

**ANALISIS SENTIMEN *ONLINE CUSTOMER REVIEW* PADA TOKO
SMARTPHONE DAERAH ISTIMEWA YOGYAKARTA DI *E-MARKETPLACE*
SHOPEE MENGGUNAKAN *LEXICON BASED* DAN *WORD CLOUD***

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-1
Pada Jurusan Teknik Industri Fakultas Teknologi Industri**



Nama : Fildzah Adlina

No. Mahasiswa : 18522273

**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA**

2022

SURAT PERNYATAAN KEASLIAN

Dengan ini saya mengakui bahwa karya ini adalah hasil kerja saya sendiri kecuali ringkasan dan kutipan setiap satunya telah saya jelaskan sumbernya. Jika kemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah dalam karya tulis dan hak intelektual maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Islam Indonesia

Yogyakarta, 12 Agustus 2022



Fildzah Adlina

NIM. 18522273

الجامعة الإسلامية
الاستدراك

LEMBAR PENGESAHAN PEMBIMBING

**ANALISIS SENTIMEN *ONLINE CUSTOMER REVIEW* PADA TOKO
SMARTPHONE DAERAH ISTIMEWA YOGYAKARTA DI *E-MARKETPLACE*
SHOPEE MENGGUNAKAN *LEXICON BASED* DAN *WORD CLOUD***

TUGAS AKHIR



Disusun Oleh

Nama : Fildzah Adlina
No. Mahasiswa : 18522273
Fakultas/Jurusan : FTI/Teknik Industri

Yogyakarta, 12 Agustus 2022

Mengetahui,

Dosen Pembimbing Tugas Akhir

Annisa Uswatun Khasanah, S.T., M.B.A., M.Sc.

LEMBAR PENGESAHAN PENGUJI

ANALISIS SENTIMEN *ONLINE CUSTOMER REVIEW* PADA TOKO *SMARTPHONE* DAERAH ISTIMEWA YOGYAKARTA DI *E-MARKETPLACE* SHOPEE MENGGUNAKAN *LEXICON BASED* DAN *WORD CLOUD*

Disusun Oleh

Nama : Fildzah Adlina

No. Mahasiswa : 18522273

Fakultas/Jurusan : FTI/Teknik Industri

Telah dipertahankan di depan sidang pengujian sebagai satu syarat untuk
memperoleh gelar Sarjana Strata-1 Teknik Industri

Yogyakarta, 24 Oktober 2022

Tim Penguji

Annisa Uswatun Khasanah, S.T., M.B.A., M.Sc.

Ketua

Wahyudhi Sutrisno, S.T., M.M., M.T

Anggota I

Sri Indrawati, S.T., M.Eng.

Anggota 2

Mengetahui

Ka. Prodi Teknik Industri

Fakultas Teknologi Industri

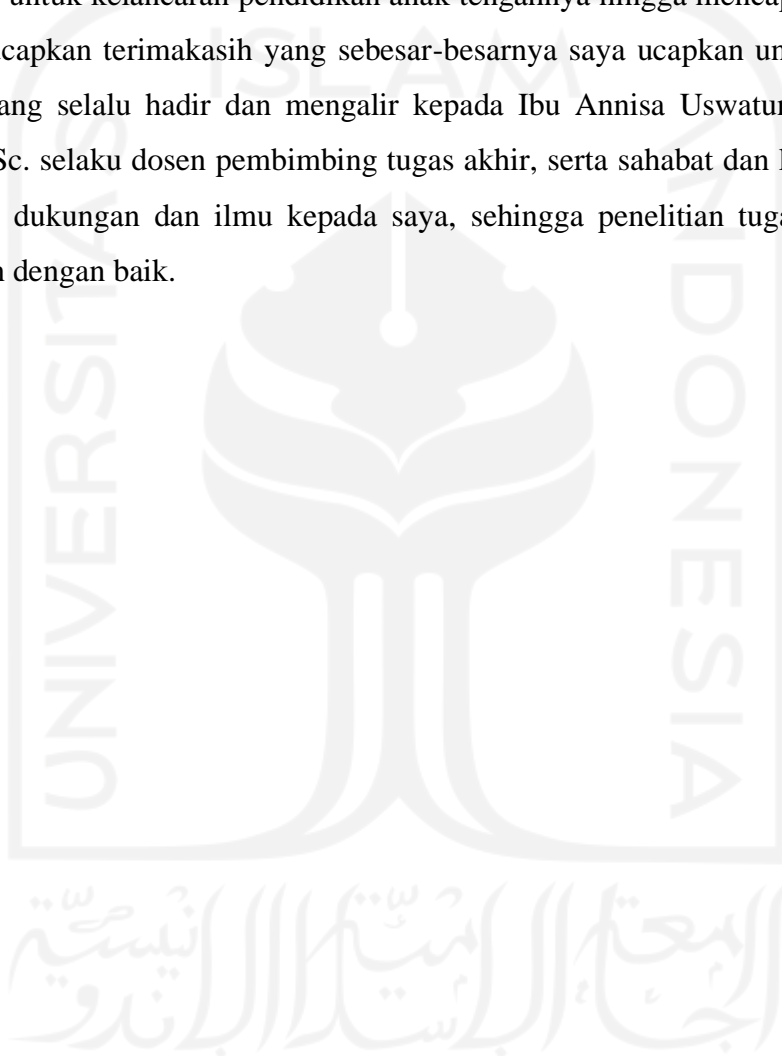
Universitas Islam Indonesia

Ir. Muhammad Ridwan And Purnomo, S.T., M.Sc., Ph.D., IPM.

HALAMAN PERSEMBAHAN

Alhamdulillahirrobbil'alamiin,

Tugas Akhir ini saya persembahkan kepada kedua orang tua tercinta, Bapak Fauzi dan Ibu Lisna Mariati, serta kakak dan adik saya, yang selalu memberikan doa dengan tulus dan tak pernah henti untuk kelancaran pendidikan anak tengahnya hingga mencapai tujuan saat ini. Serta saya ucapkan terimakasih yang sebesar-besarnya saya ucapkan untuk segala bentuk dukungan yang selalu hadir dan mengalir kepada Ibu Annisa Uswatun Khasanah, S.T., M.B.A., M.Sc. selaku dosen pembimbing tugas akhir, serta sahabat dan kerabat yang telah memberikan dukungan dan ilmu kepada saya, sehingga penelitian tugas akhir ini dapat terselesaikan dengan baik.



HALAMAN MOTTO

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya.”

(QS Al-Baqarah: 286)



KATA PENGANTAR

**Bismillahirrohmanirrohiim,
Assalamualaikum Warrahmatullahi Wabarakatuh,
Asyhadu Alla Ilahailallah Wa Asyhadu Anna Muhammadarrasulullah Allahuma
Shalli'ala Muhammad Wa'ala Alihi Washobihi Wasalam.**
Alhamdulillahirrobbil'alamiin, Segala puji dan syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya serta tak lupa pula shalawat dan salam kepada Nabi Muhammad SAW beserta keluarga, sahabat, dan para pengikutnya hingga akhir zaman, sehingga penulis dapat menyelesaikan penelitian Tugas Akhir dengan judul **“ANALISIS SENTIMEN ONLINE CUSTOMER REVIEW PADA TOKO SMARTPHONE DAERAH ISTIMEWA YOGYAKARTA DI E-MARKETPLACE SHOPEE MENGGUNAKAN LEXICON BASED DAN WORD CLOUD”** dengan lancar dan tepat waktu.

Dengan dibuatnya Laporan Tugas Akhir ini, maka sudah terselesaikannya salah satu syarat untuk memperoleh gelar Sarjana Strata-1 pada Jurusan Teknik Industri Fakultas Teknologi Industri di Universitas Islam Indonesia. Dengan itu, penulis berharap hasil pembuatan laporan ini dapat memberikan manfaat dengan menambah ilmu pengetahuan bagi pembaca maupun penyusun. Selama pelaksanaan pembuatan laporan, penulis telah mendapatkan banyak pengetahuan, bimbingan maupun saran dari berbagai pihak. Dengan segala kerendahan hati, dan penuh rasa ikhlas penulis ingin mengucapkan terimakasih sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan doa hingga penulis dapat menyelesaikan Laporan Tugas Akhir. Oleh karena itu penulis ingin mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo, M.T. selaku Dekan Fakultas Teknologi Industri Universitas Islam Indonesia.
2. Bapak Dr. Drs. Imam Djati Widodo, M.Eng.Sc. selaku Kepala Jurusan Teknik Industri Universitas Islam Indonesia.
3. Bapak Ir. Muhammad Ridwan Andi Purnomo, S.T.,M.Sc.,Ph.D.,IPM. selaku Ketua Prodi Teknik Industri Fakultas Teknologi Industri Universitas Islam Indonesia.
4. Ibu Annisa Uswatun Khasanah, S.T., M.B.A., M.Sc. selaku Dosen Pembimbing Tugas Akhir Jurusan Teknik Industri Universitas Islam Indonesia yang telah meluangkan waktunya untuk membimbing dan membantu dalam proses penyusunan laporan tugas akhir.
5. Kedua Orangtua beserta kedua saudara Saya yang senantiasa memberikan doa serta dukungannya baik secara moral maupun material.
6. Para sahabat penulis, HoS (Bela, Fajar, Dayu, dan Alya), Hayo (Dayu, Dani, Della), SaTe (Anggita, Rani, Clarissa, Ocha, dan Kiky) yang senantiasa memberikan dukungan baik secara langsung maupun tidak langsung ketika penulis merasa tidak memiliki semangat juang akan TA dan menjadi salah satu motivasi penulis untuk segera menyelesaikan tugas akhir ini.
7. Kepada teman-teman Teknik Industri Universitas Islam Indonesia angkatan 2018 dan semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu dalam pelaksanaan dan penulisan tugas akhir serta dalam menyelesaikan Pendidikan S1 Teknik Industri Universitas Islam Indonesia.

8. Kepada *Team Lead*, *Team 2*, teman-teman Strongbee, *Career Support* serta teman-teman FSDM RevoU yang telah membantu dan memberikan semangat dukungan kepada penulis untuk menyelesaikan penulisan tugas akhir.

Penulis menyadari selama penyusunan laporan tugas akhir ini masih jauh dari kata sempurna dan masih banyak kekurangan, dengan besar hati penulis menerima segala masukan dan saran yang membangun dari semua pembaca demi melengkapi kekurangan pada laporan ini. Semoga laporan tugas akhir ini dapat memberikan manfaat bagi pembaca, Aamiin Yaa Robbal ‘Aalamiin.

Wassalamualaikum Warahmatullahi Wabarakatuh.

Yogyakarta, 12 Agustus 2022



Fildzah Adlina



ABSTRAK

Kemajuan teknologi terus berkembang hingga mempengaruhi berbagai aspek, terutama pada dunia bisnis dan industri. Ditambah lagi adanya pandemi Covid-19 menyebabkan semakin banyak masyarakat lebih memilih untuk melakukan jual beli barang maupun jasa secara *online* melalui *e-marketplace* seperti Shopee yang menjadi salah satu *marketplace* populer di Indonesia. Kehadiran *e-marketplace* memudahkan masyarakat untuk berbelanja berbagai kebutuhan seperti kebutuhan elektronik yaitu *smartphone* yang dalam penggunaannya memiliki peran penting untuk mengakses informasi kapan saja dan dimana saja yang dapat menunjang pelaksanaan *work from home* maupun pembelajaran daring. Berdasarkan hasil penelusuran barang *smartphone* di Daerah Istimewa Yogyakarta pada *e-marketplace* Shopee diperoleh 3 toko yang memiliki performa terbaik yaitu Appelhouse.store, Jordan Corner dan Ifonespot yang menjadi subjek dalam penelitian ini. Adanya faktor *online customer review* menjadi salah satu pemicu terjadinya transaksi jual beli dengan melihat informasi tersebut yang telah tertera di setiap toko. Sehingga adapun tujuan penelitian tugas akhir yaitu melakukan analisis sentimen *online customer review* toko *smartphone* Daerah Istimewa Yogyakarta di *e-marketplace* Shopee menggunakan *lexicon based* dan *word cloud*. Analisis sentimen dilakukan agar data *review* pelanggan dapat dibagi ke beberapa sentimen (positif, negatif, dan netral) menggunakan kamus *SentiStrength* ID, kemudian hasil sentimen diolah kembali dengan menggunakan *word cloud* untuk mengetahui kata sering muncul untuk setiap sentimennya dan dapat diberikan rekomendasi peningkatan performa penjualan. Dari hasil pengumpulan data dari April hingga Juni 2022 menggunakan *web scraping*, diperoleh data kurang dari 500 untuk dilakukan uji penelitian. Hasil *word cloud* sentimen positif menunjukkan ulasan ‘mulus’, ‘bagus’, ‘aman’, dan ‘ramah’, sentimen netral terkait ‘bh/battery health’, ‘mulus’, ‘puas’, dan ‘awet’ serta sentimen negatif mengenai ‘bh/battery health’, ‘turun’, dan ‘kirim’. Berdasarkan hasil *fishbone diagram* dilakukan perbaikan dengan meningkatkan *quality control* produk, menyertakan kebijakan keterlambatan/kerusakan barang pada toko atau deskripsi produk dan meningkatkan kinerja admin dalam melayani pelanggan.

Kata Kunci: *smartphone, online customer review, analisis sentimen, word cloud, fishbone diagram dan SentiStrength ID.*

DAFTAR ISI

HALAMAN JUDUL	i
SURAT PERNYATAAN KEASLIAN	ii
LEMBAR PENGESAHAN DOSEN PEMBIMBING	iii
HALAMAN PERSEMBAHAN	iv
HALAMAN MOTTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Permasalahan	7
1.3 Tujuan Penelitian	8
1.4 Batasan Permasalahan	8
1.5 Manfaat Penelitian	9
1.6 Sistematika Penulisan Laporan	9
BAB II KAJIAN LITERATUR	11
2.1 Kajian Induktif	11
2.2 Kajian Deduktif	21
2.2.1 <i>E-marketplace</i>	21
2.2.2 <i>Online Customer Review dan Rating</i>	22
2.2.3 <i>Web Scraping</i>	22
2.2.4 Analisis Sentimen	23
2.2.5 <i>Text Mining</i>	23
2.2.6 <i>SentiStrength ID</i>	24
2.2.7 <i>Word Cloud</i>	25
2.2.8 Python	25
2.2.9. Asosiasi Teks	26
2.2.10 <i>Fishbone Diagram</i>	27
BAB III METODOLOGI PENELITIAN	29

3.1 Subjek Penelitian dan Objek Penelitian	29
3.2 Sumber Data Penelitian	29
3.3 Alur Penelitian	29
3.3.1 Identifikasi Masalah dan Studi Literatur	30
3.3.2 Pengumpulan Data	31
3.3.3 Pengolahan Data	31
3.3.4 Kesimpulan	32
BAB IV PENGUMPULAN DAN PENGOLAHAN DATA	33
4.1 Pengumpulan Data	33
4.2 Pengolahan Data	35
4.2.1 Pre-Processing data	36
4.2.2 Pelabelan Data Menggunakan <i>SentiStrength</i> ID	38
4.2.3 <i>Word Cloud</i>	40
BAB V PEMBAHASAN	46
5.1 Analisis Sentimen	46
5.2 <i>Word Cloud</i>	47
5.3 Rekomendasi	50
BAB VI PENUTUP	57
6.1 Kesimpulan	57
6.2 Saran	58
DAFTAR PUSTAKA	59
LAMPIRAN	66

DAFTAR TABEL

Tabel 1.1 Perbandingan Ketiga Toko	3
Tabel 2.1 Penelitian Terdahulu Mengenai Analisis Sentimen	13
Tabel 2.2 Penelitian Terdahulu Mengenai Word Cloud dari Hasil Analisis Sentimen	18
Tabel 4.1 Contoh Hasil Pengumpulan Data pada Appelhouse.store.	33
Tabel 4.2 Contoh Hasil Pengumpulan Data pada Iphonespot	34
Tabel 4.3 Contoh Hasil Pengumpulan Data pada Jordan Corner	35
Tabel 4.4 Jumlah Data Yang Akan Diolah	36
Tabel 4.5 Contoh Hasil <i>Case Folding</i>	36
Tabel 4.6 Contoh Hasil <i>Tokenizing</i>	36
Tabel 4.7 Contoh Hasil <i>Filtering</i>	37
Tabel 4.8 Contoh Hasil <i>Stemming</i>	37
Tabel 4.9 Contoh Hasil <i>SentiStrength</i> Toko Appelhouse.Store	38
Tabel 4.10 Contoh Hasil <i>SentiStrength</i> Toko Iphonespot	39
Tabel 4.11 Contoh Hasil <i>SentiStrength</i> Toko Jordan Corner	40
Tabel 5.1 Hasil <i>SentiStrength</i> pada Ketiga Toko Elektronik	46
Tabel 5.2 Perbandingan Hasil Word Cloud Sentimen Positif	47
Tabel 5.3 Perbandingan Hasil Word Cloud Sentimen Netral	48
Tabel 5.4 Perbandingan Hasil Word Cloud Sentimen Negatif	49
Tabel 5.5 Asosiasi Kata Dari Sentimen Negatif Appelhouse.store (1)	50
Tabel 5.6 Asosiasi Kata Dari Sentimen Negatif Appelhouse.store (2)	51
Tabel 5.7 Asosiasi Kata Dari Sentimen Negatif Iphonespot (1)	52
Tabel 5.8 Asosiasi Kata Dari Sentimen Negatif Iphonespot (2)	52
Tabel 5.9 Asosiasi Kata Dari Sentimen Negatif Jordan Corner (1)	53
Tabel 5.10 Asosiasi Kata Dari Sentimen Negatif Jordan Corner (2)	53

Tabel 5.11 Rencana Pemecahan Masalah Toko Appelhouse.store	56
Tabel 5.12 Rencana Pemecahan Masalah Toko Ifonespot	57
Tabel 5.13 Rencana Pemecahan Masalah Toko Jordan Corner	58



DAFTAR GAMBAR

Gambar 1.1 <i>Website E-Commerce</i> Terpopuler di Asia Tenggara	2
Gambar 1.2 Kategori Produk Terpopuler Saat Belanja <i>Online</i>	2
Gambar 1.3 Penetrasi Ponsel Pintar di Indonesia	3
Gambar 3.1 Alur Penelitian	30
Gambar 4.1 Hasil Pelabelan Sentimen Appelhouse.Store	39
Gambar 4.2 Hasil Pelabelan Sentimen Iphonespot	39
Gambar 4.3 Hasil Pelabelan Sentimen Jordan Corner	40
Gambar 4.4 <i>Word Cloud</i> Sentimen Positif pada Toko Appelhouse.store	41
Gambar 4.5 <i>Word Cloud</i> Sentimen Positif pada Toko Iphonespot	41
Gambar 4.6 <i>Word Cloud</i> Sentimen Positif pada Toko Jordan Corner	42
Gambar 4.7 <i>Word Cloud</i> Sentimen Netral pada Toko Appelhouse.store	42
Gambar 4.8 <i>Word Cloud</i> Sentimen Netral pada Toko Iphonespot	43
Gambar 4.9 <i>Word Cloud</i> Sentimen Netral pada Toko Jordan Corner	43
Gambar 4.10 <i>Word Cloud</i> Sentimen Negatif pada Toko Appelhouse.store	44
Gambar 4.11 <i>Word Cloud</i> Sentimen Negatif Iphonespot	44
Gambar 4.12 <i>Word Cloud</i> Sentimen Negatif Jordan Corner	45
Gambar 5.1 Diagram <i>Fishbone</i> Keluhan Pelanggan Toko Appelhouse.store	54
Gambar 5.2 Diagram <i>Fishbone</i> Keluhan Pelanggan Toko Iphonespot	55
Gambar 5.3 Diagram <i>Fishbone</i> Keluhan Pelanggan Toko Jordan Corner	55

BAB I PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi terus berkembang hingga mempengaruhi berbagai aspek, terutama pada dunia bisnis dan industri. Internet yang semakin berkembang juga ikut merubah berjalannya bisnis (Reynolds, 2004). Ditambah lagi adanya pandemi Covid-19 menyebabkan semakin banyak masyarakat lebih memilih untuk melakukan jual beli barang maupun jasa secara *online* dibandingkan langsung mendatangi tokonya. Karena sebagian besar sektor kehidupan diubah menjadi kegiatan dalam jaringan, sehingga kebutuhan masyarakat akan alat komunikasi elektronik seperti laptop, *handphone*, dan alat komunikasi elektronik penunjang lainnya menjadi meningkat. Perilaku masyarakat ini akan memberikan pengaruh secara langsung kepada tingkat produktivitas, efisiensi, dan pendapatan perusahaan elektronik terkait. Hal tersebut membuat para pebisnis untuk ikut terjun dalam melakukan penjualan secara *online* (Nuraeni *et al*, 2022).

