

SKRIPSI

**IMPLEMENTASI *DATA MINING* UNTUK ANALISIS
KEPUASAN KONSUMEN TERHADAP *E-COMMERCE*
SHOPEE BERDASARKAN DIMENSI *E-SATISFACTION***

**(*IMPLEMENTING DATA MINING FOR CUSTOMER
SATISFACTION ANALYSIS E-COMMERCE SHOPEE BASED
ON E-SATISFACTION DIMENSIONS*)**

Disusun dan diajukan oleh:

MUHAMMAD FAHRYL A. YANI

D0711 91 066



**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

SKRIPSI

**IMPLEMENTASI *DATA MINING* UNTUK ANALISIS
KEPUASAN KONSUMEN TERHADAP *E-COMMERCE*
SHOPEE BERDASARKAN DIMENSI *E-SATISFACTION***

***(IMPLEMENTING DATA MINING FOR CUSTOMER
SATISFACTION ANALYSIS E-COMMERCE SHOPEE BASED
ON E-SATISFACTION DIMENSIONS)***

Disusun dan diajukan oleh:

MUHAMMAD FAHRYL A. YANI

D0711 91 066



**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

LEMBAR PENGESAHAN SKRIPSI
IMPLEMENTASI DATA MINING UNTUK ANALISIS
KEPUASAN KONSUMEN TERHADAP E-COMMERCE
SHOPEE BERDASARKAN DIMENSI E-SATISFACTION

Disusun dan diajukan oleh

MUHAMMAD FAHRYL A. YANI
D071191066

Telah dipertahankan dihadapan Panitia Ujian yang dibentuk dalam rangka penyelesaian Studi Program Sarjana Program Studi Teknik Industri Fakultas Teknik Universitas Hasanuddin Pada tanggal 22 Agustus 2023 dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,


Dr. Ir. Saiful, ST., MT., IPM., ASEAN.Eng
NIP. 19810606 200604 1 004


Prof. Madya Dr. Loh Wei Ping
Pensyarah
Pusat Pengajian Kejuruteraan Mekanik
Kampus Kejuruteraan
Universiti Sains Malaysia
14300 Nibong Tebal, Pulau Pinang.

Ketua Program Studi Teknik Industri
Fakultas Teknik Universitas Hasanuddin



Ir. Kifayah Amar, S.T., M.Sc., Ph.D., IPU
NIP. 19740621 200604 2 001

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Muhammad Fahryl A. Yani
NIM : D071191066
Program Studi : Teknik Industri
Jenjang : S1

Menyatakan dengan ini bahwa karya tulisan saya berjudul

Implementasi *Data Mining* Untuk Analisis Kepuasan Konsumen Terhadap *E-Commerce* Shopee Berdasarkan Dimensi *E-Satisfaction*

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilan alihan tulisan orang lain dan bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Semua Informasi yang ditulis dalam skripsi yang berasal dari penulis lain telah diberi penghargaan, yakni dengan mengutip sumber dan tahun penerbitannya. Oleh karena itu semua tulisan dalam skripsi ini sepenuhnya menjadi tanggung jawab penulis. Apabila ada pihak manapun yang merasa ada kesamaan judul dan atau hasil temuan dalam skripsi ini, maka penulis siap untuk diklarifikasi dan mempertanggungjawabkan segala resiko.

Segala data dan informasi yang diperoleh selama proses pembuatan skripsi, yang akan dipublikasi oleh Penulis di masa depan harus mendapat persetujuan dari Dosen Pembimbing

Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan isi skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Makassar, 22 Agustus 2023

Yang Menyatakan Tanda Tangan,



Muhammad Fahryl A. Yani

ABSTRAK

MUHAMMAD FAHRYL A. YANI. *Implementasi Penambahan Data untuk Analisis Kepuasan Konsumen terhadap E-Commerce Shopee berdasarkan Dimensi E-Satisfaction* (dibimbing oleh Saiful dan Loh)

Tren *e-commerce* yang sangat pesat di Indonesia menjadikan salah satu opsi untuk meningkatkan pertumbuhan ekonomi. Shopee yang menjadi salah satu *e-commerce* dengan pengguna terbanyak di Indonesia dalam melakukan pembelian online yang diiringi kompetitor yang ada. Penilaian berupa ulasan dalam bentuk skor bintang yang masih menjadi perhatian peneliti untuk dapat melihat seperti apa persepsi yang diharapkan dalam konsep kepuasan konsumen. Penelitian ini bertujuan untuk mengetahui sentimen pengguna terhadap kepuasan pelanggan dan topik-topik yang terbentuk berdasarkan dimensi *e-satisfaction* di Shopee, serta rekomendasi perbaikan yang dapat dilakukan. Metode penelitian menggunakan analisis sentimen dan pemodelan topik. Hasil yang ditunjukkan bahwa 1) sentimen positif sebesar 98% dan sentimen negatif 2% yang dimana pada sentimen negatif lebih dominan pada dimensi *serviceability* sebesar 78% dan pada sentimen positif lebih dominan pada dimensi *convenience* sebesar 55%. Maka dari itu dengan melihat hal tersebut dimensi tersebut yang sangat banyak dikomentari melalui ulasan oleh pengguna *e-commerce* Shopee. 2) Pemodelan topik yang terbentuk adalah sistem Shopee yang sering *error*, penawaran yang diberikan tidak sesuai, fitur-fitur yang ada pada aplikasi Shopee yang kurang fungsional dan informatif, penipuan yang mengatasnamakan Shopee dan pengiriman yang *stuck* dan dana yang terlambat masuk. 3) Rekomendasi perbaikan yang dapat dilihat adalah melakukan perawatan secara berkala, memberikan informasi mengenai penawaran yang sesuai, menghapus fitur yang tidak dibutuhkan, meningkatkan keamanan dan Shopee tidak perlu bertindak sebagai perantara yang juga menyediakan jasa pengiriman.

Kata Kunci: *E-Commerce*, Kepuasan Konsumen, Shopee, Analisis Sentimen, Pemodelan Topik

ABSTRACT

MUHAMMAD FAHRYL A. YANI. *Implementing Data Mining For Customer Satisfaction Analysis E-Commerce Shopee Based on E-Satisfaction Dimensions (supervised by Saiful and Loh)*

The trend of the rapid growth of e-commerce in Indonesia makes it one of the options for increasing economic growth. Shopee, which is one of the e-commerce with the most users in Indonesia, does online shopping accompanied by existing competitors. Ratings in the form of reviews of star scores are still a concern for researchers to be able to see what the expected perception is in the concept of customer satisfaction. This study aims to determine user sentiment towards customer satisfaction and topics that are formed based on e-satisfaction dimensions on Shopee, as well recommendations for improvements that can be made. The research method uses sentiment analysis and topic modeling. The results show that 1) positive sentiment is 98% and negative sentiment is 2%, where negative sentiment is more dominant in the serviceability dimension at 78% and positive sentiment is more dominant in the convenience dimension at 55%. Therefore, by looking at this dimension which is very much commented on through reviews by Shopee users. 2) Topic modeling on Shopee consist of complaints about Shopee systems that often error, the offers provided are not suitable, the features in the Shopee application are less functional and informative, fraud on behalf of Shopee and stuck shipments and late incoming funds. 3) Recommendations for improvements that can be seen conducting regular maintenance, provide information on suitable offers, removing the features that are needed, increasing security and Shopee not needing to act as an intermediary who also provides shipping services.

