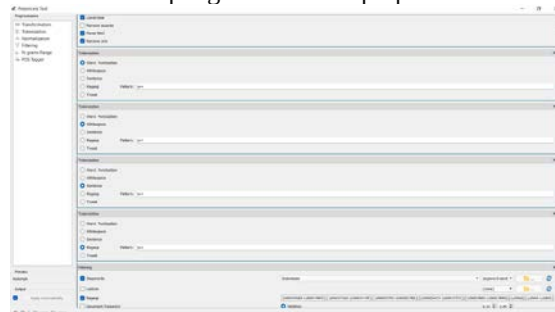
Tahapan awal untuk penelitian ini adalah pengolahan data yang telah diambil dari website Kaggle, berikut kerangka untuk pengolahan data nya:



Gambar 1. Machine Learning Orange 3

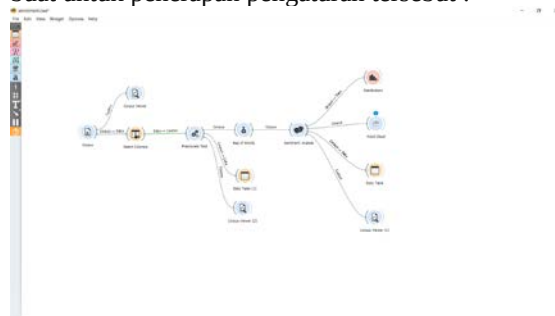
1. Sentiment Analysis

Berdasarkan kerangka pengolahan data pada gambar 1 proses dilakukan menggunakan Machine learning orange 3[9]. dimana dataset yang didapat dari website Kaggle pertama akan di proses di preprocess text, dan berikut adalah pengaturan untuk preprocess text:



Gambar 1 pengaturan preprocessing text

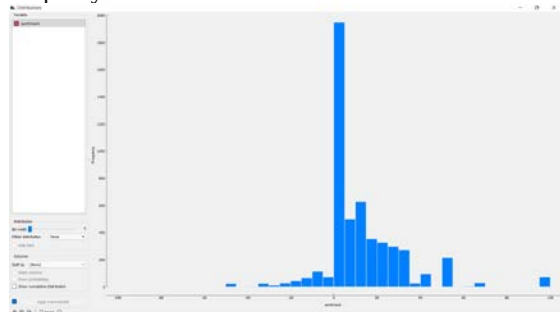
Untuk pengaturan preprocess text tersebut akan di masukan dalam rangkaian dalam machine learning Orange 3, dan berikut adalah rangkaian yang akan di buat untuk penerapan pengaturan tersebut :



Gambar 2 rangkaian Sentiment Analysis

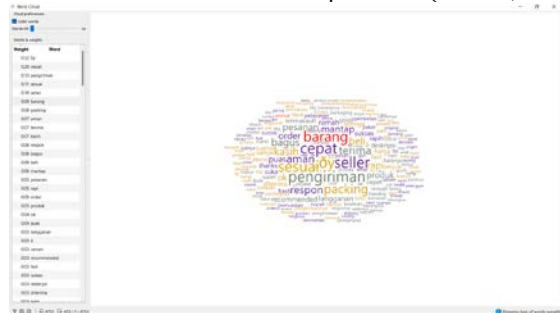
2.Hasil Sentiment Analysis

Pada gambar 3. Menunjukkan rangkaian tersebut adalah rangkaian untuk mencari sentimen dari dataset *tokopedia product reviews category food and drink* rangkaian tersebut akan menghasilkan beberapa output diantaranya *distribution sentiment*, *word cloud*, dan hasil sentimen berikut adalah hasil dari masing masing outputnya:



Gambar 3. distribution sentiment

Output tersebut menghasilkan indikasi banyak nya sentimen yang ada pada dataset yang sudah kita proses dimana “minus berarti negatif “ , “ nol berarti netral “ , dan “diatas nol berarti positif. (Hafidh, 2021)



Gambar 4. word Cloud

Output tersebut adalah word cloud dimana memvisualisasikan kata yang sering muncul dihitung dari bobot kata tersebut.



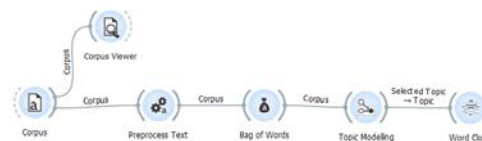
Gambar 5. data table sentiment

Hasil output tersebut adalah hasil output yang menampilkan bobot dari setiap kata yang telah di visualisasikan di word cloud dan juga angka sentiment yang kemudian akan di individualisasikan di distribution. Setelah menghitung jumlah sentiment di dapatkan angka “Positif :2903,”Neutral:585”, “Negatif :572 “.

3.Topic Modeling

Setelah data tersebut di pisah kan menjadi positif ,netral, dan negatif maka selanjutnya akan dilakukan

permodelan berikut adalah rangkaian dari topic modeling yang akan digunakan:



Gambar 6. rangkaian topic model

4. Hasil Topic Modeling

Setelah dilakukan pengolahan dari rangkaian pada gambar 7. proses tersebut menghasilkan data sebagai berikut :

- ### 1. Topic modeling Positif

Topic	Topic keywords
-------	----------------

- 1 nya, renyah, langganan, bgt, suka, terkemah, tix, sampe, diterima, pas
2 pengiriman, seller, packing, barang, respon, produk, beli, fast, ok, recommended
3 deskripsi, harga, ga, kacangnya, jd, delivery, sm, tokped, kemaren, snack
4 mantap, pesanan, kasih, yg, ya, belian, cepet, ga, udah, deh
5 banget, makasih, n, garing, top, cemilan, pelayanannya, kriuk, dikemas, tp

Gambar 7 Topic Modeling Positif



Gambar 8. Word cloud positif

Pada topic modeling sentiment positif topic yang sering muncul ialah tentang pengiriman yang vepat, respon penjual, packing, dan keadaan produk makanan tersebut[9]. Dari kata-kata yang muncul tersebut dapat disimpulkan bahwa target penelitian ini memberikan tanggapan yang positif atau bisa dikatakan puas akan barang atau pelayanannya.