Kegiatan jual beli barang dan jasa secara online atau *e-commerce* semakin berkembang pesat. Hingga saat ini, pengguna internet Indonesia berjumlah lebih dari 132 juta dengan tingkat penetrasi sebesar 50%. Namun tercatat hanya 45% dari pengguna internet tersebut yang melakukan kegiatan *e-commerce* (Hootsuite, 2018). Perkembangan *e-commerce* juga dimanfaatkan oleh berbagai pelaku bisnis sebagai peluang penghasilan baru. *E-commerce* menawarkan pengguna dalam melakukan sebuah transaksi barang maupun jasa tanpa batas jarak dan waktu selama pengguna memiliki akses internet (Işoraité & Miniotiené, 2018).

Berdasarkan data peta *E-Commerce* Indonesia yang dikeluarkan oleh iPrice, Shopee menjadi salah satu *marketplace* dengan model *e-commerce* yang populer khususnya di Indonesia dan bahkan di 5 negara di Asia Tenggara Lainnya pada kuartal pertama tahun 2022 dengan mengetahui perkembangan E-Commerce dari rataan total kunjungan *website* (*Desktop* dan *Mobile Phone*) dan *engagement* sosial (iPrice, 2022).

Top 5 Website E-Commerce Terpopuler di Asia Tenggara

Data peringkat website e-commerce berdasarkan jumlah pengunjung website terbanyak pada Q1 2022.

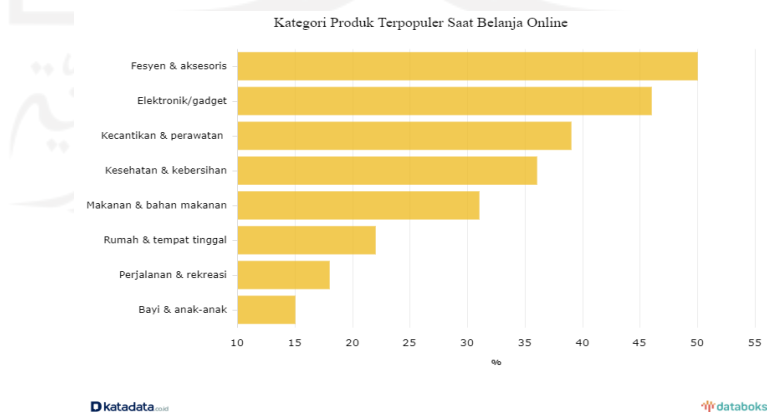
Peringkat	Indonesia	Malaysia	Singapura	Thailand	Filipina	Vietnam
1	Tokopedia	Shopee	Shopee	Shopee	Shopee	Shopee
2	Shopee	PG Mall	Lazada	Lazada	Lazada	Thế Giới Di Động
3	Lazada	Lazada	Amazon	Central Online	Zalora	Điện Máy Xanh
4	Bukalapak	Zalora	Qoo10	JD Central	Ebay	Lazada
5	Orami	GoShop	Castlery	HomePro	Beauty MNL	Tiki

Metodologi: Data peringkat e-commerce diambil berdasarkan jumlah pengunjung website terbanyak yang terdapat dalam laporan peta e-commerce iPrice pada Q1 2022 di Indonesia, Malaysia, Singapura, Thailand, Filipina dan Vietnam.
Source: iPrice Group • Created with Datawrapper

Gambar 1.1 Website E-Commerce Terpopuler di Asia Tenggara

Sumber: iPrice, 2022

Kehadiran *e-marketplace* di Indonesia semakin memudahkan masyarakat untuk berbelanja berbagai kebutuhan salah satunya adalah kebutuhan elektronik seperti laptop, handphone, dan alat komunikasi elektronik penunjang lainnya. Indonesia memiliki potensi yang cukup besar untuk ditargetkan oleh banyak industri elektronik luar negeri sebagai pangsa pasar yang besar dan tak jarang industri elektronik tersebut memusatkan beberapa kegiatan produksinya di Indonesia. Tidak salah jika industri elektronik bersama industri logam, mesin, dan industri kedokteran menempati peringkat pertama dalam “5 Besar Nilai Investasi Sektor Industri” dan menempati peringkat kelima dalam “6 Sektor Industri dengan Nilai Ekspor Terbesar” (Kemenperin, 2019). Hal tersebut pun juga diperkuat dalam survei Jajak Pendapat (Jakpat), bahwa elektronik/*gadget* menjadi produk yang paling populer kedua saat berbelanja secara *online* dengan persentase mencapai 46% (Databoks.katadata.co.id, 2022).



Gambar 1.2 Kategori Produk Terpopuler Saat Belanja Online

Sumber: Databoks.katadata, 2022

Penggunaan *smartphone* atau ponsel pintar di Indonesia diprediksi akan terus meningkat ditambah lagi dengan adanya pandemi Covid-19, membuat *smartphone* menjadi akses informasi kapan saja dan dimana saja untuk menunjang pelaksanaan *work from home* maupun pembelajaran daring di dunia pendidikan (Darmalaksana, 2020). Pada 2015, hanya terdapat 28,6% populasi di Indonesia yang menggunakan gawai tersebut. Seiring berjalannya waktu, ponsel pintar semakin terjangkau, sehingga meningkatkan penggunaannya pula. Lebih dari setengah populasi di Indonesia atau 56,2% telah menggunakan ponsel pintar pada 2018. Setahun setelahnya, sebanyak 63,3% masyarakat menggunakan ponsel pintar. Hingga 2025, setidaknya 89,2% populasi di Indonesia telah memanfaatkan ponsel pintar. Dalam kurun waktu enam tahun sejak 2019, penetrasi ponsel pintar di tanah air tumbuh 25,9% (Databoks.katadata.co.id, 2020)



Gambar 1.3 Penetrasi Ponsel Pintar di Indonesia

Sumber: Databoks.katadata, 2020

Berdasarkan hasil penelusuran barang elektronik di Daerah Istimewa Yogyakarta pada *marketplace* Shopee diperoleh 3 toko yang memiliki performa terbaik berdasarkan jumlah penjualan produk terbesar, jumlah *followers* dan penilaian produk (*rating*) yang tinggi yaitu Appelhouse.store, Jordan Corner dan Ifonespot.

Tabel 1.1 Perbandingan Ketiga Toko

Faktor Kelebihan	Appelhouse.store	Ifonespot	Jordan Corner
Jumlah Pengikut	+/-25.700	+/-26.400	+/-6.200
Jumlah Penjualan	+/-5.100	+/-2.900	+/-1.400

Faktor Kelebihan	Appelhouse.store	Ifonespot	Jordan Corner
Rating Toko	4,8/5	4,9/5	4.9/5
Performa Chat (dalam jam)	95%	94%	89%
Waktu Bergabung	35 bulan	5 Tahun	21 Bulan

David (2014) menyatakan dalam penelitiannya bahwa *e-marketplace* memiliki manfaat bagi para pelaku usaha dalam membuat strategi baru untuk melakukan pemasaran sehingga berdampak jangkauan pemasaran menjadi lebih efektif dan efisien. Mudahnnya berbelanja di *online shop* menjadi salah satu alasan terjadinya perubahan perilaku pelanggan. Kebiasaan masyarakat yang sebelumnya ketika ingin membeli suatu produk, mereka akan secara otomatis mencari informasi dengan cara melihat fisik produk baik kualitas bahan, kecocokan dengan harga, kualitas pelayanan dan lain sebagainya dapat dilihat dan dirasakan secara langsung. Berbeda ketika masyarakat akan membeli suatu produk melalui *online shop*, mereka tidak akan bisa melihat dan merasakan fisik produk serta kualitas pelayanan secara langsung, maka dari itu pembeli akan berusaha mendapatkan informasi produk dan jasa yang akan mereka beli melalui *rating* dan *online customer review*.

Review yang ada menjadi salah satu pemicu terjadinya penjualan dengan melihat informasi yang bisa diakses seperti *review* dari blog atau tulisan-tulisan yang diberikan oleh pemilik *website* tentang suatu produk, dari mulai spesifikasi, kelebihan dan kekurangan dari produk serta harga dari produk. Sedangkan *rating* juga dapat mempengaruhi keputusan pembeli dalam memutuskan pembelian pada suatu toko *online*, tinggi rendahnya *rating online shop* mencerminkan banyak tidaknya pembeli pada toko online tersebut. Penelitian sebelumnya menyebutkan 92% konsumen akan ragu untuk membeli produk atau layanan jika tidak ada ulasan pelanggan (Mangles, 2017). Dalam penelitian lain yang menggunakan pelanggan *e-commerce* Indonesia sebagai objeknya, menyebutkan bahwa *Online Customer Review* terbukti mempengaruhi minat membeli dan kepercayaan pelanggan dalam melakukan transaksi *online* (Farki & Baihaqi, 2016). Maka dari itu perlu ada perhatian lebih untuk faktor ini karena dapat mempengaruhi performa penjualan.

Penelitian ini menggunakan metode *web scraping* untuk proses pengambilan data di shopee agar dapat melakukan ekstraksi data dengan jumlah besar dan pengukuran yang bervariasi pada suatu *website* (Ayani, 2019). Hasil dari *web scraping* tersebut kemudian diolah dengan menggunakan teknik analisis sentimen untuk menghasilkan sebuah informasi berupa pandangan seseorang terhadap suatu kejadian atau isu. Pada hasil *Online Customer Review* terdapat teks yang tidak terstruktur mudah diproses dan dirasakan oleh manusia, tetapi secara signifikan lebih sulit untuk dipahami oleh mesin (Allahyari et al., 2017). Maka dari itu, analisis sentimen diperlukan untuk dapat memahami opini seseorang yang tertulis di platform online secara otomatis dengan bantuan mesin. Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau fitur atau tingkat aspek dan menentukan apakah pendapat yang dikemukakan dalam dokumen, kalimat atau fitur entitas atau aspek bersifat positif, negatif atau netral (Murnawan, 2017).

Analisis sentimen dalam pengaplikasiannya memiliki beberapa metode yang dapat digunakan yaitu *Lexicon Based*, *Machine Learning Based Approach* dan *Hybrid Approach* (gabungan antara *Lexicon Based* dan *Machine Learning Based Approach*). *Lexicon-based* menggunakan berbagai kata yang dinilai dengan skor polaritas dan subjektivitas untuk mengetahui tanggapan masyarakat/pengguna/konsumen mengenai suatu topik berdasarkan kamus leksikon dengan masing-masing kata yang telah diberi label sebagai sentimen positif, sentimen negatif, dan sentimen netral. Leksikon berarti kosa kata seseorang, bahasa atau cabang ilmu pengetahuan (Thomas, 2021). Dengan *lexicon-based* dapat diperoleh setiap data tekstual beserta label polaritas dan subjektivitas, dari sini lah diperlukan membangun model menggunakan *machine learning based* untuk mengklasifikasikan dan memprediksi *input* masa depan ke dalam berbagai kategori sentimen. *Machine learning* memiliki fitur-fitur di dalamnya yang dapat mengenali sentimen (sudut pandang seseorang) dalam sebuah teks, metode ini semakin populer karena dapat dinilai representatif (Ahmed et al, 2013). Inti dari *machine learning* untuk membuat model sistematis yang merefleksikan pola-pola data menggunakan statika atau aljabar linear dan salah satu bagian dari *Artificial Intelligence* (AI) yang berfokus pada pembuatan mesin yang dapat belajar tanpa diprogram secara eksplisit (Putra, 2019). Sementara itu *Hybrid Approach* adalah pendekatan yang menggabungkan *lexicon based* dan *machine learning based approach*. Peneliti membuktikan bahwa kombinasi dari kedua pendekatan memberikan peningkatan kinerja dari klasifikasi

(Sindhvani and Melville, 2008). Keuntungan dari pendekatan *hybrid* adalah membuat deteksi dan pengukuran sentimen pada tingkat konsep, serta akurasi tinggi dari algoritma *machine learning* yang kuat (D'Andrea *et al*, 2015). Penelitian ini menggunakan analisis sentimen dengan pendekatan *lexicon based* agar dapat mendefinisikan *online customer review* yang masih berbentuk teks, tidak terstruktur dan secara signifikan sulit untuk dipahami oleh mesin. Sehingga hasil tersebut dapat didefinisikan ke dalam sentimen positif, negatif, dan netral. *Lexicon Based* merupakan metode analisis sentimen yang sederhana, layak, dan praktis dalam melakukan analisis sentimen, karena menggunakan kamus sebagai sumber bahasa atau leksikal (Mahendrajaya *et al*, 2019). Polaritas adalah sesuatu yang dapat mengidentifikasi emosi kalimat tertentu dengan menggunakan kata-kata yang ada di dalamnya. Salah satu teknik untuk melakukan analisis sentimen dengan *Lexicon Based* adalah menggunakan *Sentiment Strength (SentiStrength) ID*. *SentiStrength* merupakan algoritma klasifikasi yang memanfaatkan pendekatan leksikon atau kamus dengan menggunakan kamus emosikon dan kamus ungkapan yang diberi bobot oleh manusia. Kamus sentimen ini didapatkan dari hasil translasi kamus sentimen berbahasa Inggris yang telah mengalami penambahan dan pengurangan kata berdasarkan dari hasil pengamatan dalam proses pengembangan sistem ini. Penelitian ini melibatkan 3 responden ahli, mahasiswa S2 Linguistik Universitas Gadjah Mada untuk memberikan bobot kekuatan sentimen. Nilai tengah dari ketiga responden yang dijadikan bobot akhir kekuatan sentimen (Wahid & Azhari, 2016). Metode ini mulai banyak digunakan dalam beberapa jurnal Indonesia karena salah satu kamus leksikon dalam bahasa Indonesia yang mudah diakses dan dapat digunakan untuk pembuatan *prototipe* cepat serta terbukti sumber pembuatan kamusnya melibatkan 3 orang *expert*. Sehingga dalam penelitian ini memilih metode *Lexicon Based* dengan kamus *SentiStrength ID* untuk melakukan analisis sentimen terhadap *review* pelanggan Shopee.

Dari hasil analisis sentimen dilanjutkan dengan memvisualisasikan kata kata yang berapa pada sentimen positif, negatif, dan netral dengan menggunakan *word cloud* agar dapat mengetahui faktor apa saja yang mempengaruhi kepuasan pelanggan pada ketiga toko *smartphone* berupa kata-kata. *Word cloud* adalah metode text mining yang menampilkan grafis dari frekuensi kata yang lebih menonjolkan kata-kata yang lebih sering muncul dalam teks sumber. Semakin besar ukuran kata dalam visual, maka semakin umum kata tersebut dalam dokumen (Alamsyah and Zuhri 2017). Sehingga dengan hasil *word cloud* pada setiap

sentimen positif, negatif dan netral diharapkan dapat dijadikan dasar untuk mengetahui performansi toko secara garis besar dengan diperolehnya kelebihan (sentimen positif dan netral) dan kelemahan (sentimen negatif) setiap tokonya. Setelah itu dilakukan analisis diagram *fishbone* untuk menemukan akar permasalahan pada kata kunci yang telah didapatkan dari ulasan negatif yang kemudian dapat dijadikan referensi untuk rencana perbaikan agar dapat meningkatkan performansi toko menjadi lebih baik.

Berdasarkan hasil latar belakang yang telah dijabarkan, maka dapat dilakukan penelitian tugas akhir terkait Analisis Sentimen *Online Customer Review* Pada Toko *Smartphone* Daerah Istimewa Yogyakarta *E-Marketplace* Shopee Menggunakan *Lexicon Based* dan *Word Cloud*.. Proses pengambilan data dilakukan dengan menggunakan metode *web scraping* terhadap *website* Shopee pada 3 toko *online smartphone* terbaik yaitu *Appelhouse.store*, *Jordan Corner & Iphonespot* pada *website* Shopee di Daerah Istimewa Yogyakarta. Kemudian data tersebut akan diolah dan dianalisis dengan menggunakan analisis sentimen dengan *library SentiStrength* agar dapat menginterpretasikan *review customer* ke dalam aspek positif, negatif, dan netral. Dari hasil analisis sentimen, maka dapat dilanjutkan pada tahap pengolahan data selanjutnya yaitu dengan memvisualisasikan kata-kata pada setiap sentimennya dengan menggunakan *word cloud* agar dapat mengetahui faktor apa saja yang mempengaruhi kepuasan pelanggan pada ketiga toko *smartphone*. Dengan harapan hasil tugas akhir ini dapat dijadikan acuan sebagai sistem rekomendasi dan informasi yang bermanfaat bagi pihak yang membutuhkan di *e-marketplace* untuk mendukung perkembangan bisnis *online* di Indonesia khususnya Daerah Istimewa Yogyakarta.