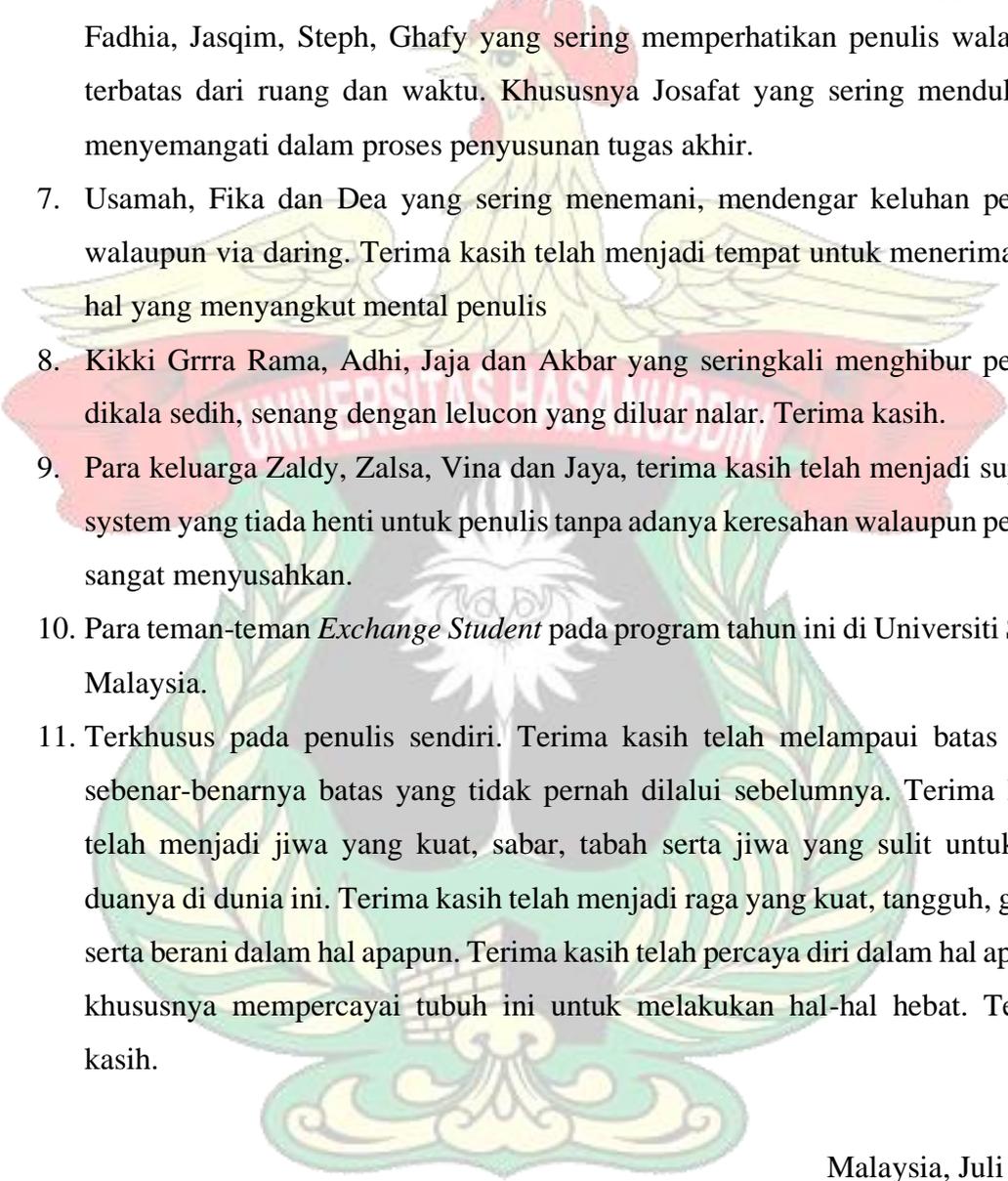
Keywords: *E-Commerce, Customer Satisfaction, Shopee, Sentimen Analysis, Topic Modeling*

KATA PENGANTAR

Bismillahirrahmanirrahim. Puji syukur saya panjatkan kehadiran Allah Subhanallahu Wa Ta' Ala karena atas berkah, rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan tugas akhir ini dengan judul “Implementasi Data Mining untuk Analisis Kepuasan Konsumen Terhadap *E-Commerce* Shopee”. Sholawat serta salam semoga tetap tercurahkan kepada baginda Nabi Muhammad Shallallahu Alaihi Wasallam, yang telah menjadi rahmatan bagi seluruh umat manusia.

Penulisan tugas akhir ini dilakukan dalam rangka memenuhi salah satu persyaratan untuk memperoleh gelar Sarjana pada Departemen Teknik Industri Fakultas Teknik Universitas Hasanuddin. Penulis menyadari bahwa dalam proses penyelesaiannya tidak lepas dari doa, dukungan, bantuan, saran dan kritik yang didapatkan oleh berbagai pihak. Oleh karena itu, dalam kesempatan ini peneliti ingin mengucapkan banyak terima kasih kepada:

1. Kedua orang tua saya, Bapak Achmad Yani HM dan Ibu Hasrat Maula yang senantiasa berdoa, merawat sedari kecil, mengajarkan dinamika kehidupan yang sangat kompleks, menanamkan ilmu kehidupan yang adaptif serta normatif dan yang terpenting mendidik penulis hingga berhasil sampai di titik ini.
2. Ibu Ir. Kifayah Amar, S.T., M.Sc., Ph.D, IPU selaku Ketua Departemen Teknik Industri Fakultas Teknik Universitas Hasanuddin
3. Bapak Dr. Ir. Saiful Mangangenre, S.T., M.T., IPM selaku pembimbing I dalam menyusun tugas akhir ini, terima kasih telah menjadi orang tua kedua di lingkup kampus, mempercayai penulis serta bimbingan dan bantuan selama proses penyusunan tugas akhir ini dimulai dari awal hingga selesai.
4. Profesor Madya Dr. Loh Wei Ping dari Universiti Sains Malaysia selaku pembimbing II dalam menyusun tugas akhir ini, terima kasih telah menjadi orang tua selama penulis berada di Negara Malaysia serta bimbingan dan bantuan selama proses penyusunan tugas akhir ini dimulai hingga selesai.

- 
5. Izul dan Fera sebagai teman seperjuangan sebagai *Exchange Student* di Negeri Jiran ini. Terima kasih telah saling mendukung, membantu, mengobati dari dalam dan luar.
 6. Sahabat OTW Sarjana Hitz Josafat, Dzaky, Una, Nini, Caca, Appi, Kiky, Fadhia, Jasqim, Steph, Ghafy yang sering memperhatikan penulis walaupun terbatas dari ruang dan waktu. Khususnya Josafat yang sering mendukung, menyemangati dalam proses penyusunan tugas akhir.
 7. Usamah, Fika dan Dea yang sering menemani, mendengar keluhan penulis walaupun via daring. Terima kasih telah menjadi tempat untuk menerima hal-hal yang menyangkut mental penulis
 8. Kikki Grrra Rama, Adhi, Jaja dan Akbar yang seringkali menghibur penulis dikala sedih, senang dengan lelucon yang diluar nalar. Terima kasih.
 9. Para keluarga Zaldy, Zalsa, Vina dan Jaya, terima kasih telah menjadi support system yang tiada henti untuk penulis tanpa adanya keresahan walaupun penulis sangat menyusahkan.
 10. Para teman-teman *Exchange Student* pada program tahun ini di Universiti Sains Malaysia.
 11. Terkhusus pada penulis sendiri. Terima kasih telah melampaui batas yang sebenar-benarnya batas yang tidak pernah dilalui sebelumnya. Terima kasih telah menjadi jiwa yang kuat, sabar, tabah serta jiwa yang sulit untuk ada duanya di dunia ini. Terima kasih telah menjadi raga yang kuat, tangguh, gagah serta berani dalam hal apapun. Terima kasih telah percaya diri dalam hal apapun khususnya mempercayai tubuh ini untuk melakukan hal-hal hebat. Terima kasih.