- ## 2. Topic modeling negatif

Topic	Topic keywords
-------	----------------

1. lecewa, lengket, rentan, kerusakan, kebocoran, jik, patah, kendur, uluh, pahit
2. bau, penyok, berat, apek, seret, serampangan, mengecewakan, asem, retak, penuh
3. parah, pecah, susah, serampangan, mengecewakan, merata, bosan, buar, tipis, lambat
4. keras, tebal, menyah, bocor, terlambat, berminyak, rusak, anah, sedih, gila
5. hancur, mahal, berbau, kering, rapuh, mengurangi, seret, mengecewakan, serampangan, disayankan

Gambar 9 Topic modeling negatif



Gambar 10 Word cloud negatif

Pada topic modeling sentiment positif topic yang sering muncul ialah tentang kerusakan, kecewa, lengket, kebocoran, jijik, patah. Yang mengdeskripsikan bahwa target penelitian ini memberikan tanggapan yang negatif atau bisa di sebut mereka kecewa.

3. Topic modeling netral

Topic	Topic keywords
1	kekurangan, puas, asem, lengkap, goreng, lancar, murah, mahal, kecewa, susah
2	mahal, murah, luka, sakit, rendah, asem, lengkap, bocor, lancar, goreng
3	lengkap, susah, habis, asem, goreng, kekurangan, lancar, mahal, murah, terbukti
4	asem, setia, terbukti, lengkap, goreng, kekurangan, lancar, murah, mahal, susah
5	lancar, goreng, salah, asem, lengkap, kekurangan, mahal, murah, susah, terbukti

Gambar 11 Topic modeling netral



Gambar 12. Word cloud netral

Pada topic modeling sentiment netral, topic yang sering muncul ialah puas, berantakan, salah, bocor, murah, kecewa, muntah, lambat. [10] Dari kata-kata tersebut dapat diambil kesimpulan bahwa target penelitian ini memberikan tanggapan positif dan negatif secara bersamaan dan hal tersebut di hitung oleh Machine Learning sebagai keadaan dengan sentiment 0 atau netral.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat ditarik kesimpulan bahwa:

- Hasil sentimen para konsumen terhadap product Tokopedia Category Food And Drink ialah "Positif :2903", "Netral:585", "Negatif :572". Yang dapat diartikan bahwa banyak respon positif pada product tersebut. Namun banyak juga yang memberikan respon kekecewaan yang rata-rata didominasi oleh kualitas product yang buruk.
- Untuk topic dan kata yang sering muncul dalam review product di Tokopedia Category Food And Drink dibagi menjadi 3. Topic pertama (positif) membicarakan tentang kualitas barang pengemasan, respon, dan pengiriman barang yang baik. Sedangkan topic kedua (negatif) membicarakan tentang kekecewaan akan product tersebut. Sedangkan topic ketiga (netral) Bagian kesimpulan harus dapat menjawab tujuan dari penelitian. Jelaskan hasil – hasil penting yang didapatkan yang terkait dengan

tujuan penelitian. Selain itu bagian ini juga menjelaskan tentang keterbatasan penelitian yang dilakukan dan gagasan penelitian berikutnya.

5. DAFTAR PUSTAKA

- J. D. S. Amory and M. Mudo, "Transformasi ekonomi digital dan evolusi pola konsumsi: Tinjauan literatur tentang perubahan perilaku belanja di era internet," *J. Minfo Polgan*, vol. 14, no. 1, pp. 28–37, 2025.
- I. Aprilianti and F. Amanta, "Memajukan Keamanan Pangan pada Layanan Pesan Antar Makanan Daring di Indonesia," 2020.
- Y. I. Christian, "Itikad Baik Dalam Klausula Baku Pada E-Commerce," *Res. J. Soc. Sci. Econ.*, vol. 1, no. 2, pp. 85–95, 2025.
- M. Yusuf and N. Ifada, *E-commerce: Konsep dan Teknologi*. Media Nusa Creative (MNC Publishing), 2022.
- D. P. Siregar and J. Putri, "Strategi pemasaran dan pengembangan bisnis halal di era digital," *AL-IQTISHAD J. Perbank. Syariah dan Ekon. Islam*, vol. 4, no. 2, pp. 29–44, 2024.
- D. A. Harahap, "Perilaku belanja online di Indonesia: Studi kasus," *JRMSI-Jurnal Ris. Manaj. Sains Indones.*, vol. 9, no. 2, pp. 193–213, 2018.
- L. N. K. Pasi and B. Sudaryanto, "Analisis pengaruh online customer reviews dan kualitas pelayanan terhadap keputusan pembelian dengan kepercayaan sebagai variabel intervening (studi pada konsumen Shopee di Kota Semarang)," *Diponegoro J. Manag.*, vol. 10, no. 4, 2021.
- M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731–5780, 2022.
- F. R. Junaedi, D. F. Zahra, and T. S. Ardan, "Analisis Ulasan Pembelian Produk Elektronik di Marketplace Tokopedia dengan menggunakan Topic Modelling," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 10, no. 4, 2023.
- L. A. Saputra, N. A. Sukarno, and S. Zulaehah, "Pengaruh Interlayer Elektroplating Zinc pada Kekuatan Mekanik Friction Stir Spot Welding Aa1100-Ss400," *Semesta Tek.*, vol. 24, no. 2, pp. 93–101, Nov. 2021.