1.2 Rumusan Permasalahan

Adapun rumusan masalah yang dapat diambil berdasarkan pembahasan pada latar belakang sebagai berikut:

1. Bagaimana hasil analisis sentimen menggunakan *Sentiment Strength* ID pada Python yang diperoleh dari *online customer review*?
2. Berdasarkan hasil *word cloud*, faktor apa saja yang paling mempengaruhi analisis sentimen dari *online customer review* toko *smartphone* Daerah Istimewa Yogyakarta di *e-marketplace* Shopee?

3. Apa rekomendasi yang dapat diberikan untuk meningkatkan performa toko *smartphone* Daerah Istimewa Yogyakarta dari hasil *online customer review* di *e-marketplace* Shopee?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan yang ingin dicapai penulis dalam penelitian tugas akhir ini adalah:

1. Menjelaskan hasil analisis sentimen menggunakan *Sentiment Strength ID* yang diperoleh dari *online customer review*.
2. Menganalisa faktor yang paling mempengaruhi analisis sentimen dari *online customer review* toko *smartphone* Daerah Istimewa Yogyakarta di *e-marketplace* Shopee berdasarkan hasil *word cloud*.
3. Menjelaskan rekomendasi yang dapat diberikan untuk meningkatkan performa toko *smartphone* Daerah Istimewa Yogyakarta dari hasil *online customer review* di *e-marketplace* Shopee.

1.4 Batasan Permasalahan

Adapun batasan penelitian tugas akhir yang perlu dilakukan sebagai berikut:

1. Pengambilan data hanya dilakukan pada *e-marketplace* yaitu Shopee terutama 3 toko *online smartphone* terbaik (Appelhouse.store, Jordan Corner & Ifonespot) di Daerah Istimewa Yogyakarta
2. Penelitian ini menggunakan metode *web scraping* (Python) untuk pengambilan data pada *website* Shopee.
3. Pengambilan data *review* pelanggan menggunakan *web scraping* dilakukan pada tanggal 20 April 2022 hingga 19 Juni 2022.
4. Menggunakan beberapa *tools* untuk membantu dalam pengolahan data yaitu Microsoft Excel, *software* R, *orange data mining* dan Jupyter Notebook dengan bahasa pemrograman yang digunakan Python versi 3.9.

1.5 Manfaat Penelitian

Adapun manfaat dari penulisan tugas akhir ini yang diharapkan dapat bermanfaat bagi perguruan tinggi, para pelaku bisnis *online* pada *e-marketplace* Shopee, dan bagi penulis sendiri sebagai berikut:

a. Bagi Penulis

Tugas akhir memberikan manfaat yang cukup besar bagi penulis dalam memahami lebih lanjut terkait pengaplikasian ilmu yang telah didapatkan selama bangku perkuliahan di jurusan Teknik Industri Universitas Islam Indonesia. Penelitian ini juga membuat penulis mendapatkan pengetahuan baru dari segala sisi terutama mengenai pengolahan data dalam penjualan produk di salah satu *e-marketplace* yang berdiri di Indonesia.

b. Bagi Pelaku Bisnis *Online*

Hasil dari penelitian tugas akhir ini diharapkan dapat dijadikan salah satu acuan sebagai perencanaan dan perancangan strategi pelaku bisnis *online* di *e-marketplace* untuk mendukung perkembangan bisnis *online* di Indonesia.

c. Bagi Peneliti Selanjutnya

Penelitian dapat dijadikan sumber referensi bagi penelitian selanjutnya dalam penerapan keilmuan *Business Intelligent* dan Statistika Industri 2 dengan topik yang serupa.

1.6 Sistematika Penulisan Laporan

Adapun sistematika penulisan laporan tugas akhir ini dibuat agar penulisan tugas akhir lebih terstruktur sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini membahas mengenai latar belakang diadakannya penelitian, rumusan masalah, tujuan dari penelitian, batasan masalah dalam penelitian, manfaat penelitian dan sistematika penulisan penelitian tugas akhir yang dibuat.

BAB II : KAJIAN LITERATUR

Isi dari bab II menjelaskan kajian literatur deduktif dan induktif dari penelitian terdahulu yang topiknya serupa sehingga dapat membuktikan

bahwa topik penelitian yang dipilih penulis telah memenuhi syarat dan kriteria yang telah dijelaskan.

BAB III : METODE PENELITIAN

Dalam bab ini berisikan informasi obyek penelitian yang digunakan dan menjelaskan metode penelitian yang digunakan penulis baik primer maupun sekunder serta penulis dapat menjelaskan alur penelitian yang digunakan dari awal hingga selesai.

BAB IV : PENGUMPULAN DAN PENGOLAHAN DATA

Bab ini memuat data yang diperoleh dari hasil pengambilan data dengan prosedur tertentu dan data tersebut akan diolah dengan menggunakan metode yang telah ditentukan, kemudian data tersebut dianalisis agar diperoleh hasil penelitiannya.

BAB V : PEMBAHASAN

Pada bab ini berisikan pembahasan kritis dari hasil yang didapatkan pada bab sebelumnya, dimana hasil pembahasan dapat dijadikan acuan dasar dalam penentuan usulan penelitian selanjutnya pada bab berikutnya.

BAB VI : KESIMPULAN DAN SARAN

Isi dari bab ini memuat kesimpulan yang didapatkan dari hasil penelitian dan saran yang dapat diberikan kepada peneliti selanjutnya agar penelitian yang serupa dapat memperluas pengembangan ilmu Teknik Industri.

DAFTAR PUSTAKA

Daftar Pustaka berisikan studi literatur terdahulu yang digunakan selama melaksanakan penelitian serta proses penulisan laporan penelitian, dapat berupa jurnal, buku, prosiding, atau kutipan-kutipan yang diperoleh dari internet.

LAMPIRAN

Lampiran memuat dokumen tambahan dapat berisikan data berupa gambar atau tabel, serta berisikan gambar dokumentasi dari pelaksanaan penelitian. Dimana lampiran digunakan sebagai penunjang dalam penulisan penelitian dan berfungsi untuk memberikan penjelasan lebih detail mengenai penelitian yang dilakukan, tetapi hal tersebut tidak dapat ditampilkan pada isi laporan karena dikhawatirkan dapat mengganggu isi dari penelitian.

BAB II KAJIAN LITERATUR

2.1 Kajian Induktif

Penelitian yang dilakukan Wahid dan Azhari (2016) bertujuan mengkombinasikan metode *SentiStrength*, *Hybrid TF-IDF* dan *Cosine Similarity* untuk mengekstraksi ringkasan sentimen positif dan negatif masyarakat terhadap topik selebriti di Twitter secara otomatis, dengan artis Agnes Monica sebagai studi kasus. Metode *SentiStrength* digunakan untuk mendapatkan skor kekuatan sentimen dan mengklasifikasi tweet ke dalam kelas positif, negatif dan netral. Tweet bersentimen positif dan negatif diringkas dengan cara pemeringkatan tweet menggunakan *Hybrid TF-IDF* yang dikombinasi dengan skor kekuatan sentimen, kemudian menghilangkan tweet yang mirip menggunakan *Cosine Similarity*. Hasil pengujian memperlihatkan bahwa kombinasi *SentiStrength*, *Hybrid TF-IDF*, dan *Cosine Similarity* mampu menghasilkan ringkasan sentimen dengan akurasi yang lebih baik dibandingkan menggunakan *Hybrid TF-IDF* saja, dengan perolehan akurasi rata-rata sebesar 60% dan f-measure sebesar 62%. Hal ini disebabkan karena penambahan kekuatan sentimen sebagai bobot peringkasan.

Untuk mengetahui opini pengguna terkait penggunaan Go-Pay melalui sosial media Twitter, Mahendrajaya *et al* (2019) melakukan penelitian Analisis Sentimen Pengguna Gopay Menggunakan Metode *Lexicon Based* Dan *Support Vector Machine*. Media twitter memiliki kelebihan tampilan simple, topik terupdate, terbuka mengakses tweet dan cepat dalam menyampaikan opini. Dari berbagai ragam komentar di Twitter diperlukan teknik untuk membagi ke dalam kelas opini positif ataupun negatif. Penelitian ini menggunakan *preprocessing* dan melabeli opini ke dalam kelas positif dan negatif dengan metode *lexicon Based*. Sedangkan untuk klasifikasinya menggunakan metode *Support Vector Machine* (SVM). Data yang digunakan berupa opini tentang ulasan Go-Pay dari media sosial Twitter yang berjumlah 1210 data. Hasil dari pelabelan dengan *Lexicon Based* berjumlah 923 untuk positif dan 287 untuk negatif. Sedangkan klasifikasi metode SVM menggunakan kernel Linear menghasilkan 89,17% dan 84,38% untuk kernel Polynomial.

Pada penelitian Khaira *et al* (2020) mengangkat permasalahan *Bullying* telah berkembang di Twitter selama beberapa tahun terakhir dengan menggunakan analisis

sentimen. Analisis sentimen dilakukan untuk mengidentifikasi unsur *bullying* dalam sebuah tweet. Sentimen adalah dibagi menjadi 3 kelas yaitu *Bullying*, *Non-Bullying* dan netral. ini menggunakan *sentiStrength*, sebuah algoritma yang menggunakan pendekatan berbasis leksikon. Leksikon *SentiStrength* ini mengandung bobot sentimennya kekuatan. Hasil penilaian dari 454 data *tweet* diperoleh 161 *tweet* non-intimidasi (35,4%), 87 *tweet* netral (19,1%), dan 206 *tweet* intimidasi (45,4%). Penelitian ini menghasilkan nilai akurasi 60,5%.

Abdillah *et al* (2021) melakukan uji perbandingan metode ekstraksi fitur dan leksikon bahasa Indonesia pada analisis sentimen teks di media sosial terkait penanganan Covid-19. Set data yang digunakan berasal dari penelitian sebelumnya yang divalidasi ahli bahasa. Penelitian ini bertujuan mengukur performa kamus sentimen kata dari leksikon InSet dan *sentistrength_id* terhadap label aktual, berikutnya untuk mengetahui pengaruh pemilihan metode ekstraksi fitur *term presence*, BoW, dan TF-IDF. Pengujian awal menunjukkan bahwa kamus sentimen kata leksikon *sentistrength_id* menghasilkan skor akurasi sedikit lebih tinggi (64,46%) dibanding InSet (62,65%) saat dikombinasikan dengan TF-IDF. Di tahap evaluasi akhir, kamus sentimen kata *sentistrength_id* masih menunjukkan performa yang relatif lebih baik berdasarkan rerata nilai akurasi, presisi, sensitivitas, dan f-measure (59,22%, 61,1%, 31,3%, 41,0%) dibanding InSet (58,98%, 62,2%, 26,1%, 36,6%) saat dikombinasikan dengan TF-IDF. Secara umum, performa kedua leksikon memang masih di bawah data dengan label aktual yang mengindikasikan kedua leksikon belum cukup kuat dalam menentukan kelas kata. Penyebabnya bisa karena jumlah data yang relatif sedikit atau karena normalisasi teks belum cukup dilakukan secara maksimal.

Pada penelitian yang dilakukan Hernikawati (2021) mengenai Kecenderungan Tanggapan Masyarakat Terhadap Vaksin Sinovac Berdasarkan *Lexicon Based Sentiment Analysis*, memiliki tujuan penelitian dengan melihat tanggapan masyarakat terhadap vaksin Sinovac, apakah sentimen yang diberikan lebih banyak yang positif, netral, atau negatif berdasarkan data Twitter. Dari hasil tersebut dibandingkan dengan tanggapan masyarakat internasional terhadap vaksin Sinovac. Metode *Lexicon Based* digunakan untuk melakukan analisis sentimen, diperoleh hasil bahwa opini netral (37.6 %) memiliki prosentase tertinggi jika dibandingkan dengan opini positif (35.4 %) dan opini negatif (27.0 %). Analisis sentimen masyarakat dunia juga mendapatkan hasil yang sama yaitu opini netral (69,4%) lebih dominan dibandingkan opini yang lain. Prosentase opini netral dunia lebih tinggi jika

dibandingkan dengan prosentase opini netral di Indonesia. Opini netral menunjukkan bahwa masyarakat tidak mendukung dan tidak menolak dengan adanya vaksin Sinovac.

Pada penelitian Aziz dan Fauziah (2022) melakukan sistem analisis sentimen dengan menggunakan Algoritma TF-IDF dan *SentiStrength* untuk mengidentifikasi opini masyarakat terhadap produk, layanan, dan kebijakan perusahaan di Twitter. Pada data uji yang diambil dengan menggunakan kata kunci “vaksin anak” sebanyak 1000 tweet, diperoleh hasil sentimen positif 54%, sentimen negatif 20% dan sentimen netral 26%. Perbandingan dengan analisis data yang sama menggunakan algoritma yang berbeda yaitu *Naïve Bayes* mendapatkan hasil sentimen positif 55%, negatif 16% dan netral 29%. *Decision Tree* mendapatkan hasil positif 61%, negatif 14% dan netral 25%.

Tabel 2.1 Penelitian Terdahulu Mengenai Analisis Sentimen

Penulis	Metode	Objek	Hasil
Wahid & Azhari (2016)	1) Metode pengambilan data: <i>Crawling Data</i> dengan API Twitter 2) Metode pengolahan data: <i>SentiStrength</i> , <i>Hybrid TF-IDF</i> dan <i>Cosine Similarity</i>	Ulasan pengguna mengenai artis Agnes Monica di Twitter	Hasil pengujian memperlihatkan bahwa kombinasi <i>SentiStrength</i> , <i>Hybrid TF-IDF</i> , dan <i>Cosine Similarity</i> mampu menghasilkan ringkasan sentimen dengan akurasi yang lebih baik dibandingkan menggunakan <i>Hybrid TF-IDF</i> saja, dengan perolehan akurasi rata-rata sebesar 60% dan f-measure sebesar 62%. Hal ini disebabkan karena penambahan kekuatan sentimen sebagai bobot peringkasan.
Mahendrajaya <i>et al</i> (2019)	1) Metode pengambilan data: <i>Crawling Data</i> dengan API Twitter 2) Metode pengolahan data: <i>Lexicon Based</i> dan	Ulasan pengguna Gopay di Twitter	Hasil analisis sentimen dari pelabelan dengan <i>Lexicon Based</i> adalah diperoleh 923 ulasan positif dan 287 ulasan negatif. Sedangkan klasifikasi metode SVM menggunakan kernel Linear menghasilkan 89,17% dan 84,38% untuk kernel Polynomial

Penulis	Metode	Objek	Hasil
	<i>Support Vector Machine</i>		
Khaira <i>et al</i> (2020)	1) Metode pengambilan data: <i>Crawling Data</i> 2) Metode pengolahan data: <i>Lexicon Based</i> menggunakan <i>SentiStrength</i>	Ulasan dengan unsur bullying di Twitter	Hasil penilaian dari 454 data <i>tweet</i> diperoleh 161 <i>tweet</i> non-intimidasi (35,4%), 87 <i>tweet</i> netral (19,1%), dan 206 <i>tweet</i> intimidasi (45,4%). Penelitian ini menghasilkan nilai akurasi 60,5%.
Abdillah <i>et al</i> (2021)	1) Metode pengambilan data: <i>Crawling Data</i> 2) Metode pengolahan data: <i>Lexicon Based</i> (InSet dan <i>sentistrength_id</i>) dan metode ekstraksi (<i>term presence</i> , BoW, dan TF-IDF)	Ulasan mengenai penanganan Covid-19 di Twitter	Pengujian awal menunjukkan bahwa kamus sentimen kata leksikon <i>sentistrength_id</i> menghasilkan skor akurasi sedikit lebih tinggi (64,46%) dibanding InSet (62,65%) saat dikombinasikan dengan TF-IDF. Di tahap evaluasi akhir, kamus sentimen kata <i>sentistrength_id</i> masih menunjukkan performa yang relatif lebih baik berdasarkan rerata nilai akurasi, presisi, sensitivitas, dan f-measure (59,22%, 61,1%, 31,3%, 41,0%) dibanding InSet (58,98%, 62,2%, 26,1%, 36,6%) saat dikombinasikan dengan TF-IDF.
Hernikawati (2021)	1) Metode pengambilan data: <i>Crawling Data</i> 2) Metode pengolahan	Ulasan Masyarakat di Twitter terkait vaksin Sinovac	Dari hasil analisis sentimen menggunakan <i>Lexicon Based</i> , didapatkan bahwa opini netral (37.6 %) memiliki prosentase tertinggi jika dibandingkan dengan opini positif (35.4 %) dan opini negatif (27.0 %).

Penulis	Metode	Objek	Hasil
	data: <i>Lexicon Based</i>		Analisis sentimen masyarakat dunia juga mendapatkan hasil yang sama yaitu opini netral (69,4%) lebih dominan dibandingkan opini yang lain.
Aziz & Fauziah (2022)	1) Metode pengambilan data: API Twitter 2) Metode pengolahan data: Algoritma TF-IDF, <i>SentiStrength</i> , <i>Naïve Bayes</i> dan <i>Decision Tree</i>	Opini masyarakat terhadap produk, layanan, dan kebijakan perusahaan di Twitter.	Pada data uji yang diambil dengan menggunakan kata kunci “vaksin anak” sebanyak 1000 tweet, diperoleh hasil sentimen positif 54%, sentimen negatif 20% dan sentimen netral 26%. Perbandingan dengan analisis data yang sama menggunakan algoritma yang berbeda yaitu <i>Naïve Bayes</i> mendapatkan hasil sentimen positif 55%, negatif 16% dan netral 29%. <i>Decision Tree</i> mendapatkan hasil positif 61%, negatif 14% dan netral 25%.