Malaysia, Juli 2023

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
PERNYATAAN KEASLIAN.....	ii
ABSTRAK	iii
KATA PENGANTAR.....	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL.....	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	5
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Kepuasan Konsumen	7
2.2 E-Satisfaction.....	8
2.3 E-Commerce	10
2.4 Pra-Pemrosesan Data	10
2.5 Analisis Sentimen	12
2.6 Algoritma Naïve Bayes Classification.....	12
2.7 Evaluasi.....	13
2.8 Pemodelan topik.....	14
2.9 Latent Dirichlet Allocation (LDA)	14
2.10 RapidMiner Studio.....	15
2.11 WEKA.....	16
2.12 Penelitian Terdahulu	16
BAB III METODOLOGI PENELITIAN	21
3.1 Objek Penelitian.....	21
3.2 Jenis Data	21
3.3 Metode Pengumpulan Data.....	21
3.4 Flowchart Penelitian	23
3.5 Kerangka Penelitian	27

BAB IV PENGUMPULAN DAN PENGOLAHAN DATA	29
4.1 Deskripsi Data.....	29
4.2 Pra-Pemrosesan Data	30
4.3 Klasifikasi Data.....	37
4.4 Pemodelan Topik	43
BAB V ANALISA DAN PEMBAHASAN.....	48
5.1 Analisis Sentimen	48
5.2 Pemodelan Topik	53
5.3 Evaluasi Perbaikan.....	54
BAB VI KESIMPULAN DAN SARAN.....	57
6.1 Kesimpulan	57
6.2 Saran	58
DAFTAR PUSTAKA	59
LAMPIRAN.....	64



DAFTAR GAMBAR

Gambar 1 <i>Flowchart</i> Penelitian	23
Gambar 2 Kerangka Penelitian	27
Gambar 3 Tahap Pra-Pemrosesan Teks Menggunakan RapidMiner Studio.....	33
Gambar 4 Pra-Pemrosesan Dokumen pada Data	34
Gambar 5 Proses Klasifikasi Data berdasarkan Dimensi <i>E-Satisfaction</i>	37
Gambar 6 Proses Analisis Sentimen pada <i>Tweets E-Commerce</i> Shopee.....	39
Gambar 7 Pra-Pemrosesan Data Analisis Sentimen	39
Gambar 8 <i>Cross Validation</i>	40
Gambar 9 Klasifikasi Dimensi <i>E-Satisfaction</i> dengan Naïve Bayes dengan menggunakan WEKA	41
Gambar 10 Visualisasi PyLDAvis Dimensi <i>Convenience</i>	43
Gambar 11 Visualisasi PyLDAvis Dimensi <i>Site design</i>	44
Gambar 12 Visualisasi PyLDAvis Dimensi <i>Site Design</i>	45
Gambar 13 Visualisasi PyLDAvis Dimensi <i>Security</i>	46
Gambar 14 Visualisasi PyLDAvis Dimensi <i>Serviceability</i>	47
Gambar 15 Persentasi Klasifikasi Data <i>E-Commerce</i> Shopee Berdasarkan Dimensi	49
Gambar 16 <i>Confussion Matrix</i> Analisis Sentimen <i>E-Commerce</i> Shopee pada RapidMiner Studio.....	50
Gambar 17 <i>Confussion Matrix</i> Analisis Sentimen <i>E-Commerce</i> Shopee pada WEKA.....	50
Gambar 18 Visualisasi (a)Sentimen Ulasan Pengguna <i>E-Commerce</i> Shopee, (b) Data Sentimen Negatif, dan (c) Data Sentimen Positif	52

DAFTAR TABEL

Tabel 1 Perbandingan Persepsi Kinerja dan Harapan	7
Tabel 2 Penelitian Terdahulu	13
Tabel 3 Penelitian Terdahulu	18
Tabel 4 Hasil Pengumpulan Data.....	29
Tabel 5 Klasifikasi Data E-Commerce Shopee berdasarkan E-Satisfaction.....	31
Tabel 6 Klasifikasi Data E-Commerce Shopee berdasarkan Sentimen	32
Tabel 7 Pembagian Jenis Data berdasarkan Dimensi E-Satisfaction.....	33
Tabel 8 Struktur Data Tweet Sebelum dan Setelah Tahap Case Folding	35
Tabel 9 Struktur Data Tweet Sebelum dan Setelah Tahapan Tokenizing	35
Tabel 10 Struktur Data Tweet Sebelum dan Setelah Tahap Stopwords	36
Tabel 11 Struktur Data Tweet Sebelum dan Setelah Tahapan Stemming	36
Tabel 12 Pemodelan topik LDA Dimensi Convenience	43
Tabel 13 Pemodelan topik LDA Dimensi Site design	44
Tabel 14 Pemodelan topik LDA Dimensi Site Design	45
Tabel 15 Pemodelan topik LDA Dimensi Security.....	46
Tabel 16 Pemodelan topik LDA Dimensi Serviceability.....	47
Tabel 17 Nilai Kinerja Pengklasifikasian Naïve Bayes berdasarkan Dimensi <i>E-Satisfaction</i>	48
Tabel 18 Hasil Klasifikasi Naïve Bayes berdasarkan Dimensi <i>E-Satisfaction</i> dengan RapidMiner Studio	49

BAB I PENDAHULUAN

1.1 Latar Belakang

Roda perekonomian seringkali dijumpai peran industri dalam proses penjualan produk hingga jasa yang memerlukan beberapa proses sebelum sampai ke tangan konsumen. Salah satu proses yang sering memiliki kendala serta tanggapan dari konsumen adalah kepuasan dari masyarakat dalam menggunakan media sosial untuk menginterpretasikan keluhannya sehingga menjadi bahan evaluasi dari situs jual beli online tersebut.

Menurut lembaga riset pasar e-Marketer, populasi netter di dunia mencapai angka sebesar 5.3 miliar orang di tahun 2022 dan pada populasi netter di Indonesia mencapai angka 112.6 juta orang yang awalnya di peringkat ke-5 di tahun 2014 dan pada tahun 2022 meningkat di peringkat ke-4 yang menyentuh angka 204.7 juta (katadata.co.id). Hal ini menunjukkan bahwa antusias masyarakat Indonesia dalam menggunakan internet lebih meningkat dari 5 tahun sebelumnya. Tidak menutup kemungkinan ditahun-tahun berikutnya netter Indonesia akan meningkat secara drastis. Hal ini sejalan dengan perkembangan dari situs online mulai dari media sosial hingga *e-commerce*.

Perkembangan *e-commerce* di Indonesia sendiri telah mengikuti perkembangan internet sejak pertama kali masuk ke Indonesia pada tahun 1990-an. Saat ini aktivitas *e-commerce* di Indonesia telah merambah ke berbagai aktivitas bisnis. Dari skala industri kecil hingga industri besar. Berbelanja melalui *e-commerce* juga memudahkan pembeli yang sibuk dan tidak punya waktu untuk membeli barang yang mereka cari atau butuhkan, dan tentunya menghemat waktu dan tenaga.