Penelitian mengenai Analisis *Text Clustering* Akun *Fanpage* Shopee Indonesia Dengan Komentar *Followers* Menggunakan *Tools Orange Data Mining* yang dilakukan oleh Sentiya dan Suroyo (2019) yang memanfaatkan tahapan *Preprocess Text* terdiri dari *transformation*, *tokenization*, *normalization*, dan *filtering* berfungsi agar teks dari komentar *followers* dapat dibaca dan dianalisis oleh aplikasi *Orange Data Mining* sehingga diperoleh topik pembahasan dari kata yang sering muncul dari isi komentar *followers* Shopee Indonesia. Kumpulan teks komentar yang berjumlah banyak dikumpulkan dengan menggunakan Web Scraper hasil data *fanpage* Shopee Indonesia dapat dilakukan analisis untuk mengetahui gambaran umum akan topik yang dibahas dan kata yang sering muncul dari status dan komentar dari *followers* Shopee Indonesia sehingga dapat dilakukan pembuatan *cluster* untuk setiap topik pembicaraannya. Metode *clustering* yang digunakan yaitu *k-means* yang menghasilkan 3 *cluster* membahas topik pembicaraan tentang kuis, tebak-tebakan dan *retweet* yang memberikan *handphone* gratis di *fanpage* Shopee Indonesia.

Dari hasil analisis *text clustering* diperoleh topik pembicaraan yang dibahas pada *fanpage* Shopee Indonesia mengenai hadiah, iphone, vote.

Rahman dan Suroyo (2019) juga melakukan penelitian mengenai Analisis Teks Komentar Produk Pada E-Commerce Dengan Menggunakan Algoritma Teks *Clustering* Berbasis *Python*. Dengan data yang diteliti yaitu teks komentar, angka penjualan dan skor *rating* bintang diperoleh hasil analisis teks komentar produk menggunakan *word cloud* pada produk *Smartphone low cost* menunjukkan data komentar *marketplace* shopee Indonesia baik di *smartphone* dengan *low cost* maupun yang *high cost* cenderung memiliki pola *word cloud* yang mirip dengan kata-kata dominan yang muncul cenderung bersifat netral dan positif, sedangkan kata-kata yang bersifat negatif tidak terlalu dominan. Kata-kata yang sering muncul adalah barang, mantap, cepat, kirim dan bagus dengan presentase dari nilai akurasi sebesar 92%. Sedangkan hasil *wordcloud* dari *smartphone* dengan *medium-high cost* diperoleh kata yang dominannya adalah kirim, cepat, dan bagus dengan nilai akurasi 94%. Berdasarkan hasil grafik *clustering* data *k-means* menunjukkan bahwa angka penjualan 0 sampai 1000 memperoleh skor *rating* bintang tertinggi, sementara itu penjualan dengan skor *rating* bintang terendah yang memiliki angka penjualan 1500 sampai 2000 ke atas.

Penelitian Muktafin *et al* (2020) menggunakan algoritma KNN dan TF-IDF dengan pendekatan NLP untuk mengklasifikasikan ulasan produk “hijab instan” ke dalam 2 kelas (positif dan negatif). Klasifikasi dengan pendekatan NLP menghasilkan nilai akurasi sebesar 76,92%, presisi 80,00% dan *recall* 74,07%, sedangkan tanpa pendekatan NLP hanya memperoleh nilai akurasi sebesar 69,23%, presisi 80,00% dan *recall* 64,52%. Dari hasil tersebut dapat diperoleh kata yang sering muncul pada ulasan membahas tentang penilaian pembeli secara umum pada produk. Pada ulasan positif menunjukkan kepuasan pembeli terhadap kualitas, kecepatan pengiriman dan harga barang, sedangkan ulasan negatif menunjukkan kekecewaan pembeli terhadap warna, dan jumlah barang yang dikirim tidak sesuai dengan pesanan pembeli.

Penelitian Limbong *et al* (2021) mengenai Analisis Klasifikasi Sentimen Ulasan Pada E-Commerce Shopee Berbasis *Word Cloud* Dengan Metode *Naive Bayes* Dan *K-Nearest Neighbor* menggunakan data sebesar 500 ulasan. Kemudian ulasan tersebut diklasifikasikan menggunakan metode *Naive Bayes* dan *K-Nearest Neighbor* (KNN) pada aplikasi *orange data mining*. Kemudian hasil data diolah lagi dengan metode *word cloud* untuk mengetahui

topik pembicaraan apa yang sering dibahas pelanggan ada ulasan di Shopee. Hasil klasifikasi yang diperoleh dengan metode *Naive Bayes* memperoleh nilai *accuracy* 0,914, *precision* 0,915, *recall* 0,914 dan *F1 score* 0,916. Sedangkan hasil klasifikasi dengan metode KNN didapatkan nilai *accuracy* 0,928, *precision* 0,929, *recall* 0,928, dan *F1 score* 0,926. Hal ini membuktikan bahwa dalam penelitian ini kinerja metode KNN lebih baik. Kemudian dari data ulasan tersebut diperoleh informasi kata dengan sentimen positif yang paling sering diulas menggunakan *word cloud* terkait kata: gratis, bagus, suka, murah, mudah, dan cepat. Sedangkan informasi sentimen negatif cenderung membahas topik mengenai kata : kecewa, jelek, mahal, bohong, ribet, dan perbaiki.

Pada penelitian yang dilakukan Sihombing *et al* (2021) bertujuan untuk menganalisis sentimen customer produk shopee menggunakan algoritma *naïve bayes classifier*. Data yang digunakan diperoleh dari hasil penerapan teknik *Web Scraping* pada *customer review* dari produk Xiaomi Redmi Note 9 di *website* Shopee Indonesia. Algoritma klasifikasi yang digunakan yaitu *Naïve Bayes Classifier* karena algoritma tersebut dikenal populer dan efektif untuk melakukan klasifikasi data. Penelitian ini juga menggunakan metode *Knowledge Discovery in Text* (KDT) untuk menggali informasi dari sebuah teks. Hasil klasifikasi dengan algoritma *Naïve Bayes* memperoleh nilai akurasi sebesar 85%. Dengan menerapkan teknik analisis sentimen, penelitian ini membuktikan kepada pelaku bisnis dapat mengetahui opini customer yang bisa digunakan sebagai bahan evaluasi dengan melakukan optimalisasi terhadap produk dan layanan yang diberikan. Hasil visualisasi *word cloud* dapat mempermudah user dalam mengamati topik utama dalam sebuah data dengan hasil yang diperoleh menunjukkan sepuluh besar kata yang paling sering muncul mengenai barang, pengiriman, cepat, bagus, aman, produk, hp, pengiriman cepat, semoga, sesuai.

Fahrudin *et al* (2022) melakukan uji penelitian akan analisis sentimen dengan judul Analisis Video Keluhan Pelanggan Menggunakan *Automatic Speech Recognition* dan Analisis Polaritas Sentimen. Hasil penelitian menunjukkan beberapa keluhan pelanggan pada video yang dianimasikan pada lokasi restoran dan minimarket. Nilai *compound* pada video keluhan pelanggan di restoran pada potongan video ke-7 sebesar -0.4747, potongan video ke-10 sebesar -0.8664, dan potongan video ke-11 sebesar -0.6801, sedangkan nilai *compound* pada video keluhan pelanggan di minimarket pada potongan video ke-1 sebesar -0.1027, potongan video ke-2 sebesar -0.2023, dan potongan video ke-5 sebesar -0.5563. Nilai

compound tersebut merepresentasikan hasil keluhan pelanggan cenderung ke arah sentimen negatif. Hasil visualisasi *word cloud* pada video keluhan pelanggan di restoran menunjukkan pelanggan memang mengeluhkan rasa dan ukuran dari menu makanan yang disajikan, misalnya garam yang terlalu banyak sehingga terasa asin ditunjukkan pada percakapan dengan kata-kata 'let', 'try', 'good', 'think', 'eat', 'slices', 'salt', 'spaghetti', 'chicken', 'taste', 'spicy', 'broccoli', 'ordered', 'expected', dan 'frozen'. Sementara itu pada video keluhan pelanggan di *minimarket* menunjukkan bahwa pelanggan mengeluhkan ketidaksesuaian harga yang ditagihkan dengan label harga *item* produk yang tertera ditunjukkan pada percakapan dengan kata-kata 'see', 'bill', 'sorry', 'discount', 'cooker', 'minute', 'included', 'advertisement', 'amount', 'wrong', 'mango', 'price', 'incorrect', dan 'wait'.

Tabel 2.2 Penelitian Terdahulu Mengenai *Word Cloud* dari Hasil Analisis Sentimen

Penulis	Metode	Objek	Hasil
Sentiya dan Suroyo (2019)	1) Metode pengambilan data: <i>Web Scraping</i> 2) Metode pengolahan data: <i>Text Clustering</i> menggunakan K-Means.	Komentar <i>Followers</i> pada Akun <i>Fanpage</i> Shopee Indonesia	Dari hasil <i>clustering</i> dengan <i>k-means</i> menghasilkan 3 <i>cluster</i> membahas tentang kuis, retweet yang memberikan <i>handphone</i> , dan tebak-tebakan dari <i>fanpage</i> Shopee Indonesia. Sehingga diperoleh topik pembicaraan dari analisis <i>text clustering</i> pada komentar yang mengerucut pada hadiah, <i>iphone</i> , dan <i>vote</i> .
Rahman dan Suroyo (2019)	1) Metode pengambilan data: <i>Web Scraping</i> 2) Metode pengolahan data: Algoritma Teks <i>Clustering</i> Berbasis <i>Python</i>	Teks komentar, angka penjualan dan skor rating bintang pada <i>e-commerce</i> .	Hasil penelitian ini memperoleh kata yang sering muncul adalah barang, mantap, cepat, kirim dan bagus dengan presentase dari nilai akurasi sebesar 92%. Sedangkan hasil <i>wordcloud</i> dari <i>smartphone</i> dengan <i>medium-high cost</i> diperoleh kata yang dominannya adalah kirim, cepat, dan bagus dengan nilai akurasi 94%.

Penulis	Metode	Objek	Hasil
Muktafin <i>et al</i> (2020)	1) Metode pengambilan data: <i>Web Scraping</i> 2) Metode pengolahan data: <i>word cloud</i> dengan algoritma KNN dan TF-IDF dengan pendekatan NLP	Ulasan produk “hijab instan”.	Dari hasil klasifikasi dengan pendekatan NLP memperoleh nilai akurasi, presisi dan <i>recall</i> yang lebih besar dibandingkan klasifikasi tanpa pendekatan NLP. Kemudian hasil tersebut diperoleh kata yang sering muncul pada ulasan membahas tentang penilaian pembeli secara umum pada produk. Pada ulasan positif menunjukkan kepuasan pembeli terhadap kualitas, kecepatan pengiriman dan harga barang, sedangkan ulasan negatif menunjukkan kekecewaan pembeli terhadap warna, dan jumlah barang yang dikirim tidak sesuai dengan pesanan pembeli.
Limbong <i>et al</i> (2021)	1) Metode pengambilan data: <i>Web Scraping</i> 2) Metode pengolahan data: <i>Word Cloud</i> Dengan Metode <i>Naive Bayes</i> Dan <i>K-Nearest Neighbor</i>	Ulasan pelanggan pada <i>e-commerce</i> Shopee.	Hasil penelitian ini diperoleh nilai klasifikasi dengan metode KNN memiliki kinerja lebih baik dibandingkan metode <i>Naive Bayes</i> . Kemudian dari data ulasan tersebut diperoleh informasi kata dengan sentimen positif yang paling sering diulas menggunakan <i>word cloud</i> terkait kata: gratis, bagus, suka, murah, mudah, dan cepat. Sedangkan informasi sentimen negatif cenderung membahas topik mengenai kata : kecewa, jelek, mahal, bohong, ribet, dan perbaiki.
Sihombing <i>et al</i> (2021)	1) Metode pengambilan	<i>Customer review</i> dari produk	Hasil klasifikasi menggunakan algoritma

Penulis	Metode	Objek	Hasil
	data: <i>Web Scraping</i> 2) Metode pengolahan data: Analisis sentimen menggunakan <i>naïve bayes classifier</i> dan <i>word cloud</i>	Xiaomi Redmi Note 9 pada website Shopee Indonesia.	<i>Naïve Bayes</i> didapatkan nilai akurasi nya sebesar 85%. Kemudian dilanjutkan visualisasi <i>customer review</i> dengan <i>word cloud</i> yang menunjukkan sepuluh besar kata yang paling sering muncul membahas tentang barang, pengiriman, cepat, bagus, aman, produk, hp, pengiriman cepat, semoga, sesuai.
Fahrudin <i>et al</i> (2022)	1) Metode pengambilan data: <i>Web Scraping</i> 2) Metode pengolahan data: Automatic Speech Recognition dan Analisis Polaritas Sentimen dengan <i>word cloud</i>	Keluhan pelanggan pada video yang dianimasikan bertempat di restoran dan minimarket.	Nilai <i>compound</i> dari kedua video keluhan pelanggan yang berlokasi di restoran dan <i>minimarket</i> ini merepresentasikan keluhan yang mengarah ke sentimen negatif. Hasil visualisasi <i>word cloud</i> pada video keluhan pelanggan di restoran menunjukkan pelanggan memang mengeluhkan rasa dan ukuran dari menu makanan yang disajikan, misalnya garam yang terlalu banyak sehingga terasa asin ditunjukkan pada percakapan dengan kata-kata ' <i>let</i> ', ' <i>try</i> ', ' <i>good</i> ', ' <i>think</i> ', ' <i>eat</i> ', ' <i>slices</i> ', ' <i>salt</i> ', ' <i>spaghetti</i> ', ' <i>chicken</i> ', ' <i>taste</i> ', ' <i>spicy</i> ', ' <i>broccoli</i> ', ' <i>ordered</i> ', ' <i>expected</i> ', dan ' <i>frozen</i> '. Sementara itu pada video keluhan pelanggan di <i>minimarket</i> menunjukkan bahwa pelanggan mengeluhkan ketidaksesuaian harga yang ditagihkan dengan label harga <i>item</i> produk yang tertera

Penulis	Metode	Objek	Hasil
			ditunjukkan pada percakapan dengan kata-kata 'see', 'bill', 'sorry', 'discount', 'cooker', 'minute', 'included', 'advertisement', 'amount', 'wrong', 'mango', 'price', 'incorrect', dan 'wait'.

Dari hasil penjabaran penelitian terdahulu yang memiliki topik dan metode yang serupa dengan penelitian tugas akhir yang diambil, dapat disimpulkan bahwa sebagian besar penelitian terdahulu menggunakan analisis sentimen untuk melakukan pelabelan data pada hasil ulasan pengguna menggunakan metode *Lexicon Based*, sebagian besar menggunakan kamus *SentiStrength* ID. Dari hasil pelabelan dengan *SentiStrength* menghasilkan nilai polaritas yang berbeda-beda dari setiap penelitian terdahulu yang digunakan tergantung dengan konteks yang diangkat pada ulasan yang diberikan pengguna, tetapi seluruh penelitian tetap mencapai tujuan digunakannya analisis sentimen agar dapat membagi data ke dalam kelas sentimen positif, negatif, dan netral.

Pada hasil analisa penelitian terdahulu yang memiliki topik serupa yaitu memvisualisasi hasil ulasan pelanggan ke bentuk *word cloud* yang sangat membantu dalam mengetahui pokok pembahasan yang dominan pada ulasan pelanggan, sehingga dari hasil tersebut sebagian besar penelitian terdahulu menggunakan *word cloud* agar dapat memberikan kesimpulan dan rekomendasi untuk tindakan selanjutnya berupa solusi atas permasalahan yang dihadapi pengguna. Berdasarkan hasil tersebut maka dilakukan penelitian tugas akhir ini melakukan visualisasikan hasil dari analisis sentimen *online customer review* yang terbagi atas sentimen positif, negatif, dan netral untuk mengetahui pokok pembahasan yang paling dominan disetiap sentimennya sehingga dapat diperoleh kelemahan dan kelebihan dari performansi penjualan toko dengan menggunakan *word cloud* pada *orange data mining*. Kemudian dari hasil tersebut dapat diperoleh pokok permasalahan (kelemahan) dari *word cloud* sentimen negatif, sehingga dapat diberikan rekomendasi yang tepat untuk meningkatkan performansi penjualan untuk setiap tokonya.

2.2 Kajian Deduktif

Kajian deduktif berisikan landasan teori dari beberapa ahli yang memiliki fungsi sebagai kerangka teori/ dasar teori dan tersusun secara sistematis untuk membantu menyelesaikan penelitian yang dilakukan oleh peneliti. Adapun beberapa landasan teori yang digunakan peneliti dalam melaksanakan penelitian yaitu:

2.2.1 *E-marketplace*

E-marketplace merupakan sebuah tempat berinteraksinya penjual dan pembeli dengan melakukan komunikasi terkait produk yang diperjualkan baik dari harga, spesifikasi produk hingga keduanya sampai ke tahap transaksi produk melalui alat komunikasi elektronik pada sistem informasi yang didalamnya terdapat organisasi-organisasi. *E-marketplace* mengidentifikasi struktur sosial, konsep ekonomi pasar, dan penggunaan teknologi dalam memberikan peluang untuk berbisnis dan melakukan transaksi melalui saluran elektronik, umumnya menggunakan suatu platform yang berbasis internet (Rahmadi,2016).

E-marketplace memiliki kegunaan dalam memperluas jaringan dan meningkatkan penjualan terutama pada usaha kecil maupun mandiri yang memiliki modal usaha menengah ke bawah ataupun wirausaha yang baru merintis usahanya. Sehingga *e-marketplace* menjadi pilihan yang sangat tepat sebagai sarana dalam mengenalkan produk baru hingga menjaga komunikasi dengan pelanggan maupun meningkatkan hubungan relasi bisnis, karena biaya yang dibutuhkan tidak berjumlah besar layaknya pasar yang memiliki berbentuk fisiknya. *E-marketplace* tidak terhalang oleh waktu dan kondisi geografis sehingga dapat diakses kapanpun dan dimanapun serta dapat meningkatkan penjualan dan pengenalan produk dalam lingkup yang lebih luas (Raharjo *et al*, 2017).

2.2.2 *Online Customer Review dan Rating*

Definisi *Online Customer Review* menurut (Thakur, 2018) mencakup pengalaman pembeli tentang pelayanan yang diberikan oleh penjual dari segala aspek di *platform retail online* maupun *marketplace*. *Online rating dan review* telah menjadi fitur standar untuk perusahaan *e-commerce* serta perlu adanya deskripsi kegunaan produk dan informasi kepada calon pembelinya (Kang & Namkung, 2019).