Menurut Wong (2010) bahwa definisi dari *e-commerce* adalah pembelian, penjualan dan pemasaran barang serta jasa melalui sistem elektronik, seperti radio, televisi dan jaringan komputer atau internet, sedangkan menurut Widodo dan Avania (2023) mendefinisikan bahwa *e-commerce* merupakan arena

transaksi atau pertukaran informasi antara penjual dan pembeli di dunia maya. Siapapun yang mempunyai jaringan internet dapat berpartisipasi dalam kegiatan *e-commerce*.

Dengan banyaknya *e-commerce* di Indonesia yang berkembang, opini dan ulasan terkait *e-commerce* pun semakin banyak yang berangkat dari media cetak maupun elektronik. Dengan data bahwa netter di Indonesia yang semakin meningkat saat ini pastinya sering menggunakan media sosial dalam memberikan ulasan atau komentar mengenai masalah, khususnya dari suatu produk yang diperjual-belikan. Berbicara mengenai kepuasan konsumen yang diberikan oleh pengguna *e-commerce* menjadi penentu bahwa citra perusahaan yang diperoleh dari ulasan atau komentar pengguna melalui media sosial akan berpengaruh signifikan terhadap keputusan pembelian di *e-commerce* tersebut hal ini diperkuat oleh penelitian sebelumnya. Pada dasarnya, keputusan pembelian berangkat dari citra perusahaan yang diperoleh dari ulasan pengguna dan ulasan pengguna berasal dari tingkat kepuasan konsumen. Menurut Hanaysha (2018) terdapat dua faktor yang mempengaruhi keputusan pembelian yaitu informasi produk dan ulasan produk secara online. Maka dari itu, salah satu indikator untuk mengukur kepuasan konsumen dapat diperoleh dari ulasan pengguna. Menurut Zagoto dan Taryo (2022) pengaruh citra terhadap keputusan pembelian pada *e-commerce* Shopee memiliki nilai korelasi sebesar 0.799 berarti hubungan diantara kedua variabel tersebut tergolong kuat. Hal ini berarti dengan adanya ulasan yang diberikan oleh pengguna mengindikasikan tingkat dari kepuasan konsumen. Secara empiris, menurut Nurakhmawati, dkk (2022) Shopee memiliki nilai ulasan yang mencapai 4.7 dari total 5 bintang. Jika diperbandingkan dengan kompetitor yang sejenis ternyata memiliki nilai ulasan sebesar 4.8 dari total 5 bintang. Hal ini jika dipandang secara empiris berarti *e-commerce* Shopee masih tertinggal dari kompetitornya dari segi kepuasan konsumen maka dianggap perlu untuk mengangkat penelitian ini sebagai langkah perusahaan untuk menyaingi para kompetitornya.

Dengan ulasan yang diberikan oleh konsumen terkait produk yang dibeli di *e-commerce* tertentu pastinya selain langsung dari pelayanan pelanggan yang berfungsi untuk pelayanan dari pengaduan terkait kepuasan dari konsumen, fungsi dari sosial media adalah tempat atau wadah untuk memberikan suara mengenai komentar-komentar entah dari segi komentar yang positif, netral bahkan negatif sekalipun Pasal UU ITE Pasal 27 ayat (3) tentang pencamaran nama baik sebagai landasan hukum dalam memberikan komentar. Media sosial memiliki koneksi yang sangat luas sehingga para penegak hukum akan sulit untuk mengawasi hal tersebut. pengguna *ecommerce* Shopee menurut data dari similarweb per Agustus 2022 sebanyak 190,7 juta kunjungan dari Indonesia yang diperkuat dengan data pendapatan Shopee menunjukkan kuartil 1 tahun 2022 meningkat sebesar 64% dibandingkan pada kuartil 1 tahun 2021 dengan pendapatan mencapai US\$ 2,9 miliar (databoks.co.id).

Pada kasus seperti ini tidak dipungkiri bahwa dengan banyaknya pengguna *e-commerce* Shopee pastinya sering memberikan ulasan terkait kepuasan mereka di media sosial yang mereka gunakan. Twitter salah satu dari media sosial yang sering digunakan oleh masyarakat terkait komentar-komentar mereka dari produk yang mereka beli. Jumlah pengguna aktif harian mencapai 237,8 juta pengguna (dataindonesia.id). Dengan adanya sosial media khususnya Twitter, pihak perusahaan dapat melihat komentar yang bersifat membangun bahkan merugikan sehingga bisa menjadi bahan evaluasi dari perusahaan itu sendiri.

Dengan banyaknya penelitian terdahulu yang menyangkut dengan penelitian peneliti tentang kepuasan konsumen. Peneliti mengkaji aspek kepuasan konsumen dari pengguna *e-commerce* Shopee dari tweet yang mereka berikan berupa ulasan dan diolah dengan menggunakan 2 metode dari konsep penambangan data, yaitu Naïve Bayes Classification sebagai alat untuk mengelompokkan ulasan pengguna dan Latent Dirichlet Allocation sebagai alat untuk melihat topik yang sering dibicarakan oleh pengguna dari *e-commerce* Shopee. Dengan ini mengelompokkan ulasan dan melihat topik

dari tiap-tiap pengelompokan dari pengguna *e-commerce* Shopee masih belum terjawab.

Maka dari itu, ulasan dari pelayanan Shopee yang diberikan di sosial media Twitter sebagai tempat pengambilan data dianggap perlu untuk di olah secara terstruktur. Selanjutnya dilakukan klasifikasi sentimen positif dan negatif serta melihat ulasan apa saja yang diberikan paling banyak oleh pengguna Shopee itu.

1.2 Rumusan Masalah

Seperti yang telah diuraikan dari latar belakang, maka penelitian ini memiliki rumusan masalah sebagai berikut.

1. Bagaimana mengelompokkan kepuasan konsumen dengan menggunakan metode *sentimen analysis*?
2. Bagaimana menganalisis hasil pengelompokan kepuasan konsumen dengan menggunakan metode pemodelan topik Latent Dirichlet Allocation (LDA)?
3. Bagaimana rekomendasi perbaikan yang diberikan kepada *e-commerce* Shopee?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini, yaitu untuk:

1. Mengetahui dan mengelompokkan dari ulasan kepuasan konsumen *e-commerce* Shopee.
2. Menganalisis hasil pengelompokan kepuasan konsumen *e-commerce* Shopee dengan metode pemodelan topik Latent Dirichlet Allocation (LDA)
3. Memberikan rekomendasi perbaikan terhadap *e-commerce* Shopee.

1.4 Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian ini ditujukan bagi beberapa pihak sebagai berikut.

a. Bagi Penulis

Penulis dapat menambah wawasan dan kemampuan dalam mengaplikasikan ilmu-ilmu Teknik Industri dalam memecahkan permasalahan khususnya pada penambangan data.

b. Bagi Pihak Perusahaan

Sebagai bentuk evaluasi yang ditampilkan oleh penulis agar pihak perusahaan dapat meningkatkan tingkat kepuasan penggunaannya.

1.5 Batasan Masalah

Agar permasalahan tidak meluas, maka diambil batasan-batasan masalah yang akan diteliti adalah sebagai berikut.

1. Data yang dianalisis adalah data dari media sosial Twitter berbahasa Indonesia.
2. *Tweet* yang dianalisis sentimen hanya *tweet* tentang kepuasan konsumen dalam hal dimensi *e-satisfaction e-commerce* Shopee.
3. Data yang diperoleh mulai bulan Januari hingga Maret 2023.