Online consumer review mempunyai pengaruh yang signifikan kepada perilaku konsumen dan telah menjadi sumber informasi penting bagi konsumen, dengan persentase pengaruh sebanyak 20-50% dari semua *online purchase decision* (Thakur, 2018). Semakin banyak *online intermediary platforms* seperti *e-commerce* untuk menaruh *reviews* yang ditulis oleh pembeli atau konsumen *online* yang tertuang pada halaman deskripsi produk atau servis, hal tersebut memiliki tujuan agar pembeli dapat memperoleh informasi tambahan mengenai produk atau servis tersebut, sehingga pembeli membuat opini tentang produk yang diinginkan dan semakin menumbuhkan minat beli dari pembeli. Hal ini juga berguna bagi penjual untuk mendapatkan data konsumen dan juga *feedback* mengenai produk & servis yang mereka sediakan kepada konsumen (Xie, Chen, & Wu, 2016).

Rating adalah bagian dari *review* yang menggunakan bentuk simbol bintang daripada bentuk teks dalam mengekspresikan pendapat dari pelanggan. *Rating* dapat diartikan sebagai penilaian dari pengguna pada preferensi suatu produk terhadap pengalaman mereka mengacu pada keadaan psikologis dan emosional yang mereka jalani saat berinteraksi dengan produk virtual dalam lingkungan di mediasi (Li, N. and Zhang, P. 2002). Menurut penelitian Moe dan Schweidel, (2012) menghubungkan antara *rating* terhadap tingkat pengambilan keputusan pelanggan. Mereka menemukan bahwa pengaruh penilaian pelanggan terhadap *rating* sebelum memutuskan membeli sesuatu tergantung pada seberapa sering *rating* atau penilaian dilakukan oleh pelanggan terhadap suatu produk dan menjadi representasi dari opini konsumen dengan skala yang spesifik (Lakkermaier et al, 2013).

2.2.3 Web Scraping

Web scraping adalah teknik yang digunakan untuk mendapat informasi suatu *website* yang bertujuan mencari dan mengumpulkan informasi dalam *website* yang baru. *Web scraping* akan melakukan ekstraksi data dengan pengukuran yang bervariasi (Ayani, 2019). *Web Scraping* adalah proses yang melibatkan pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya halaman web dalam bahasa *markup* seperti HTML (Setiawan et al, 2020). Kegunaan *web scraping* dapat juga memproses dan mengumpulkan data dengan jumlah besar seperti pengambilan data pada suatu produk dari beberapa *e-marketplace* (Levi, 2020). Untuk melakukan pengambilan data tersebut dapat memanfaatkan fungsi dari Python. Sehingga

dengan memanfaatkan teknik *web scraping* dapat mengumpulkan informasi barang yang dicari dari berbagai toko *online* maupun *e-marketplace*.

2.2.4 Analisis Sentimen

Analisis sentimen adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen positif maupun sentimen negatif (Nasukawa & Yi, 2003). Pendapat yang sama juga disampaikan oleh (Coletta *et al*, 2014), di mana analisis sentimen adalah proses yang digunakan untuk menentukan opini, emosi dan sikap yang dicerminkan melalui teks, dan biasanya diklasifikasikan menjadi opini negatif dan positif.

2.2.5 Text Mining

Text mining adalah satu langkah dari analisis teks yang dilakukan secara otomatis oleh komputer untuk menggali informasi yang berkualitas dari suatu rangkaian teks yang terangkum dalam sebuah dokumen (Han & Kamber, 2006). Tahapan dalam *text mining* adalah *case folding*, *tokenizing*, *filtering*, dan *stemming*, tahapan ini disebut tahap *text preprocessing*. Tahap *text preprocessing* adalah tahapan dimana aplikasi melakukan seleksi data yang akan diproses pada setiap dokumen.

1. Case Folding

Case folding merupakan proses mengubah semua huruf dalam suatu dokumen atau kalimat menjadi huruf kecil (Valatehan, *at al*, 2016). *Case folding* digunakan untuk mempermudah pencarian. Tidak semua data konsisten dalam penggunaan huruf kapital.

2. Tokenisasi

Tokenisasi merupakan proses pemisahan suatu rangkaian karakter berdasarkan karakter spasi, dan mungkin pada waktu yang bersamaan dilakukan juga proses penghapusan karakter tertentu, seperti tanda baca (Amin, 2012).

3. Filtering

Filtering / eliminasi *stopwords* memiliki banyak keuntungan, yaitu akan mengurangi space pada tabel *term index* hingga 40% atau lebih (Baeza & Neto, 1999). Proses

stopword removal merupakan proses penghapusan *term* yang tidak memiliki arti atau tidak relevan (Amin, 2012).

4. *Stemming*

Kata-kata yang sudah diubah menjadi huruf kecil perlu dilakukan pengecekan. *Stemming* digunakan untuk menyeragamkan kata sehingga mengurangi daftar kata yang ada pada data latih (Manning, 2008).

2.2.6 *SentiStrength ID*

Menurut Wahid & Azhari (2016) *SentiStrength* merupakan algoritma klasifikasi yang memanfaatkan pendekatan leksikon atau kamus dengan menggunakan kamus emosikon dan kamus ungkapan yang diberi bobot oleh manusia. Kamus sentimen ini didapatkan dari hasil translasi kamus sentimen berbahasa Inggris yang telah mengalami penambahan dan pengurangan kata berdasarkan dari hasil pengamatan dalam proses pengembangan sistem ini. Penelitian ini melibatkan 3 responden ahli, mahasiswa S2 Linguistik Universitas Gadjah Mada untuk memberikan bobot kekuatan sentimen. Kamus sentimen ini berisi kumpulan kata yang telah diberi bobot dengan kekuatan sentimen 1 (tidak memiliki sentimen positif) s.d. 5 (memiliki sentimen positif yang sangat kuat), dan -1 (tidak memiliki sentimen negatif) s.d. -5 (memiliki sentimen negatif yang sangat kuat) (Wahid & Azhari, 2016). Skor akhir pada sebuah kalimat ditentukan dari skor positif tertinggi dan skor negatif tertinggi dari kata-kata penyusunnya. Skor akhir tersebut dapat berubah jika memenuhi aturan-aturan tambahan pada ekstraksi fitur. Pembobotan term selain menggunakan kamus sentiment juga dapat menggunakan kamus emosikon dan kamus ungkapan yang masing-masing kata dan karakter pada kamus diberi bobot tertentu. Kemudian keputusan akhir sentimen didasarkan pada aturan (Aziz & Fauziah, 2022):

- Jika nilai positif > nilai negatif maka termasuk sentimen positif.
- Jika nilai positif < nilai negatif maka termasuk sentimen negatif.
- Jika nilai positif = nilai negatif maka termasuk sentimen netral.

2.2.7 *Word Cloud*

Word cloud berguna untuk memvisualisasikan data dalam bentuk teks secara visual, tujuannya untuk memudahkan dalam memberikan pemahaman informasi kepada pengguna (Pradana, 2020). Visualisasi *word cloud* ini menggunakan frekuensi kata dari teks yang direpresentasikan dalam bentuk ukuran kata. Semakin besar ukuran kata yang divisualisasikan, maka semakin tinggi frekuensi kata tersebut. Sebaliknya, semakin kecil ukuran kata yang divisualisasikan, maka semakin rendah frekuensi kata tersebut. *Word cloud* (awan kata) adalah kumpulan kata-kata yang paling banyak muncul dalam data teks yang dianalisis. Kata-kata tersebut terkumpul seperti sebuah gumpalan awan yang berisi kata-kata. Intensitas banyaknya kata yang digunakan, ditunjukkan dengan ukuran huruf pada kata. Semakin besar huruf dari kata yang terdapat di awan kata menunjukkan semakin sering kata tersebut muncul. Tampilan awan kata ini lebih menarik serta cepat untuk menemukan kata-kata yang sering muncul, akan tetapi awan kata memiliki kekurangan yaitu tidak dapat menunjukkan frekuensi kata-kata muncul dalam suatu teks yang dianalisis (Adiyana & Hakim, 2015).

2.2.8 Python

Python adalah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif (Cahyono, 2018). Menurut Syahrudin dan Kurniawan (2018) input dan output adalah proses dasar dalam sebuah pemrograman, tanpa adanya input dan output mustahil program dapat berjalan dan berinteraksi dengan user, dalam mempelajari bahasa pemrograman yang harus kita pahami yaitu struktur pemrograman bukan menghafal coding, karena setiap adanya pembaruan/update terkadang terdapat syntax/coding yang berubah.

2.2.9 Asosiasi Teks

Asosiasi teks dilakukan dengan pendekatan untuk perhitungan dari nilai korelasi. Nilai korelasi dinyatakan dengan mengetahui hubungan antara dua atau lebih variabel kuantitatif, namun pada asosiasi teks nilai korelasi lebih diartikan sebagai keeratan hubungan yang diperoleh antar dua atau lebih variabel kualitatif (Ulwan, 2016). Korelasi bertujuan untuk

menemukan tingkat hubungan antara variabel bebas (X) dan variabel bebas (Y), dalam ketentuan data memiliki syarat-syarat tertentu (Fadlisyah, 2014).

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\{n \sum x_i^2 - (\sum x_i)^2\} \{n \sum y_i^2 - (\sum y_i)^2\}}}$$

(2.1)

dengan

r = nilai korelasi antara variabel x dan variabel y

n = banyaknya pasangan data x dan y

$\sum x_i$ = jumlah nilai pada variabel x

$i = 1, 2, 3, \dots, n$

$\sum y_i$ = jumlah nilai pada variabel y

$\sum x_i^2$ = kuadrat dari total nilai variabel x

$\sum y_i^2$ = kuadrat dari total nilai variabel y

$\sum x_i \sum y_i$ = jumlah dari hasil perkalian antara nilai variabel x dan variabel y

Dalam perhitungan asosiasi teks, pertama-tama data teks ditransformasikan ke dalam *document term matrix* (dtm). Adapun simulasi perhitungan dilakukan pada enam data berikut.

kata1

kata1 kata2

kata1 kata2 kata3

kata1 kata2 kata3 kata4

kata1 kata2 kata3 kata4 kata5

kata1 kata2 kata3 kata4 kata5 kata6

Kemudian ke 6 kata tersebut diubah dalam *document term matrix*.

Docs	kata1	kata2	kata3	kata4	kata5	kata6
1	1	0	0	0	0	0
2	1	1	0	0	0	0
3	1	1	1	0	0	0
4	1	1	1	1	0	0
5	1	1	1	1	1	0
6	1	1	1	1	1	1

Setelah diperoleh nilai *document term matrix*, selanjutnya dilakukan perhitungan nilai asosiasi. Nilai asosiasi diperoleh dengan menghitung rumus korelasi seperti pada simulasi kata 3 dan kata 5 berikut.

Docs	Kata3	Kata5	Kata3^2	Kata5^2	kata3*kata5
1	0	0	0	0	0
2	0	0	0	0	0
3	1	0	1	0	0
4	1	0	1	0	0
5	1	1	1	1	1
6	1	1	1	1	1
Total	4	2	4	2	2

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\{n \sum x_i^2 - (\sum x_i)^2\} \{n \sum y_i^2 - (\sum y_i)^2\}}}$$

$$r = \frac{(6 \times 2) - (4 \times 2)}{\sqrt{\{(6 \times 4) - (4^2)\} \{(6 \times 2) - (2^2)\}}}$$

$$r = \frac{4}{\sqrt{64}} = \frac{4}{8} = 0,5$$

Jadi, nilai korelasi kata3 dan kata5 sebesar 0,5. Hal ini menunjukkan bahwa besarnya asosiasi atau hubungan kata3 dengan kata5 sebesar 0,5 atau 50%.

2.2.10 Fishbone Diagram

Fishbone Diagram dikenal juga dengan sebutan *the Cause and Effect Diagram* atau Ishikawa Diagram diperkenalkan pertama kali oleh pencetusnya yaitu Kaoru Ishikawa (1915- 1989), yang merupakan warga negara Jepang. Diagram sebab-akibat adalah sebuah teknik grafis yang digunakan untuk mengurutkan dan menghubungkan interaksi antara faktor-faktor yang berpengaruh dalam suatu proses. Diagram ini berguna untuk menganalisis, mengidentifikasi dan mengorganisir sebab-sebab yang mungkin muncul dengan memisahkan akar penyebabnya dan menyebutkan beberapa permasalahan yang muncul dalam menentukan karakteristik kualitas *output* kerja. Efek ini dapat bernilai "baik" dan dapat bernilai "buruk". Jadi dengan diketahui sebab dari efek yang terjadi, diharapkan hasil dari proses produksi bisa diperbaiki dengan mengubah faktor terkontrol dari suatu proses. Diagram *fishbone* memfokuskan pada penekanan masalah atau gejala yang merupakan akar penyebab

masalah(Fauziah, 2009). Sehingga dengan menggunakan Diagram *fishbone* peneliti dapat memberikan kontribusinya dengan memberikan rekomendasi perbaikan akan permasalahan yang muncul dari hasil penggunaan teknik ini.



BAB III

METODOLOGI PENELITIAN

3.1 Subjek Penelitian dan Objek Penelitian

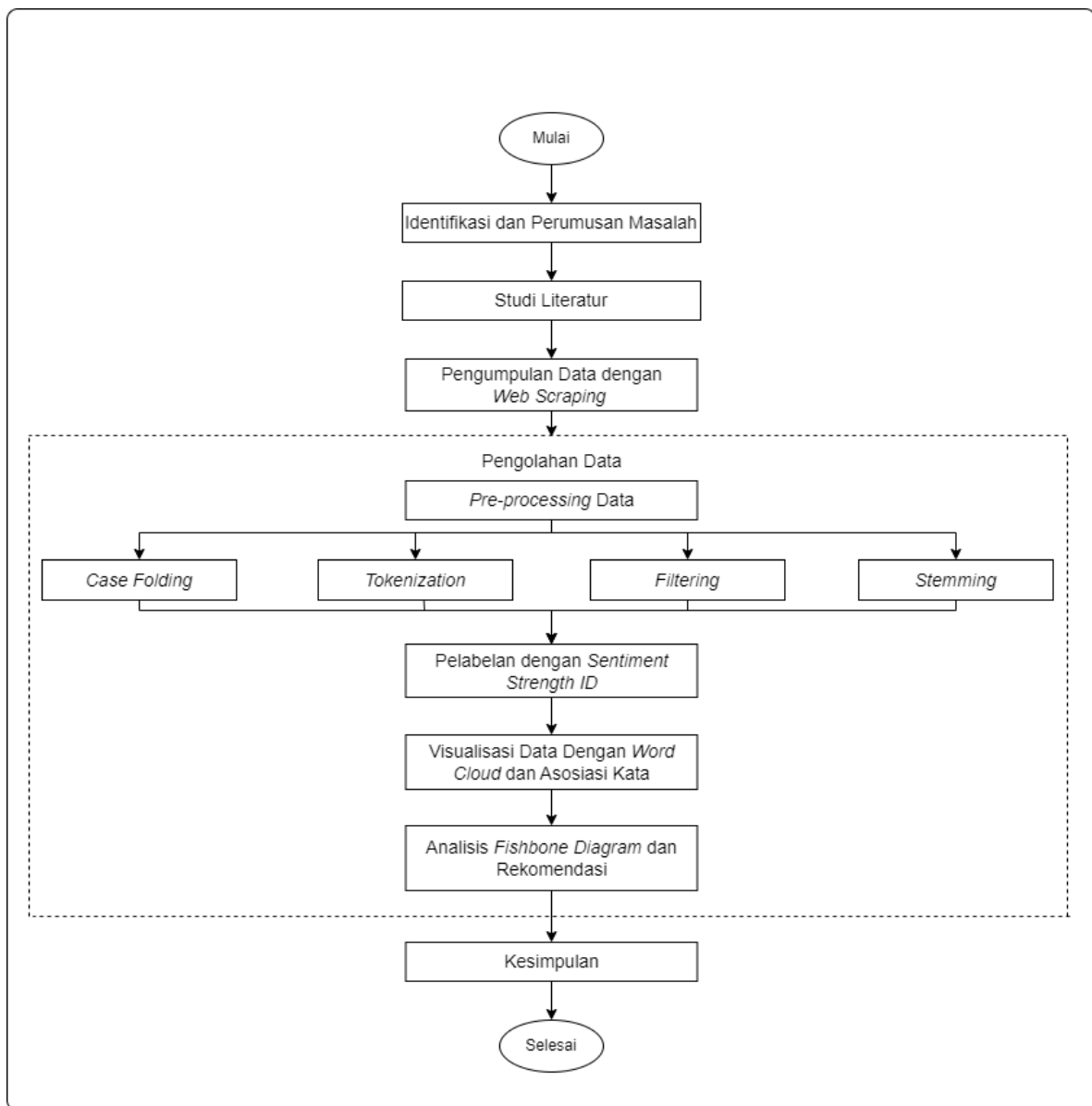
Subjek pada penelitian ini adalah pelanggan yang telah melakukan transaksi jual beli *online* pada *e-marketplace* Shopee terutama di 3 toko *online* elektronik Appelhouse.store, Jordan Corner dan Iphonespot. Dalam hal ini, subjek penelitian dijadikan sebagai subjek yang memberikan penilaian dari produk yang telah dibeli dan akan menjadi objek dari penelitian ini. Sehingga dapat diketahui objek penelitian ini adalah hasil penilaian berupa *online customer review*. Pada objek penelitian *online customer review* dan *rating* dilakukan pengambilan data pada tanggal 20 April hingga 19 Juni 2022 berdasarkan ulasan yang diberikan pelanggan terhadap produk yang telah diterima hasil transaksi jual beli *online* di 3 toko *online* elektronik pada *e-marketplace* Shopee.

3.2 Sumber Data Penelitian

Sumber data yang digunakan pada penelitian ini adalah data sekunder dimana data – datanya diperoleh dari penelitian – penelitian terdahulu maupun informasi yang terdapat di internet. Pada penelitian ini, data sekunder diambil dari hasil *online customer review* dari pelanggan terhadap suatu produk tertentu di 3 toko *online* elektronik Appelhouse.store, Jordan Corner dan Iphonespot pada *e-marketplace* Shopee serta beberapa jurnal ilmiah dan artikel yang bersangkutan dengan penelitian sebagai referensi peneliti.

3.3 Alur Penelitian

Metodologi penelitian bertujuan untuk menyelesaikan masalah yang ada secara terstruktur. Berikut pembagian dari alur penelitian ditunjukkan pada gambar 3.1.