1.6 Sistematika Penulisan

Sistematika yang akan digunakan dalam tugas akhir ini dibagi menjadi beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab pendahuluan membahas mengenai latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan Tugas Akhir mengenai analisis sentimen ulasan pengguna *e-commerce* Shopee menggunakan Naïve Bayes Classification dan Latent Dirichlet Allocation.

BAB I TINJAUAN PUSTAKA

Pada bab ini menjelaskan mengenai keseluruhan teori-teori yang digunakan yang berkaitan tentang penelitian ini yang menggunakan metode Naïve Bayes Classification dan Latent Dirichlet Allocation, yaitu mulai dari kepuasan konsumen, *e-commerce*, Shopee, Twitter, analisis sentimen, penambahan data dan Latent Dirichlet Allocation.

BAB III METODOLOGI PENELITIAN

Pada bab ini menyajikan garis besar penyelesaian masalah tugas akhir yang diawali dengan objek penelitian, jenis data yang digunakan, metode pengumpulan data, dan prosedur penelitian dalam bentuk *flowchart*.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menampilkan hasil dan pembahasan dari pengolahan data yang telah dilakukan sehingga dapat memenuhi dari rumusan masalah yang diusung.

BAB V PENUTUP

Pada bab ini berisi tentang kesimpulan dari penelitian ini beserta saran yang dapat diajukan guna pengembangan dari analisis sentimen Naïve Bayes Classification dan pemodelan topik Latent Dirichlet Allocation dan akan dikembangkan oleh peneliti kedepannya.

BAB II

TINJAUAN PUSTAKA

2.1 Kepuasan Konsumen

Menurut Tjiptono dan Chandra (2016) kepuasan konsumen merupakan ukuran kinerja 'produk total' sebuah organisasi dibandingkan serangkaian keperluan pelanggan. Kepuasan konsumen tidak lagi konsep yang mutlak, melainkan relatif atau bergantung dengan ekspektasi konsumen itu sendiri sedangkan menurut Hasan (2014) kepuasan dan ketidakpuasan merupakan bentuk evaluasi setelah konsumen melakukan transaksi, dimana solusi yang dipilih hampir sama ataupun lebih dari ekspektasi konsumen, dan ketidakpuasan dapat terjadi jika hasil yang tidak sesuai dari ekspektasi itu sendiri. Adapun perbandingan persepsi kinerja dan harapan terhadap kepuasan pelanggan menurut Hasan (2014):

Tabel 1 Perbandingan Persepsi Kinerja dan Harapan

Kinerja > Harapan	Sangat Puas	Persepsi Positif
Kinerja < Harapan	Tidak Puas	Persepsi Negatif

Sumber : Hasan (2014)

Menurut Putri dan Astuti (2017) bahwa Salah satu cara untuk menentukan tingkat kepuasan konsumen adalah dengan melihat nilai pelanggan. Nilai pelanggan dapat didefinisikan sebagai besarnya perbedaan yang dirasakan oleh konsumen terhadap produk atau jasa yang diberikan oleh perusahaan dibandingkan dengan biaya yang harus dibayar oleh konsumen untuk mendapatkan produk atau jasa tersebut. Dengan kata lain, nilai pelanggan dapat didefinisikan sebagai jumlah manfaat dan kualitas yang dirasakan oleh konsumen dibandingkan dengan pengorbanan yang telah dilakukan oleh konsumen untuk mendapatkan produk atau jasa tersebut. Selain itu, faktor-faktor yang memengaruhi keputusan pembelian juga akan memengaruhi seberapa puas pelanggan dengan produk, layanan, dan nilai yang ditawarkan. kepuasan yang baik cenderung menyebabkan pelanggan tertarik untuk membeli produk lain.

Menurut Kotler (2002) ciri-ciri konsumen yang merasa puas adalah sebagai berikut:

- a. Loyal terhadap produk, konsumen yang memiliki rasa puas lebih loyal dimana mereka akan menimbulkan minat beli ulang dari produsen yang sama.
- b. Adanya komunikasi dari mulut ke mulut yang bersifat positif. Sifat komunikatif konsumen yang bersifat positif berupa informasi dari mulut ke mulut akan mengakibatkan rekomendasi kepada calon konsumen lain dan mengatakan hal-hal lain yang positif mengenai produk atau jasa perusahaan.
- c. Perusahaan menjadi prioritas utama. Dengan melihat nilai pelayanan yang diberikan oleh konsumen, konsumen akan mendahulukan perusahaan tersebut sebagai opsi pertama ketika ingin memberi produk yang lain.

Menurut Lupiyoadi (2013) adapun faktor-faktor yang mempengaruhi kepuasan konsumen:

- a. Kualitas produk adalah tiap konsumen akan merasa puas jika hasil produk yang mereka gunakan berkualitas.
- b. Kualitas pelayanan. Faktor ini merupakan konsumen dapat merasakan kepuasan jika pelayanan yang mereka dapatkan baik ataupun dapat memenuhi harapannya.
- c. Emosi. Pada faktor ini yang dimaksudkan adalah konsumen akan merasakan perasaan senang serta kepercayaan bahwa orang lain akan kagum pada mereka jika produk dengan merek tertentu digunakan yang mempunyai kepuasan yang lebih bagus.
- d. Harga. Faktor ini adalah value yang lebih baik jika produk memiliki kualitas sama akan tetapi harga yang sedikit lebih murah.
- e. Biaya. Berbeda pada faktor harga, biaya yang dimaksudkan disini adalah para konsumen tidak perlu memberi biaya yang lebih atau mengulur waktu agar produk tersebut dapat dimiliki.

2.2 E-Satisfaction

Menurut Ranjbarian, dkk (2012) menjelaskan bahwa *e-satisfaction* atau kepuasan pelanggan online adalah hasil dari persepsi konsumen terhadap kenyamanan online, perdagangan /cara transaksi, desain situs, keamanan dan

pelayanan. Menurut Kim, dkk (2009) mendefinisikan *e-satisfaction* sebagai akumulasi dari kepuasan yang konsumen dapatkan pada saat pembelian barang atau jasa dari waktu ke waktu pada sebuah situs online.

Menurut Ranjbarian, dkk (2012) adapun lima dimensi utama yang mempengaruhi *e-satisfaction* ini sebagai berikut:

- a. *Convenience* menjelaskan tentang belanja secara online dapat mengefisienkan waktu serta tenaga dengan konsumen mudah mendapatkan penjual, produk dan penawaran yang diinginkan. Faktor-faktor yang mempengaruhi adanya *convenience* adalah berlangsungnya proses jual beli, nyamannya saat berbelanja serta mudahnya mengeksplor hal-hal yang diinginkan konsumen secara online.
- b. *Site design* menjelaskan tentang informasi yang lebih luas dan berkualitas tinggi yang disediakan secara online guna mendapatkan keputusan pembelian yang lebih baik dan tingkat kepuasan elektronik yang lebih tinggi. Adapun Faktor-faktor yang mempengaruhi adanya *site design* yaitu kuantitas penawaran, variasi dari penawaran, kualitas dan jumlah informasi yang ada dan biaya yang lebih minim dalam pencarian.
- c. *Site design* menjelaskan tentang desain situs web yang baik adalah tentang mengorganisir konten yang baik dan kemudahan dalam pencarian dalam hal tampilan layer yang sistematis, sederhana dan presentasi yang singkat. Faktor-faktor yang mempengaruhi adanya *site design* adalah struktur navigasi, tampilan yang saling terintegrasi dalam hal desain web, presentasi yang gesit, dan konten yang relevan serta aktualnya informasi yang disediakan.
- d. *Security* menjelaskan tentang transaksi online yang berprinsip keamanan sebagai indikator utama saat memutuskan pembelian secara online. Faktor-faktor yang mempengaruhi adanya *security* adalah data pribadi dan keamanan keuangan.
- e. *Serviceability* menjelaskan tentang komentar secara umum mengenai kompetitifnya harga produk, tersedianya stok produk, kondisi produk dan pengiriman yang sesuai estimasi atau tepat waktu. Faktor-faktor yang mempengaruhi adanya *serviceability* adalah pengaduan yang direspon,

dukungan pelanggan, proses pengemasan, waktu pengiriman, layanan pos, harga yang kompetitif, kebijakan dari pengembalian produk atau dana, email konfirmasi, proses pembelian dan umpan balik situs.

2.3 E-Commerce

Menurut Insana & Johan (2020) bahwa *e-commerce* merupakan salah satu keunggulan dari internet, hingga akhirnya di era sekarang ini (yang disebut dengan era digital) *e-commerce* memiliki banyak sebutan yaitu *internet commerce*, *e-com*, *e-commerce* atau *immerce* yang berarti membeli atau menjual secara elektronik, dan kegiatan ini dilakukan pada jaringan internet.

Menurut (Ginting dkk., 2022) *e-commerce* adalah distribusi, pembelian, penjualan, dan pemasaran barang dan jasa melalui media elektronik seperti internet, televisi, atau jaringan internet lainnya. Pertukaran data elektronik, transfer dana elektronik, dan sebagainya adalah beberapa aspek dari *e-commerce*. *E-commerce* dapat melakukan banyak hal, seperti pemasaran, penjualan, pembelian, dan distribusi produk dengan menggunakan sistem elektronik seperti internet atau jaringan komputer yang lain. Pembayaran untuk pembelian dapat dilakukan melalui transfer bank, dan produk dapat dikirim dengan menggunakan jasa pengiriman seperti JNE, JNT, dan lainnya. Saat ini hampir semua barang dapat dibeli melalui platform ini, sehingga *e-commerce* bisa berkembang pesat.

2.4 Pra-Pemrosesan Data

Menurut (Muttaqin & Mukaharil Bachtiar, 2016) bahwa pra-pemrosesan adalah proses untuk menyiapkan data yang masih mentah yang selanjutnya akan di proses oleh sistem. Secara umum, pada proses ini data diproses dengan mengeliminasi data yang tidak berkaitan ataupun mengubah data menjadi bentuk yang lebih mudah diproses daripada sebelumnya. Dikarena data tidak boleh diterapkan secara membabi buta selama proses pra-pemrosesan, kumpulan data harus diperiksa dan dikarakterisasi sebelum memulai proses ini. Dengan eksplorasi dan karakterisasi data, kita dapat melakukan Teknik tokenizing, cleaning data,

hingga stemming. Adapun tahapan yang dilakukan pada saat pra-pemrosesan data adalah sebagai berikut:

2.4.1 Tokenizing

Tokenizing merupakan proses pemisahan atau memotong data baik berupa frasa, klausa atau kalimat menjadi kata perkata atau dengan kata lain potongan kata tunggal. Pada tahap ini kalimat ataupun paragraf yang akan dilakukan dengan Teknik ini akan dipisahkan seperti contoh “Saya adalah mahasiswa Teknik Industri” menjadi: “Saya”, “adalah”, “mahasiswa”, “Teknik”, “Industri” (Muttaqin & Mukaharil Bachtiar, 2016).

2.4.2 Cleaning Data

Pada proses pra-pemrosesan data, adapun teknik yang harus dilakukan adalah pembersihan data. Pembersihan data ini menangani data yang kotor karena kumpulan data yang kotor dapat menyebabkan masalah untuk eksplorasi dan analisis data. Teknik ini telah dikembangkan untuk membersihkan data dengan mengisi nilai yang hilang (*input* data), menghaluskan data yang bising, mengidentifikasi dan/atau menghapus *outlier* dan menyelesaikan data yang inkonsisten. Data *cleaning* ini berguna untuk mengatasi data atau nilai yang hilang, *noise* dan data yang tidak konsisten (Sulastri & Gufroni, 2017).

2.4.3 Stemming

Stemming merupakan tahap mencari akar kata dari tiap kata hasil dari filtering. Pada tahap ini dilakukan proses pengambilan berbagai bentukan kata ke dalam suatu representasi yang sama. Tahap ini kebanyakan dipakai teks berbahasa Inggris dan lebih sulit diterapkan pada teks berbahasa Indonesia. Hal ini dikarenakan bahasa Indonesia tidak memiliki rumus bentuk baku yang permanen. Contoh tahapan stemming ini seperti “menyukai” menjadi “suka”, “melihat” menjadi “lihat” (Muttaqin & Mukaharil Bachtiar, 2016)

2.5 Analisis Sentimen

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data yang dituangkan dalam teks secara otomatis guna informasi sentimen yang terkandung dalam suatu kalimat dapat diperoleh. Analisis sentimen dilakukan guna melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung beropini negatif atau positif (Rozi, dkk., 2013).

Analisis sentimen adalah salah satu tugas klasifikasi dalam bentuk teks, yang bertujuan untuk mendapatkan informasi yang subjektif berdasarkan kalimat yang ada, apakah sentimen positif atau negatif terdapat pada kalimat tersebut atau tidak. Analisis sentimen akan mengekstra informasi kontekstual dalam dari sebuah bentuk teks dan kemudian menentukan sentimen teks dengan memanfaatkan algoritma tertentu dan perhitungan seperti pembelajaran mesin. Analisis sentimen adalah bidang dimana kita dapat menggunakan berbagai metode untuk menyelesaikan tugas (Diekson dkk., 2023).

2.6 Algoritma Naïve Bayes Classification

Klasifikasi merupakan proses guna mendapatkan kategori yang tidak diketahui sebelumnya dari suatu objek. menentukan. Naïve Bayes Classification (NBC) merupakan salah satu algoritma metode klasifikasi pembelajaran yang diawasi dan bekerja pada algoritma Bayes. Algoritma ini merupakan algoritma yang efektif untuk melakukan permodelan prediktif dan dapat digunakan untuk menyelesaikan klasifikasi prediktif. Salah satu keunggulan dari algoritma ini adalah dapat bekerja dengan baik dengan data yang berukuran besar. Selain itu, algoritma ini dijuluki sebagai predictor yang cepat yang dapat membuat perkiraan berdasarkan kemungkinan suatu objek. Prinsip umum pada algoritma ini, yaitu semua prediktor independent satu sama lain. Artinya bahwa setiap pasangan fitur untuk diklasifikasikan, tidak bergantung sama lain. Oleh karena itu, algoritma klasifikasi ini mengasumsikan bahwa tidak ada hubungan antara fitur tertentu dengan fitur lain di kelasnya (Amin dkk., 2023).

2.7 Evaluasi

Menurut Syakuro (2017) evaluasi performansi dilakukan guna melakukan pengujian hasil dari klasifikasi dengan mengukur nilai performansi dari sistem yang telah dibuat oleh perangkat lunak. Parameter pengujian yang digunakan untuk evaluasi yaitu akurasi yang perhitungannya dari tabel *confusion matrix*. Menurut (Narayan & Sathiyamoorthy, 2022) bahwa *confusion matrix* digunakan untuk menunjukkan jumlah total yang salah dan benar dari prediksi atau hasil dengan model apapun dibandingkan dengan klasifikasi pasti dari kumpulan data. Contoh kelas yang diprediksi ditampilkan pada bagian baris *confusion matrix* serta contoh kelas yang sebenarnya ditampilkan pada bagian begitu pun sebaliknya. Dalam tabel 2 ini terdapat istilah *true positif* (TP), *false positif* (FP), *false negatif* (FN) dan *true negatif* (TN). *True positif* (TP) adalah nilai aktualnya dan model yang diprediksi bernilai positif, *false positif* (FP) adalah Prediksinya bernilai positif, tetapi hasilnya salah, *false negatif* (FN) adalah prediksinya bernilai negatif dan hasilnya juga salah serta *true negatif* (TN) adalah nilai aktualnya dan model diprediksi bernilai negatif. Adapun tabel dibawah ini menunjukkan bentuk dari *confusion matrix*.