Gambar 3.1 Alur Penelitian

3.3.1 Identifikasi Masalah dan Studi Literatur

Pada tahap ini dilakukan observasi terhadap permasalahan yang dapat diamati dan diidentifikasi sehingga permasalahan tersebut dapat diangkat agar dapat diselesaikan dengan metode ataupun analisa terbaru yang diperkuat dengan dasar-dasar ilmu pengetahuan dari studi literatur yang digunakan peneliti agar dapat membuktikan bahwa topik penelitian yang dipilih penulis telah memenuhi syarat dan kriteria yang telah dijelaskan.

3.3.2 Pengumpulan Data

Melakukan penelusuran terhadap kategori produk *smartphone* di Daerah Istimewa Yogyakarta dengan menggunakan filter produk terlaris, kemudian menentukan dan melakukan analisa kepada 3 toko yang memiliki performa terbaik berdasarkan jumlah pengikut, jumlah penjualan, rating toko, performa chat dan waktu bergabung toko dengan Shopee. Setelah itu diperoleh 3 toko *online smartphone* terbaik di Daerah Istimewa Yogyakarta yaitu Appelhouse.store, Jordan Corner dan Iphonespot. Kemudian dilakukan proses pengambilan dan pengumpulan data *review* pelanggan menggunakan metode *web scraping* dengan bahasa pemrograman Python 3.9 pada 3 toko *online* tersebut.

3.3.3 Pengolahan Data

Setelah data dikumpulkan, kemudian peneliti melakukan perapihan dan pengolahan data sederhana agar data siap untuk dilakukan analisis. Terdapat 5 aktivitas dalam mengolah data dari hasil pengumpulan data, yaitu:

- 1) Pada aktivitas pertama dilakukan *pre-processing* data untuk melakukan seleksi data yang akan diproses pada setiap dokumen dengan beberapa tahapan yaitu *case folding, tokenizing, filtering, dan stemming*.
- 2) Kemudian data akan dilakukan pelabelan data menggunakan kamus *Senti Strength ID* pada Python agar dapat dapat mengklasifikasikan hasil *online customer review* ke dalam aspek positif, netral dan negatif
- 3) Hasil analisis sentimen positif, negatif, dan netral dari *online customer review* divisualisasikan dengan menggunakan *word cloud* agar dapat diketahui kelebihan (sentimen positif dan netral) dan kelemahan (sentimen negatif) setiap tokonya. Kemudian dilakukan asosiasi kata berdasarkan hasil *word cloud* untuk mengetahui faktor kata apa saja yang memicu sering munculnya kata pada sentimen negatif
- 4) Hasil asosiasi kata tersebut kemudian dijabarkan dengan diagram *fishbone* agar dapat mengetahui akar permasalahan yang dihadapi setiap toko dan dapat dilakukan perencanaan perbaikan pada setiap performansi toko, sehingga dapat diberikan rekomendasi yang tepat untuk menunjang performansi ketiga toko menjadi lebih baik

3.3.4 Kesimpulan

Di tahap ini memuat kesimpulan berdasarkan hasil analisis pada penelitian yang telah dilakukan. Hasil kesimpulan diharapkan dapat membantu seller di *e-marketplace* Shopee dalam perencanaan dan perancangan strategi agar dapat lebih memahami minat membeli dari pelanggan, sehingga dapat mendukung perkembangan bisnis *online* di Indonesia khususnya Daerah Istimewa Yogyakarta.



BAB IV

PENGUMPULAN DAN PENGOLAHAN DATA

4.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari hasil *review* pengguna terhadap penjualan produk barang elektronik berupa *smartphone* pada 3 toko *online* elektronik Appelhouse.store, Jordan Corner dan Iphonespot pada *e-marketplace* Shopee. Untuk proses pengumpulan data dilakukan dengan *web scraping* menggunakan bahasa pemrograman Python melalui *jupyter notebook*.

Dikarenakan keterbatasan akses data yang dapat dikumpulkan pada aplikasi Shopee dengan menggunakan Python, sehingga sampel data yang dapat diperoleh terbatas yaitu sebanyak 324 data penjualan terakhir untuk setiap toko nya. Berikut ini adalah contoh hasil pengumpulan data menggunakan *web scraping* pada toko *online* Appelhouse.store.

Tabel 4.1 Contoh Hasil Pengumpulan Data pada Appelhouse.store.

No	Nama Pengguna	Produk	Review	Rating	Waktu Transaksi
1	eminaksuryadin	iPhone 11 64GB/128GB/256G B Bekas Original 100% NORMAL MULUS FULLSET	Sumpah bagus banget di pilihankan yg terbaik ngk nyesel beli i phone di toko nh di kash BH yg 89% thanks you	5	2022-06-19
2	n*****9	AKSESORIS	Bagus pokonya , adminnya juga fast respon thanks yaa sukses sllu	5	2022-06-18
3	listyadhisty__	iPhone X 64GB/256GB Bekas Second Original 100% NORMAL MULUS FULLSET	Baru pertama kali beli hp online, puass banget ga request bh karena pasti dikirim yg terbaikk!! Top lah appelhouse Cuma masalah di ekspedisi pengiriman aja si, selebihnya okee bgt	5	2022-06-17
4	ajengmaulidiaa	iPhone 11 Pro Max 64GB/256GB/512G B Bekas Second Original 100% NORMAL MULUS FULLSET	Kondisi semua baik, tp minta yg spacegrey d kirim yg midnight green. Dpt yg HB 93%	4	2022-06-16
5	nauva05	iPhone 11 64GB/128GB/256G B Bekas Original	Layarnya baret dan speaker kayak kemasukan air. Untuk yg laun itsokay	3	2022-06-15

No	Nama Pengguna	Produk	Review	Rating	Waktu Transaksi
		100% NORMAL MULUS FULLSET			

Berikut ini adalah contoh hasil pengumpulan data menggunakan *web scraping* pada toko *online* Iphonespot.

Tabel 4.2 Contoh Hasil Pengumpulan Data pada Iphonespot

No	Nama Pengguna	Produk	Review	Rating	Waktu Transaksi
1	s*****u	iPhone 11 64GB SECOND ORIGINAL	Fitur Terbaik:baterai 100% Sepadan dengan Harga:masih bagus banget Ada bagian iphone yang lecet- lecet tapi keseluruhan bagus. Admin responsif juga. Bisa tanya admin kalau ada warna yang sold bisa di-update. Bisa tanya dan request BH juga (meski kudu nambah biaya sih). Pakaging kayu aman banget, nambah biaya untuk itu jadi aman.	5	2022-06-15
2	eyra_22	FAST CHARGING SET ORIGINAL	Fitur Terbaik:fungsi normal fast charging gak tau beneran ori atau gak tpi bagus buat cas cepat	4	2022-06-14
3	eyra_22	iPhone 11 64GB SECOND ORIGINAL	Fitur Terbaik:silent kamera on, face id normal, imei sama, layar baik belum check di 3utools barang sesuai harga, pemakaian belum 1 bulan bh udah turun padahal pakai carger dari toko yg sama, dapet bh 91%, kirain bisa req beneran ternyata gak	4	2022-06-14
4	asep.hermawan03	iPhone XS 64GB SECOND ORIGINAL	keren bagus barangnya terimakasih next order lagi	5	2022-06-14
5	nauva05	iPhone 11 64GB/128GB/256G B Bekas Original 100% NORMAL MULUS FULLSET	Layarnya baret dan speaker kayak kemasukan air. Untuk yg laun itsokay	3	2022-06-15

Berikut ini adalah contoh hasil pengumpulan data menggunakan *web scraping* pada toko *online* Jordan Corner.

Tabel 4.3 Contoh Hasil Pengumpulan Data pada Jordan Corner

No	Nama Pengguna	Produk	Review	Rating	Waktu Transaksi
1	d*****a	iPhone 11 64 GB Second Original	Alhamdulillah ip nya sangat worth it sekali dipakai 😊 sengaja baru kasih ulasan pas bgt udah seminggu pakai. Semua ok baterai ok bgt tahan seharian dpt bh 99% awet wet blm turun samsek. Admin ramah mau bgt direpotin(ditanya2) xixi bakalan order lg sih next time. Thnx seller & kurir! ❤️	5	2022-06-19
2	nadiaaprinda	Iphone 11 Pro 64 GB SECOND ORIGINAL	Barang nya bagus tidak mengecewakan. Seler ramah . Pokoknya recomended banget toko ini	5	2022-06-18
3	p*****8	IPHONE 12 MINI 128 GB SECOND ORIGINAL	Sepadan dengan Harga:yessss Fitur Terbaik:mantul	5	2022-06-17
4	reza.azifa27	iphone Xs 64 GB Second Original	Kualitas sudah tidak usah di ragukan lagi. Second original mwantuuuuuuulllllll. Adminnya sangat ramah dan mengarahkan. Jordan memang mantep sih. Tengkyuuu min	5	2022-06-17
5	wakhiddonins	IPHONE 11 PRO MAX 256 GB SECOND ORIGINAL	Fitur Terbaik:Pelayanan sangat baik , adminnya ramah balas chatnya fast respon ,detail banget ngirim fotobarangnya Sepadan dengan Harga:sepadan dengan harga yang diberikan puas pokoknya	5	2022-06-16

4.2 Pengolahan Data

Setelah melakukan proses pengumpulan data, kemudian dilakukan pengolahan data sementara dengan mengeliminasi beberapa jenis produk yang tidak relevan dengan produk yang akan diteliti (*smartphone*) dan juga produk yang tidak memiliki ulasan sehingga diperoleh jumlah data yang dapat dilanjutkan untuk tahap pengolahan data selanjutnya yaitu

153 data untuk toko Appelhouse.store, 100 data untuk toko Ifonespot, dan 161 data untuk toko Jordan Corner.

Tabel 4.4 Jumlah Data Yang Akan Diolah

Keterangan	Appelhouse.store	Ifonespot	Jordan Corner
Jumlah Data	153	100	161

Kemudian data tersebut dilakukan proses pengolahan data selanjutnya yaitu *pre-processing* data untuk melakukan seleksi data yang akan diproses pada setiap dokumen dengan beberapa tahapan yaitu *case folding*, *tokenizing*, *filtering*, dan *stemming*.

4.2.1 Pre-Processing data

Kemudian data tersebut dilanjutkan pada tahap *pre-processing* data menggunakan *jupyter notebook* dengan bahasa pemrograman python 3.9 terdiri dari 4 proses yaitu *case folding*, *tokenizing*, *filtering*, dan *stemming*.

4.2.1.1 Case Folding

Tahap ini adalah terjadi proses keseragaman huruf yang berkaitan dengan huruf kapital atau bukan, selain itu dilakukan penghapusan tanda baca `[?|$.|!^_;"')(-+,.]'` dan emoji yang digunakan pengguna untuk memberikan review. Berikut ini hasil sebelum dan sesudah dilakukan *case folding*.

Tabel 4.5 Contoh Hasil *Case Folding*

Sebelum <i>Case Folding</i>	Sesudah <i>Case Folding</i>
Pesan hari ini langsung kirim, mantep respon penjual dan barang bagus mulus	pesan hari ini langsung kirim mantep respon penjual dan barang bagus mulus
Barang original mulus, pengemasan istimewa, admin ramah, sukses selalu appelhouse	barang original mulus pengemasan istimewa admin ramah sukses selalu appelhouse
hp nya tiba-tiba tidak berfungsi dengan baik, baru satu hari nyampe langsung rusak	hp nya tiba-tiba tidak berfungsi dengan baik baru satu hari nyampe langsung rusak

4.2.1.2 Tokenizing

Tokenizing merupakan proses yang dilakukan untuk memotong atau memecah kalimat menjadi beberapa bagian. Berikut ini hasil sebelum dan sesudah dilakukan *tokenizing*.

Tabel 4.6 Contoh Hasil *Tokenizing*

Sebelum Tokenizing	Sesudah Tokenizing
pesan hari ini langsung kirim mantep respon penjual dan barang bagus mulus	'pesan', 'hari', 'ini', 'langsung', 'kirim', 'mantep', 'respon', 'penjual', 'dan', 'barang', 'bagus', 'mulus'
barang original mulus pengemasan istimewa admin ramah sukses selalu appelhouse	'barang', 'original', 'mulus', 'pengemasan', 'istimewa', 'admin', 'ramah', 'sukses', 'selalu', 'appelhouse'
hp nya tibatiba tidak berfungsi dengan baik baru satu hari nyampe langsung rusak	'hp', 'nya', 'tibatiba', 'tidak', 'berfungsi', 'dengan', 'baik', 'baru', 'satu', 'hari', 'nyampe', 'langsung', 'rusak'

4.2.1.3 Filtering

Filtering adalah tahap menghilangkan kata-kata yang muncul dalam jumlah besar tetapi dianggap tidak memiliki makna (*stopwords*). Pada dasarnya, daftar *stopwords* adalah sekumpulan kata yang banyak digunakan dalam berbagai bahasa. Alasan menghapus kata yang terkait dengan penambahan teks adalah karena penggunaannya yang terlalu umum, sehingga pengguna dapat fokus pada kata-kata lain yang jauh lebih penting. Berikut ini hasil sebelum dan sesudah dilakukan *filtering*.

Tabel 4.7 Contoh Hasil *Filtering*

Sebelum Filtering	Sesudah Filtering
'pesan', 'hari', 'ini', 'langsung', 'kirim', 'mantep', 'respon', 'penjual', 'dan', 'barang', 'bagus', 'mulus'	'pesan', 'langsung', 'kirim', 'mantep', 'respon', 'penjual', 'barang', 'bagus', 'mulus'
'barang', 'original', 'mulus', 'pengemasan', 'istimewa', 'admin', 'ramah', 'sukses', 'selalu', 'appelhouse'	'barang', 'original', 'mulus', 'pengemasan', 'istimewa', 'admin', 'ramah', 'sukses', 'appelhouse'
'hp', 'nya', 'tibatiba', 'tidak', 'berfungsi', 'dengan', 'baik', 'baru', 'satu', 'hari', 'nyampe', 'langsung', 'rusak'	'hp', 'tibatiba', 'tidak', 'berfungsi', 'baik', 'baru', 'nyampe', 'langsung', 'rusak'

4.2.1.4 Stemming

Proses *preprocessing* data yang terakhir yaitu proses *stemming* berfungsi untuk menjadikan kata *suffix* menjadi kata dasar sesuai dengan aturan Indonesia yang benar. Berikut ini hasil sebelum dan sesudah dilakukan *stemming*.

Tabel 4.8 Contoh Hasil *Stemming*

Sebelum Stemming	Sesudah Stemming
'pesan', 'langsung', 'kirim', 'mantep', 'respon', 'penjual', 'barang', 'bagus', 'mulus'	'pesan', 'langsung', 'kirim', 'mantap', 'respon', 'jual', 'barang', 'bagus', 'mulus'

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
'barang', 'original', 'mulus', 'pengemasan', 'istimewa', 'admin', 'ramah', 'sukses', 'appelhouse'	'barang', 'original', 'mulus', 'kemas', 'istimewa', 'admin', 'ramah', 'sukses', 'appelhouse'
'hp', 'tibatiba', 'tidak', 'berfungsi', 'baik', 'baru', 'nyampe', 'langsung', 'rusak'	'hp', 'tibatiba', 'tidak', 'fungsi', 'baik', 'baru', 'sampai', 'langsung', 'rusak'

Setelah tahapan *pre-processing data* telah selesai dilakukan, maka data tersebut menjadi lebih terstruktur seperti yang ditunjukkan pada Tabel 4.8. Hal ini dapat mempermudah tahapan selanjutnya untuk melakukan proses klasifikasi.

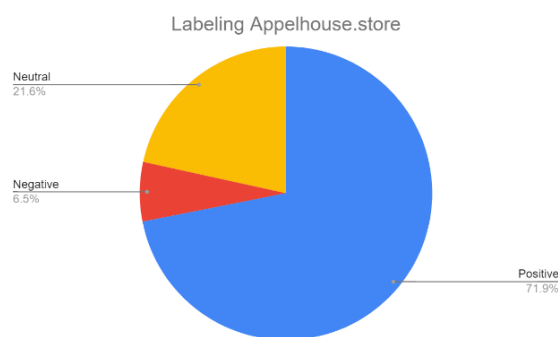
4.2.2 Pelabelan Data Menggunakan *SentiStrength* ID

Pelabelan data dihitung dengan menggunakan kamus *Sentimen Strength* Indonesia untuk mendapatkan nilai *polarity dari review*. Nilai *polarity dari review* <0 , akan dilabeli sebagai sentimen negatif untuk *polarity dari review* $=0$ akan dilabeli sebagai sentimen netral dan untuk *polarity dari review* >0 akan dilabeli sebagai sentimen positif (Aziz & Fauziah, 2022). Adapun hasil pelabelan dari ketiga toko yang diteliti ditunjukkan oleh Tabel 4.9, Tabel 4.10, dan Tabel 4.11.

Berdasarkan hasil pelabelan sentimen yang dilakukan pada toko Appelhouse.store dengan menggunakan kamus *Sentimen Strength* Indonesia, maka didapatkan jumlah pembagian yang diperoleh yaitu label sentimen positif sebanyak 110, label sentimen negatif sebanyak 10 dan label sentimen netral sebanyak 33. Hasil pelabelan sentimen tersebut divisualisasikan dalam grafik *pie* pada Gambar 4.1.

Tabel 4.9 Contoh Hasil *SentiStrength* Toko Appelhouse.Store

<i>Review</i>	<i>Polarity dari review</i>	Label
pesan langsung kirim mantep respon jual barang bagus mulus	0.7184	<i>Positive</i>
barang original mulus emas istimewa admin ramah sukses appelhouse	0.9517	<i>Positive</i>
sip	0	<i>Neutral</i>
pesan warna hitam dateng warna merah	0	<i>Neutral</i>
hp tibatiba tidak fungsi baik baru sampe langsung rusak	-0.7184	<i>Negative</i>

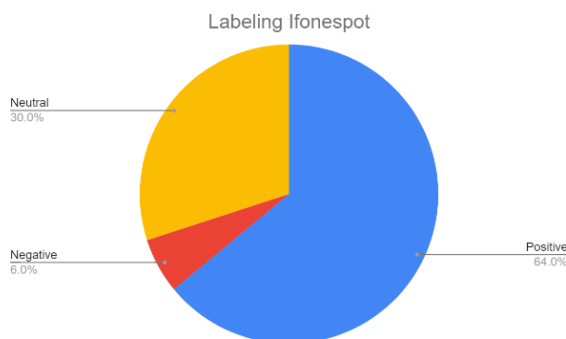


Gambar 4.1 Hasil Pelabelan Sentimen Appelhouse.Store

Sementara itu hasil pelabelan sentimen yang dilakukan pada toko Iphonespot dengan menggunakan kamus *Sentimen Strength* Indonesia, diperoleh jumlah pembagian yang diperoleh yaitu label sentimen positif sebanyak 64, label sentimen negatif sebanyak 6 dan label sentimen netral sebanyak 30. Hasil pelabelan sentimen tersebut divisualisasikan dalam grafik *pie* pada Gambar 4.2.

Tabel 4.10 Contoh Hasil SentiStrength Toko Iphonespot

Review	Polarity	Label
puas belanja unit dua beli sini	0.6124	<i>Positive</i>
mantap banget jogja makassar aman kendali	0.7184	<i>Positive</i>
keren admin ramah baik packing aman bagus kualitas oke barang original	0.9682	<i>Positive</i>
iphone 11 64gb purple dapat bh 96 1 bulan udah turun 87 dapat dari japan tidak bisa silent camera	-0.7184	<i>Negative</i>
mulus	0	<i>Neutral</i>

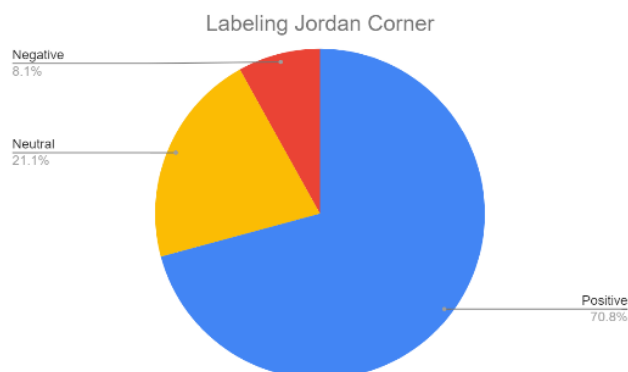


Gambar 4.2 Hasil Pelabelan Sentimen Iphonespot

Sedangkan hasil pelabelan sentimen yang dilakukan pada toko Jordan Corner dengan menggunakan *Sentimen Strength* Indonesia, diperoleh jumlah pembagian yang diperoleh yaitu label sentimen positif sebanyak 114, label sentimen negatif sebanyak 13 dan label sentimen netral sebanyak 34. Hasil pelabelan sentimen tersebut divisualisasikan dalam grafik *pie* pada Gambar 4.3.

Tabel 4.11 Contoh Hasil *SentiStrength* Toko Jordan Corner

Review	Polarity	Label
barang waktu estimasi packing respon admin	0	<i>Neutral</i>
udah nambah uang tempered glass case tidak kasih suruh ambil store wonosobo udah ke store tidak kasih karyawan store satu ketus banget tidak hargai customer	-0.7184	<i>Negative</i>
fitur baik mantap oke	0.6124	<i>Positive</i>
bagus banget sesuai request	0.875	<i>Positive</i>
jual ramah barang original packing aman rapih	0.9001	<i>Positive</i>



Gambar 4.3 Hasil Pelabelan Sentimen Jordan Corner

4.2.3 Word Cloud

Dari hasil pelabelan data pada ketiga toko, maka dapat dilanjutkan pengolahan data dengan memvisualisasikan hasil *review* sentimen positif, negatif dan netral pada toko Appelhouse.store, Iphonespot, dan Jordan Corner dengan menggunakan *word cloud* dengan aplikasi *orang data mining*, sehingga diperoleh hasilnya sebagai berikut.

4.2.3.1 Sentimen Positif



Gambar 4.4 *Word Cloud* Sentimen Positif pada Toko Appelhouse.store

Dari hasil *word cloud* untuk sentimen positif pada toko Appelhouse.store dapat diketahui bahwa 5 kata yang paling sering muncul adalah bagus, barang, mulus, aman, dan bh (*battery health*) yang berarti kualitas pelayanan dan kualitas produk toko Appelhouse.store sudah sangat baik terlihat dari produk yang diperoleh masih aman (tidak ada kecacatan produk) sehingga kualitas barang yang diperoleh pelanggan produk masih mulus, dan *battery health* dari *smartphone* masih bagus.



Gambar 4.5 *Word Cloud* Sentimen Positif pada Toko Ifonespot

Sementara itu hasil *word cloud* untuk sentimen positif pada toko Ifonespot dapat diketahui bahwa 7 kata yang paling sering muncul adalah bagus, barang, beli, mulus, bh (*battery health*), layanan, dan ramah yang berarti kualitas produk, kualitas pelayanan dan kualitas pengiriman barang toko Ifonespot sudah sangat baik dapat terlihat dari pelayanan admin dalam menanggapi pertanyaan pelanggan yang ramah, dan kualitas pengiriman barang yang bagus sehingga dari kualitas *smartphone* yang diperoleh pelanggan masih mulus dan bagus.



Gambar 4.6 *Word Cloud* Sentimen Positif pada Toko Jordan Corner

Berdasarkan hasil *word cloud* untuk sentimen positif pada toko Jordan Corner dapat diketahui bahwa 6 kata yang paling sering muncul adalah bagus, barang, bh (*battery health*), kirim, aman, dan ramah yang berarti kualitas pelayanan dan kualitas produk toko Jordan Corner sudah sangat baik dapat terlihat dari pelayanan admin dalam menanggapi pertanyaan pelanggan yang ramah, dan kualitas pelayanan dan pengiriman produk yang diperoleh masih aman (tidak ada kecacatan produk) sehingga dari kualitas *smartphone* yang diperoleh pelanggan masih bagus.

4.2.3.1 Sentimen Netral



Gambar 4.7 *Word Cloud* Sentimen Netral pada Toko Appelhouse.store

Hasil *word cloud* untuk sentimen netral pada toko Appelhouse.store dapat diketahui bahwa 6 kata yang paling sering muncul adalah barang, mulus, warna, moga, awet dan bh (*battery health*) yang berarti kualitas pelayanan dan kualitas produk toko Appelhouse.store sudah cukup baik karena dapat dilihat dari barang yang diperoleh masih mulus (tidak ada

kecacatan produk) sehingga kualitas *smartphone* yang diperoleh pelanggan memiliki warna dan *battery health* (bh) *smartphone* sesuai keterangan produk.



Gambar 4.8 *Word Cloud* Sentimen Netral pada Toko Ionespot

Sementara itu hasil *word cloud* untuk sentimen netral pada toko Ionespot dapat diketahui bahwa 6 kata yang paling sering muncul adalah beli, barang, mulus, hp, bh (*battery health*) dan puas yang berarti kualitas produk dibeli, kualitas pelayanan toko dan kualitas pengiriman barang toko Ionespot sudah cukup baik dapat dilihat dari pelanggan yang merasa puas akan produk yang diperoleh, dan kualitas pengiriman barang yang bagus sehingga membuat pelanggan dapat melakukan pembelian kembali (*repeat order*) karena kualitas *smartphone* yang diperoleh pelanggan masih mulus (tidak ada kecacatan produk) dan kualitas *smartphone* memiliki *battery health* (bh) sesuai keterangan produk.

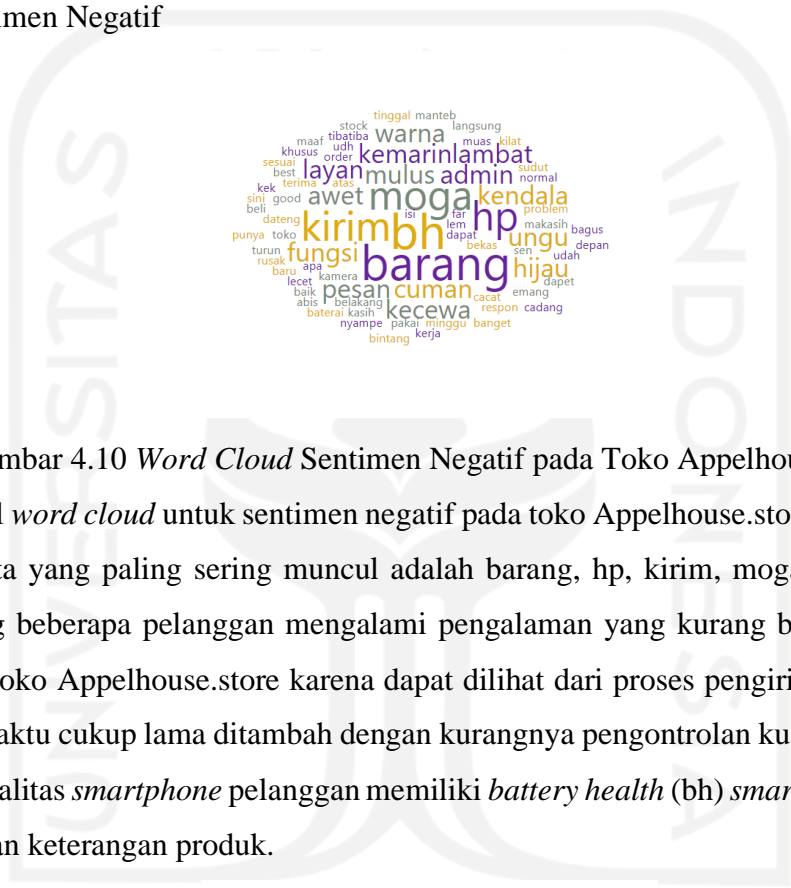


Gambar 4.9 *Word Cloud* Sentimen Netral pada Toko Jordan Corner

Berdasarkan hasil *word cloud* untuk sentimen netral pada toko Jordan Corner dapat diketahui bahwa 6 kata yang paling sering muncul adalah awet, barang, admin, bh (*battery*

health), moga, dan terimakasih yang berarti pelanggan menghargai dan bersyukur karena barang yang datang pelanggan cukup sesuai dengan barang yang diharapkan, dapat dilihat pada *battery health* (bh) yang sesuai dengan keterangan produk *smartphone* dan pelanggan berharap semoga *smartphone* yang ia beli dapat awet.

4.2.3.1 Sentimen Negatif



Gambar 4.10 *Word Cloud* Sentimen Negatif pada Toko Appelhouse.store

Hasil *word cloud* untuk sentimen negatif pada toko Appelhouse.store dapat diketahui bahwa 5 kata yang paling sering muncul adalah barang, hp, kirim, moga dan bh (*battery health*) yang beberapa pelanggan mengalami pengalaman yang kurang baik dari transaksi jual beli di toko Appelhouse.store karena dapat dilihat dari proses pengiriman barang yang memakan waktu cukup lama ditambah dengan kurangnya pengontrolan kualitas produk yang membuat kualitas *smartphone* pelanggan memiliki *battery health* (bh) *smartphone* yang tidak sesuai dengan keterangan produk.



Gambar 4.11 *Word Cloud* Sentimen Negatif pada Toko Ifonespot

Sementara itu hasil *word cloud* untuk sentimen negatif pada toko Ifonespot dapat diketahui bahwa 5 kata yang paling sering muncul adalah barang, turun, pakai, bh (*battery*

health) dan kendala yang berarti kualitas produk yang diperoleh pelanggan memiliki kualitas yang kurang baik, dapat dilihat pada kata yang paling sering muncul yang menunjukkan bahwa pelanggan memiliki kendala terhadap produk yang baru ia gunakan ternyata memiliki kualitas *battery health* (bh) yang rendah ditunjukkan pada kata turun sehingga *battery health* dari *smartphone* yang dibeli pelanggan baru digunakan beberapa hari telah mengalami penurunan *battery health*.



Gambar 4.12 *Word Cloud* Sentimen Negatif pada Toko Jordan Corner

Berdasarkan hasil *word cloud* untuk sentimen negatif pada toko Jordan Corner dapat diketahui bahwa 5 kata yang paling sering muncul adalah kirim, bh (*battery health*), pakai, fungsi, dan lecet yang berarti kualitas produk dan kualitas pengiriman barangnya kurang baik dapat dilihat pada *battery health* (bh) yang diperoleh pelanggan tidak sesuai dengan keterangan produk *smartphone* dan barang yang diperoleh terdapat lecet (ada kecacatan produk) ditambah lagi dengan kualitas pengiriman yang memakan waktu cukup lama.

BAB V PEMBAHASAN

5.1 Analisis Sentimen

Hasil penelitian analisis sentimen terhadap 3 toko *online* elektronik (Appelhouse.store, Iphonespot, dan Jordan Corner) pada *website* Shopee di Daerah Istimewa Yogyakarta dengan menggunakan python. Dari hasil pengolahan analisis sentimen pada 153 data untuk Appelhouse.store, 100 data untuk toko Iphonespot, dan 161 data untuk toko Jordan Corner menggunakan pendekatan *lexicon based* dengan kamus *SentiStrength* diperoleh sentimen positif, negatif, dan netral pada ketiga toko berdasarkan skor *polarity dari review* ditunjukkan pada Tabel 5.2 sebagai berikut.

Tabel 5.1 Hasil *SentiStrength* pada Ketiga Toko Elektronik

Kelas Sentimen	Appelhouse.store	Iphonespot	Jordan Corner
Positif	110	64	114
Netral	33	30	34
Negatif	10	6	13
Total	153	100	161

Dari hasil analisis sentimen pada Tabel 5.2, maka dapat disimpulkan bahwa toko Appelhouse.store memperoleh nilai pelabelan sentimen pada *review* pelanggan sebesar 71,9% sentimen positif, 21,6% sentimen netral, dan 6,5% sentimen negatif. Sementara itu pada toko Iphonespot mendapatkan nilai pelabelan sentimen pada *review* pelanggan sebesar 64% sentimen positif, 30% sentimen netral, dan 6% sentimen negatif. Sedangkan toko Jordan Corner memiliki nilai pelabelan sentimen positif sebesar 70,8%, sentimen netral sebesar 21,1%, dan sentimen negatif sebesar 8,1%. Dari hasil analisis sentimen pada *review* Shopee untuk ketiga toko menggunakan *SentiStrength ID* diperoleh nilai sentimen yang dominan yaitu sentimen positif, kemudian dilanjutkan dengan sentimen netral dan yang terendah adalah sentimen negatif.

Dari hasil penelitian sebelumnya oleh Aziz dan Fauziah (2022) dengan menggunakan metode analisis sentimen yang sama menggunakan *SentiStrength ID* untuk mengidentifikasi opini masyarakat terhadap produk, layanan dan kebijakan perusahaan di media sosial Twitter diperoleh nilai sentimen sebesar 54% untuk sentimen positif, 20% sentimen negatif dan 26% sentimen netral. Dari hasil analisis dengan data yang berbeda tetapi menggunakan algoritma

yang sama didapatkan hasil label yang sama yaitu, sentimen yang paling dominan adalah sentimen positif, urutan kedua sentiment netral dan yang paling rendah adalah sentimen negatif. Sedangkan pada penelitian Abdillah *et al* (2021) diperoleh hasil analisis sentimen menggunakan *SentiStrength* ID sebesar 67,11% sentimen negatif dan 32,89% sentimen positif. Sementara itu dengan metode yang sama juga untuk menganalisis sentimen terhadap *cyberbullying* di Twitter (Khaira *et al*, 2020) diperoleh nilai sentimen positif (*non-bullying*) sebesar 36,2%, sentimen negatif (*bullying*) sebesar 45,3% dan sentimen netral 18,5%. Dari hasil analisis antara ketiga penelitian sebelumnya dengan penelitian tugas akhir ini dapat diberikan kesimpulan bahwa jumlah sentimen dengan label positif, negatif, dan netral tidak dapat ditentukan sentimen dominannya karena ulasan yang ingin diteliti dengan menggunakan analisis sentimen pada konteks yang berbeda akan menghasilkan jumlah persentase label sentimen yang berbeda juga, ditambah dengan semakin besarnya jumlah data yang diteliti maka semakin beragam juga nilai sentimen yang diperoleh untuk tiap labelnya. Dari hasil pelabelan sentimen pada *review* pelanggan di ketiga toko, kemudian dapat dilanjutkan dengan visualisasi kata dari hasil sentimen positif, negatif, dan netral dengan *word cloud* sehingga diperoleh kata yang sering muncul untuk dapat dilakukan perencanaan perbaikan pada setiap toko kedepannya.

5.2 Word Cloud

Berdasarkan hasil analisis sentimen untuk sentimen positif, negatif dan netral, maka dapat dilanjutkan dengan memvisualisasikan ketiga sentimen tersebut menggunakan *word cloud* pada *tools python* untuk ketiga toko (Appelhouse.store, Iphonespot, dan Jordan Corner) sehingga diperoleh hasil *word cloud* sebagai berikut.

Tabel 5.2 Perbandingan Hasil *Word Cloud* Sentimen Positif

Perbandingan Hasil <i>Word Cloud</i>	Hasil <i>Word Cloud</i> Sentimen Positif
Appelhouse.store	Hasil <i>word cloud</i> untuk sentimen positif pada toko Appelhouse.store dapat diketahui bahwa 5 kata yang paling sering muncul adalah bagus, barang, mulus, aman, dan bh (<i>battery health</i>)

Perbandingan Hasil <i>Word Cloud</i>	Hasil <i>Word Cloud</i> Sentimen Positif
Iphonespot	Hasil <i>word cloud</i> untuk sentimen positif pada toko Iphonespot dapat diketahui bahwa 7 kata yang paling sering muncul adalah bagus, barang, beli, mulus, bh (<i>battery health</i>), layan, dan ramah.
Jordan Corner	Hasil <i>word cloud</i> untuk sentimen positif pada toko Jordan Corner dapat diketahui bahwa 6 kata yang paling sering muncul adalah bagus, barang, bh (<i>battery health</i>), kirim, aman, dan ramah.
Muktafin <i>et al</i> (2020)	Pada ulasan positif menunjukkan pembeli puas terhadap kualitas, kecepatan pengiriman dan harga barang.
Limbong <i>et al</i> (2021)	Hasil <i>word cloud</i> yang diperoleh dari kata dengan sentimen positif oleh pelanggan terkait gratis, bagus, suka, murah, mudah, dan cepat.
Sihombing <i>et al</i> (2021)	Hasil visualisasi <i>word cloud</i> menunjukkan kata yang paling sering muncul adalah barang, pengiriman, cepat, bagus, aman, produk, hp, pengiriman cepat, semoga, sesuai.