Tabel 2 Penelitian Terdahulu

Actual Class	Predicted Class	
	C1	C2
C1	<i>True Positif</i>	<i>False Positif</i>
C2	<i>False Negatif</i>	<i>True Negatif</i>

Di dalam *confusion matrix* terdapat. evaluasi performansi, yaitu *precision*, *recall* dan *accuracy*. *Precision* merupakan tingkat keakuratan antara informasi yang diminta oleh pengguna dan jawaban yang disediakan oleh sistem. *Recall* merupakan tingkat keberhasilan sistem dalam mencari informasi kembali. *Accuracy* merupakan tingkat kedekatan antara nilai prediksi dan aktual dari suatu nilai yang dihasilkan (Widodo & Aji, 2023).

2.8 Pemodelan topik

Pemodelan topik merupakan metode *clustering* yang termasuk dalam unsupervised learning. Dalam unsupervised learning tidak ada label untuk suatu objek. Pemodelan topik adalah suatu metode yang memiliki pendekatan dapat menganalisis dokumen yang telah dikumpulkan dalam bentuk teks dan dikelompokkan di tiap-tiap topik. Pendekatan tersebut masuk dalam pendekatan clustering dalam studi perangkat lunak (Nurlayli & Nasichuddin, 2019).

Terdapat 3 tipe clustering yaitu *hard clustering*, *hierarchical clustering*, dan *soft/fuzzy clustering*. *Hard clustering* adalah metode pengelompokan di mana setiap objek ditugaskan secara jelas ke dalam satu kluster, *hierarchical clustering* adalah metode pengelompokan hierarkis menghasilkan representasi data berbasis pohon yang dinamakan dendrogram, dan *soft/fuzzy clustering* teknik yang memberikan eksplisit yang memberikan representasi ketidakpastian sesuai dengan tingkatannya. Dengan melihat 3 tipe *clustering* tersebut, *soft/fuzzy clustering* merupakan tipe yang cocok pada pemodelan topik ini karena setiap objeknya dapat dimiliki lebih dari satu kluster dengan tingkatan yang ada (Doig, 2015).

2.9 Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) adalah suatu metode yang bertujuan untuk menganalisis dokumen yang memiliki kuantitas yang sangat besar. Metode ini pun dapat melakukan hal-hal seperti meringkas, melakukan pengelompokan, menghubungkan hingga menjalankan data. Dalam LDA ini seringkali bertujuan untuk melakukan analisis tren topik yang akan dihasilkan dan divisualisasikan sehingga lebih informatif dan mudah dipahami oleh pengguna. LDA telah banyak dikembangkan dalam penelitian terkait menganalisis topik dari sebuah teks maupun dokumen. Penelitian terkait dengan menggunakan metode LDA banyak cara untuk menganalisis tren sebuah topik dengan berbagai sumber, yaitu dapat mengakses dari situs-situs *website* dan media sosial. Model berbasis LDA dapat dilihat sebagai model generatif. Model-model ini mengasumsikan proses generatif korpus dan memperoleh estimasi berdasarkan metode inferensi. LDA dapat

digunakan secara efisien untuk menemukan topik laten dari koleksi besar dokumen dan kelompok kata keluaran termasuk aspek. Aspek-aspek tersebut secara alami dikumpulkan untuk topik terkait. Adapun aliran penting penelitian LDA dalam penggalian opini menyangkut penggalian aspek dan kata-kata sentimen secara simultan (Guo dkk., 2022).

Latent Dirichlet Allocation (LDA) yang diperkenalkan oleh Blei dan Jordan, dijelaskan sebagai model probabilistik generatif untuk mencari struktur semantik dari kumpulan korpus yang berdasarkan *hierarchical Bayesian analysis*. LDA merupakan kumpulan dokumen sebagai topik campuran yang berisi kata-kata dengan probabilitas tertentu. Adapun alur kerja dari LDA atau prosedur cara kerja LDA adalah sebagai berikut:

- (1) Inisialisasi beberapa parameter, termasuk jumlah dokumen, topik, dan iterasi. Dalam LDA, parameter yang paling penting adalah jumlah topik k .
- (2) Menetapkan kata untuk topik tertentu secara acak sesuai dengan distribusi *dirichlect*.
- (3) Mengulangi masing-masing alur proses untuk semua kata dalam korpus. (Nurlayli & Nasichuddin, 2019).

2.10 RapidMiner Studio

Persiapan data, perangkat lunak, *deep learning*, penambahan teks, analisis prediktif, dan analisis bisnis semuanya didukung oleh RapidMiner Studio, yang menawarkan lingkungan terpadu. RapidMiner Studio mendukung semua proses pembelajaran mesin, termasuk persiapan data, visualisasi hasil, validasi model, dan pengoptimalan. Software ini biasanya digunakan untuk bisnis, aplikasi komersial, penelitian, pendidikan, pelatihan, pembuatan *prototype* cepat, dan pengembangan aplikasi. Kelebihan dari software ini adalah dataset dapat disimpan dalam memori selama mungkin oleh RapidMiner Studio. Oleh karena itu, perangkat lunak ini tidak akan menghapus hasil proses yang lama jika masih ada memori yang tersisa. RapidMiner Studio dipuji karena mendukung fungsionalitas yang luas dan mendalam, dan dengan hal itu sangat dekat dengan pasar bisnis (Pynam, dkk., 2018)

2.11 WEKA

WEKA mengandung algoritme pembelajaran mesin yang membantu metode pembelajaran mesin pada kumpulan data baru. Perangkat lunak ini sangat membantu dalam seluruh proses penggalian data eksperimental, seperti menyiapkan data input, mengevaluasi skema pembelajaran secara statistik, dan menampilkan data input dan hasil pembelajaran. Untuk mencapai hal ini, berbagai metode prapemrosesan dan berbagai algoritme untuk mempelajari berbagai jenis konsep digunakan. Melalui antarmuka yang sama, kumpulan alat yang beragam dan luas ini dapat digunakan. Ini memungkinkan pengguna membandingkan metode yang berbeda dan menemukan metode yang paling sesuai untuk masalah yang dihadapi. Salah satu keunggulan dari machine learning ini adalah modularitasnya. Hal ini memungkinkan algoritma untuk digabungkan dengan berbagai cara. Sebagai contoh, seseorang dapat menggabungkan beberapa algoritma untuk klasifikasi, *clustering* dan sebagainya (Frank, dkk., 2010).

2.12 Penelitian Terdahulu

Penelitian Sutherland, dkk (2020) memiliki tujuan yaitu untuk membandingkan topik diantara berbagai jenis akomodasi dan lokasi sehingga dapat dibandingkan dengan pendekatan deduktif tradisional dengan LDA. Penelitian ini mendapatkan lebih banyak dimensi daripada yang ditemukan oleh penelitian tradisional karena dengan pendekatan statistik dan induktif, topik dapat berdasarkan perspektif konsumen. Selain itu, hasil menunjukkan bahwa topik-topik yang berbeda memiliki signifikansi yang berbeda tergantung pada karakteristik akomodasi, yang menunjukkan topik mana yang lebih penting berdasarkan urbanisasi kota atau jenis akomodasi. Hal berbeda yang ada pada penelitian tugas akhir ini adalah penelitian penulis mengkhususkan bukan pada ulasan produk yang dijual oleh *seller*, akan tetapi ulasan dari *e-commerce* yang menyediakan jasa ataupun produk yang ada didalamnya.

Penelitian Wang, dkk (2018) memiliki tujuan yaitu untuk menganalisis topik ulasan online untuk dua produk yang saling kompetitif dan mendapatkan

pemahaman tentang daya saing pada masing-masing produknya. Pada penelitian ini yang dilakukan dengan pendekatan Latent Dirichlet Allocation (LDA), menemukan bahwa desainer produk dapat lebih mudah memahami aspek-aspek produk. Selain itu, bisnis yang berpartisipasi dalam pasar produk kompetitif dapat menemukan fitur khusus, iklan kompetitif, serta keunggulan dan kekurangan produk mereka dibandingkan dengan pesaing mereka. Hal yang berbeda pada penelitian tugas akhir ini adalah penelitian penulis mengelompokkan ulasan positif dan negatif dari kinerja *e-commerce* Shopee berbeda dengan penelitian Wang, dkk (2018) yang membandingkan kedua produk yang bersifat kompetitif.

Penelitian Arianto dan Anuraga (2020) memiliki tujuan yaitu untuk penggunaan pendekatan Latent Dirichlet Allocation (LDA) untuk mengelompokkan *tweet* masyarakat yang berkaitan dengan layanan dari ruangguru. Dengan melihat hasil penelitiannya bahwa dengan metode permodelan topik LDA pada aplikasi ruangguru dapat mengelompokkan data Twitter menjadi 28 topik, dengan diskon ruangguru sebagai topik yang paling populer. Hal yang berbeda pada penelitian tugas akhir penulis adalah penulis mengelompokkan kepuasan konsumen *e-commerce* Shopee dan melihat topik yang terbanyak diberikan oleh pengguna dan hal ini berbeda dengan penelitian Bagus Wicaksono Arianto dan Gangga Anuraga (2020) dari segi objek penelitian.

Penelitian Masrury, dkk (2019) memiliki tujuan yaitu untuk mengetahui dimensi yang terbentuk berdasarkan kualitas layanan pada aplikasi *mobile* penyedia agen travel *online* yang berasal dari ulasan pengguna di Google *PlayStore*. Dengan menggunakan Latent Dirichlet Allocation (LDA) mendapatkan hasil bahwa terdapat sentimen positif dan negatif pada masing-masing agen *travel online* terhadap dimensi *MappSql (mobile application service quality)*. *Topic modelling* digunakan untuk menemukan kata dominan yang berhubungan dengan setiap dimensi kualitas layanan aplikasi *mobile*. Hal yang berbeda pada penelitian tugas akhir penulis adalah penulis mengelompokkan kepuasan konsumen *e-commerce* Shopee dengan menggunakan dimensi *e-satisfaction* dan melihat topik yang terbanyak diberikan oleh pengguna dan hal ini pun berbeda dengan penelitian Masrury, dkk (2019) dari segi objek penelitian

Penelitian Kabiru dan Sari (2019) memiliki tujuan yaitu untuk mengetahui klasifikasi sentimen yang dibicarakan oleh pengguna sosial media. Penelitian tersebut menggunakan metode Latent Dirichlet Allocation (LDA) dan *Naïve Bayes*. Hasil penelitian menunjukkan bahwa dengan adanya analisis sentimen dan pemodelan topik pada data komentar di sosial media Twitter Shopee, adapun topik yang diungkapkan oleh pengguna ini adalah sentimen positif berupa acara yang bagus dan menarik serta sentimen negatif yang diindikasikan melalui kekecewaan pengguna. Pada penelitian tugas akhir ini memiliki kesamaan dengan penelitian Kabiru dan Sari (2019) dari segi pengolompokkan ulasan konsumen *e-commerce* Shopee, akan tetapi penelitian penulis melihat topik yang terbanyak yang diberikan oleh konsumen pada kelompok komentar positif dan negatif. Selain itu, tempat pengambilan data dari penelitian tugas akhir ini berbeda, yakni pada sosial media Twitter.

Tabel 3 Penelitian Terdahulu

No	Nama Peneliti dan Tahun	Tujuan Penelitian	Hasil Penelitian
1.	Sutherland, dkk (2020)	Untuk membandingkan topik antara berbagai jenis akomodasi dan lokasi, pemahaman yang lebih baik tentang kepuasan pelanggan sehingga dapat dibandingkan dengan pendekatan deduktif tradisional dengan memanfaatkan LDA.	Penelitian ini mengidentifikasi lebih banyak dimensi daripada yang telah diidentifikasi oleh penelitian tradisional karena topik topik tersebut ditentukan secara statistik dan induktif berdasarkan perspektif konsumen. Hasil juga menunjukkan bahwa pentingnya topik berbeda berdasarkan karakteristik akomodasi, yang menunjukkan topik mana yang lebih penting berdasarkan urbanisasi kota atau tipe akomodasi.
2.	Wang, dkk (2018)	Untuk menganalisis topik ulasan online untuk dua produk yang bersaing dan memahami daya saing dari	Penelitian tersebut mengusulkan pendekatan baru untuk menganalisis keuntungan dan kerugian kompetitif dari dua

	<p>tiap-tiap produk yang bersumber dari komentar konsumen.</p>	<p>produk yang bersaing dari ulasan produk online. Hasil dari penelitian memungkinkan desainer produk untuk memahami aspek-aspek produk dengan lebih mudah. Selain itu, perusahaan dalam kompetisi pasar produk mampu mengidentifikasi aspek unik, iklan kompetitif, serta keuntungan dan kelemahan produk dengan pesaing.</p>
3. Arianto dan Anuraga (2020)	<p>Penelitian ini bertujuan untuk mengelompokkan <i>tweet</i> masyarakat mengenai kepuasan konsumen ruangguru dengan <i>topic modelling</i> LDA.</p>	<p>Metode klasifikasi dari <i>topic modelling</i> LDA pada aplikasi ruangguru mengelompokkan data Twitter menjadi 28 buah topik dengan topik yang sering diperbincangkan adalah diskon ruangguru</p>
4. Masrury, dkk (2019)	<p>Untuk mengetahui dimensi kualitas layanan pada aplikasi <i>mobile</i> penyedia agen travel <i>online</i> yang bersumber dari <i>review</i> konsumen di Google <i>Play Store</i></p>	<p>Hasil penelitian menunjukkan bahwa terdapat sentimen positif dan negatif pada masing-masing agen <i>travel online</i> terhadap dimensi MappSql (<i>mobile application service quality</i>). <i>Topic modelling</i> digunakan untuk menemukan kata dominan yang berhubungan dengan setiap dimensi kualitas layanan aplikasi <i>mobile</i>.</p>
5. Kabiru dan Sari (2019)	<p>Penulis menetapkan tujuan penelitian guna mengetahui klasifikasi sentimen yang diutarakan oleh pengguna sosial media.</p>	<p>Dengan melihat analisis sentimen dan pemodelan topik pada data komentar Instagram Shopee, adapun topik yang diungkapkan oleh pengguna ini adalah sentimen positif berupa event yang bagus, sentimen negatif berupa kekecewaan.</p>