Dari hasil *word cloud* sentimen positif pada tabel 5.2 dapat disimpulkan bahwa penelitian yang dilakukan pada ketiga toko yaitu *appelhouse.store*, *ifonespot*, dan *jordan corner* memiliki inti ulasan yang serupa terkait kualitas produk, kualitas pelayanan dan kualitas pengiriman yang diberikan ketiga toko sudah sangat baik karena masih produk masih terjaga kualitasnya (tidak ada kecacatan produk), pelayanan admin dalam menanggapi pertanyaan pelanggan yang ramah, dan kualitas pengiriman barang yang bagus sehingga dari kualitas *smartphone* yang diperoleh pelanggan masih mulus dan bagus. Hasil tersebut selaras dengan penelitian sebelumnya yang menunjukkan hasil ulasan positif yang serupa mengenai kualitas produk, kualitas pelayanan yang diberikan dan kualitas pengiriman produk yang baik.

Tabel 5.3 Perbandingan Hasil *Word Cloud* Sentimen Netral

Perbandingan Hasil <i>Word Cloud</i>	Hasil <i>Word Cloud</i> Sentimen Netral
<i>Appelhouse.store</i>	Hasil <i>word cloud</i> untuk sentimen netral pada toko <i>Appelhouse.store</i> dapat diketahui bahwa 6 kata yang paling sering muncul adalah barang, mulus, warna, moga, awet dan bh (<i>battery health</i>).

Perbandingan Hasil <i>Word Cloud</i>	Hasil <i>Word Cloud</i> Sentimen Netral
Iphonespot	Hasil <i>word cloud</i> untuk sentimen netral pada toko Iphonespot dapat diketahui bahwa 6 kata yang paling sering muncul adalah beli, barang, mulus, hp, bh (<i>battery health</i>) dan puas.
Jordan Corner	Hasil <i>word cloud</i> untuk sentimen netral pada toko Jordan Corner dapat diketahui bahwa 6 kata yang paling sering muncul adalah awet, barang, admin, bh (<i>battery health</i>), moga, dan terimakasih.

Berdasarkan hasil *word cloud* sentimen netral pada tabel 5.3 dapat disimpulkan bahwa penelitian yang dilakukan pada ketiga toko yaitu *appelhouse.store*, *ifonespot*, dan *jordan corner* memiliki inti ulasan yang serupa terkait kualitas produk dan kualitas pelayanan yang diberikan ketiga toko cukup baik karena barang yang diperoleh masih mulus (tidak ada kecacatan produk) sehingga kualitas *smartphone* yang diperoleh pelanggan memiliki warna dan *battery health* (bh) *smartphone* sesuai keterangan produk. Penelitian ini belum dapat dibandingkan dengan penelitian terdahulu, karena penelitian terdahulu tidak mencantumkan sentimen netral dalam memvisualisasikan ulasan dengan menggunakan *word cloud*, dan hanya visualisasikan sentimen positif dan netral, sehingga perbandingan hasil *word cloud* hanya ketiga toko *online* yang diteliti saja.

Tabel 5.4 Perbandingan Hasil *Word Cloud* Sentimen Negatif

Perbandingan Hasil <i>Word Cloud</i>	Hasil <i>Word Cloud</i> Sentimen Negatif
<i>Appelhouse.store</i>	Hasil <i>word cloud</i> untuk sentimen negatif pada toko <i>Appelhouse.store</i> dapat diketahui bahwa 5 kata yang paling sering muncul adalah barang, hp, kirim, nya dan bh (<i>battery health</i>)
Iphonespot	Hasil <i>word cloud</i> untuk sentimen negatif pada toko Iphonespot dapat diketahui bahwa 7 kata yang paling sering muncul adalah barang, turun, pakai, banget, normal, bh dan kendala.
Jordan Corner	Hasil <i>word cloud</i> untuk sentimen negatif pada toko Jordan Corner dapat diketahui bahwa 5 kata yang paling sering muncul adalah kirim, bh, pakai, normal, dan lecet yang berarti kualitas produk dan kualitas pengiriman barangnya kurang baik

Perbandingan Hasil <i>Word Cloud</i>	Hasil <i>Word Cloud</i> Sentimen Negatif
Muktafin <i>et al</i> (2020)	Pada ulasan negatif pembeli kecewa pada warna, dan jumlah barang yang dikirim tidak sama dengan yang dipesan.
Limbong <i>et al</i> (2021)	Informasi sentimen negatif yang diperoleh seperti kata : kecewa, jelek, mahal, bohong, ribet, dan perbaiki.
Fahrudin <i>et al</i> (2022)	Hasil ulasan negatif menunjukkan pelanggan mengeluhkan ketidaksesuaian harga yang ditagihkan dengan label harga item produk yang tertera.

Hasil *word cloud* sentimen negatif pada tabel 5.4 dapat disimpulkan bahwa penelitian yang dilakukan pada ketiga toko yaitu *appelhouse.store*, *ifonespot*, dan *jordan corner* memiliki inti ulasan yang serupa terkait kualitas produk dan kualitas pengiriman yang diberikan ketiga toko kurang baik pada beberapa pelanggan karena proses pengiriman barang yang memakan waktu cukup lama ditambah dengan kualitas *smartphone* yang diperoleh masih kurang mulus (beberapa produk mengalami kecacatan produk) yang disebabkan oleh kurang telitinya pengontrolan kualitas produk sebelum *smartphone* di tangan pelanggan sehingga kualitas *smartphone* yang diperoleh pelanggan memiliki *battery health* (bh) yang tidak sesuai dengan keterangan produk.

5.3 Rekomendasi

5.3.1 Asosiasi Kata

Berdasarkan hasil *word cloud* sentimen negatif dari *online customer review* terhadap produk yang dibeli pelanggan pada ketiga toko *online* elektronik di *e-marketplace* Shopee, dapat dijadikan dasar untuk menemukan asosiasi kata agar dapat memberikan rekomendasi perbaikan untuk meningkatkan kualitas dari ketiga toko. Berikut ini adalah hasil asosiasi antar kata yang berhubungan dengan kata yang sering muncul secara bersamaan dan didapatkan hasilnya dengan bantuan *tool software* R sebagai berikut.

Tabel 5.5 Asosiasi Kata Dari Sentimen Negatif *Appelhouse.store* (1)

Barang		Hp		Kirim	
order	0,39	bh (<i>battery health</i>)	0,50	lambat	0,43
pesan	0,31	fungsi	0,29	hp	0,36
mulus	0,20	kirim	0,24	cacat	0,22

Barang		Hp		Kirim
sesuai	0,16	hijau	0,14	

Tabel 5.6 Asosiasi Kata Dari Sentimen Negatif Appelhouse.store (2)

Moga		bh (<i>battery health</i>)	
awet	0,46	awet	0,52
fungsi	0,27	hp	0,32
bh (<i>battery health</i>)	0,19		

Dari hasil tabel 5.5 dan 5.6 menunjukkan hasil asosiasi teks pada toko Appelhouse.store pada kata-kata yang berasosiasi dengan kata ‘barang’ dapat diperoleh informasi tentang barang yang diterima pelanggan dari hasil pemesanan produk memiliki kualitas produk yang kurang sesuai dengan deskripsi produk yang dicantumkan penjual dan pelanggan mengharapkan barang sampai dengan kondisi mulus walaupun barang yang dijual adalah barang *second*.

Sedangkan kata-kata yang berasosiasi dengan kata ‘hp’ menunjukkan bahwa kualitas fungsi hp atau fitur dari *smartphone* yaitu *battery health* memiliki kualitas yang rendah, karena dalam masa penggunaan yang belum terlalu lama sudah menyebabkan penurunan yang cukup signifikan pada *battery health* hp, dan juga warna *smartphone* yang dipesan pelanggan tidak sesuai, kemudian diganti dengan warna lain (hijau) tanpa persetujuan pelanggan. Kemudian pelanggan mengalami kendala akan keterlambatan pengiriman hp yang memakan waktu cukup lama dari waktu yang dijanjikan penjual.

Pada kata-kata yang berasosiasi dengan kata ‘kirim’ menunjukkan keluhan pelanggan akan kualitas pengiriman produk dari ekspedisi yang digunakan memakan waktu cukup lama agar barang/hp dapat sampai di tangan pelanggan. Selain itu juga barang yang sampai di pelanggan akibat proses pengiriman mengalami kecacatan atau kerusakan produk.

Kemudian pada kata ‘moga’ berasosiasi dengan kata kata yang menyiratkan harapan pelanggan akan produk yang diterima agar dapat bertahan dalam jangka waktu yang lama (awet), pelanggan juga berharap agar fungsi (fitur) dan *battery health* dari produk tidak mengalami kendala dan penurunan kualitas pada saat penggunaannya.

Kata-kata yang berasosiasi dengan kata ‘bh/ *battery health*’ menunjukkan bahwa produk hp yang diterima pelanggan masih memiliki kualitas *battery health* yang tidak sesuai pada keterangan produk/ yang dijanjikan penjual dan juga *battery health* produk tersebut

tidak awet karena sudah mengalami penurunan jumlah *battery health* hanya dalam penggunaan beberapa minggu saja.

Tabel 5.7 Asosiasi Kata Dari Sentimen Negatif Iphonespot (1)

Barang		Turun		Pakai	
awet	0,33	bh (<i>battery health</i>)	0,64	bh (<i>battery health</i>)	0,43
mulus	0,25	banget	0,41	fungsi	0,39
second	0,21	pakai	0,37	hari	0,28

Tabel 5.8 Asosiasi Kata Dari Sentimen Negatif Iphonespot (2)

Kendala		Bh (<i>battery health</i>)	
bh (<i>battery health</i>)	0,70	turun	0,44
turun	0,41	hp	0,26
silent	0,13		

Berdasarkan hasil tabel 5.7 dan 5.8 menunjukkan hasil asosiasi teks pada toko ifonespot pada kata-kata yang berasosiasi dengan kata ‘barang’ dapat diperoleh informasi tentang barang yang diterima pelanggan dari hasil pemesanan produk memiliki kualitas produk yang kurang sesuai dengan deskripsi produk yang dicantumkan penjual, selain itu juga barang yang diterima pelanggan kurang awet karena dalam penggunaan beberapa minggu sudah terjadi penurunan kualitas produk. Pelanggan mengharapkan barang sampai dengan kondisi mulus walaupun barang yang dijual adalah barang *second* akan tetapi kenyataannya barang yang diterima pelanggan belum memiliki kondisi yang mulus.

Pada kata-kata yang berasosiasi dengan kata ‘turun’ menunjukkan bahwa kualitas produk *smartphone* yang diterima memiliki kualitas *battery health* yang rendah, karena dalam masa pemakaian yang belum terlalu lama sudah menyebabkan penurunan yang cukup signifikan pada *battery health* produk.

Kemudian kata-kata yang berasosiasi dengan kata ‘pakai’ memberikan informasi terkait keluhan pelanggan akan kualitas produk dalam masa penggunaan pada beberapa hari yang ternyata menunjukkan bahwa beberapa fungsi hp atau fitur dari *smartphone* seperti *battery health* memiliki kualitas yang rendah, karena dalam masa penggunaan yang belum terlalu lama sudah menyebabkan perubahan angka *battery health* yang cukup signifikan.

Sementara itu pada kata ‘kendala’ berasosiasi dengan kata kata yang menyiratkan kendala apa saja yang didapatkan pelanggan setelah produk telah sampai seperti kualitas *battery health* produk yang ternyata mengalami kendala seperti terjadi penurunan kualitas

yang cukup signifikan, kemudian salah satu fitur *silent* dari *smartphone* mengalami kendala yaitu produk tidak bisa disunyikan (*silent*) untuk notifikasinya.

Kata-kata yang berasosiasi dengan kata ‘bh/ *battery health*’ menunjukkan bahwa produk hp yang diterima pelanggan masih memiliki kualitas *battery health* yang tidak sesuai pada keterangan produk/ yang dijanjikan penjual dan juga *battery health* produk tersebut sudah mengalami penurunan yang cukup drastis selama masa penggunaannya.

Tabel 5.9 Asosiasi Kata Dari Sentimen Negatif Jordan Corner (1)

Pakai		Kirim		Fungsi	
bh (<i>battery health</i>)	0,48	lambat	0,32	panas	0,29
panas	0,25	cacat	0,26	kamera	0,21
lemah	0,22	lecet	0,23	hp	0,18

Tabel 5.10 Asosiasi Kata Dari Sentimen Negatif Jordan Corner (2)

Lecet		Bh (<i>battery health</i>)	
kondisi	0,37	cepat	0,41
hp	0,32	turun	0,36
		hp	0,27
		barang	0,22

Berdasarkan hasil tabel 5.9 dan 5.10 menunjukkan hasil asosiasi teks pada toko Jordan corner pada kata-kata yang berasosiasi dengan kata kata ‘pakai’ memberikan informasi terkait keluhan pelanggan akan kualitas produk dalam masa pemakaian ternyata *smartphone* cepat panas dan baterai lemah. Kemudian *battery health* dari produk memiliki kualitas yang rendah, karena dalam masa penggunaan yang belum terlalu lama sudah menyebabkan perubahan angka *battery health* yang cukup signifikan.

Kemudian kata-kata yang berasosiasi dengan kata ‘kirim’ menunjukkan keluhan pelanggan akan kualitas pengiriman produk dari ekspedisi yang digunakan memakan waktu cukup lama agar barang dapat sampai di tangan pelanggan. Selain itu juga barang yang sampai di pelanggan mengalami kecacatan atau kerusakan produk seperti kondisi *smartphone* yang lecet akibat proses pengiriman.

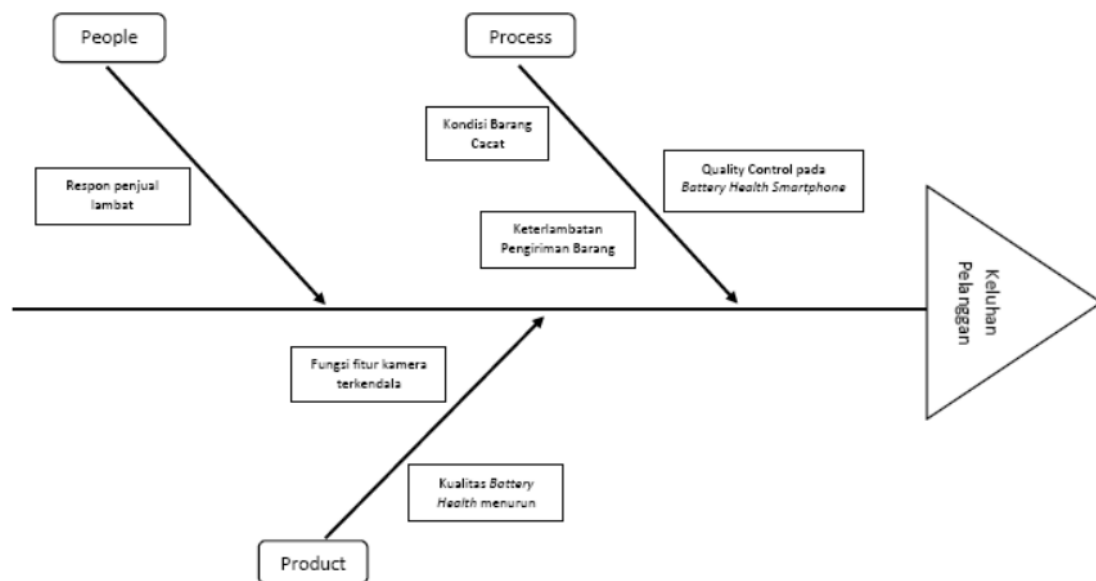
Sedangkan kata-kata yang berasosiasi dengan kata ‘fungsi’ memberikan informasi terkait keluhan pelanggan akan kualitas produk pada fungsi (fitur) hp yang cepat panas selama masa penggunaannya. Selain itu juga salah satu fitur hp yaitu kamera mengalami kendala karena fitur kamera tidak bisa disunyikan suaranya saat melakukan aktivitas yang membutuhkan kamera dari hp.

Pada kata ‘lecet’ berasosiasi dengan kata kata yang menyiratkan kualitas produk yang diterima pelanggan dalam keadaan kondisi hp kurang baik. Hal ini dapat disebabkan oleh kualitas pengemasan produk yang kurang aman, kurangnya *quality control* akan produk yang dijual, dan juga dapat disebabkan oleh keadaan yang kurang baik saat terjadi proses pengiriman

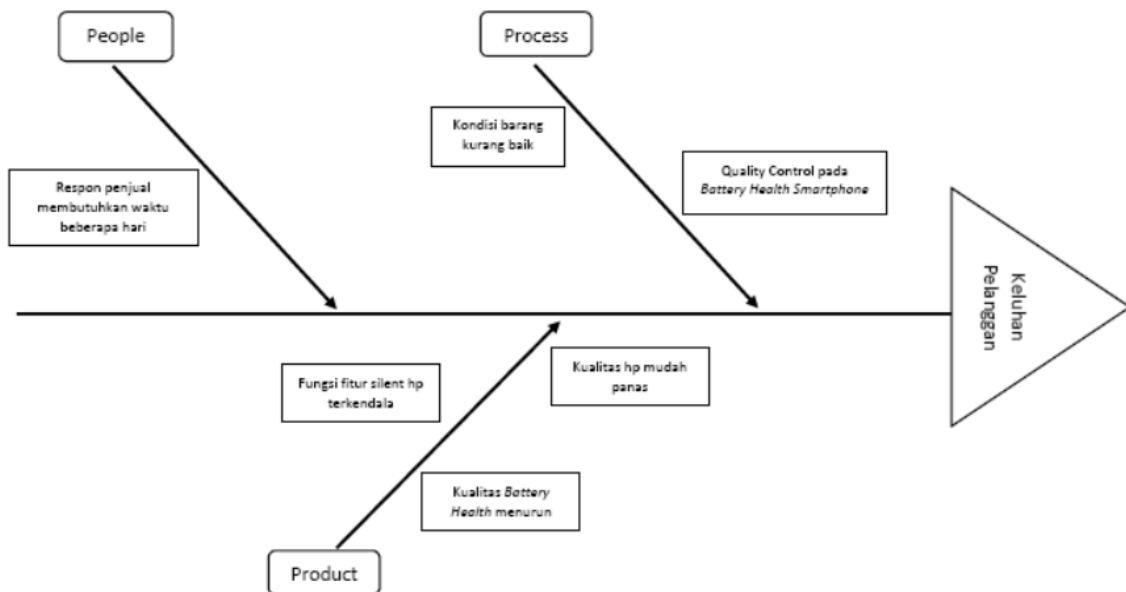
Kata-kata yang berasosiasi dengan kata ‘bh/ *battery health*’ menunjukkan bahwa barang yang diterima pelanggan masih memiliki kualitas *battery health* yang tidak sesuai pada keterangan produk/ yang dijanjikan penjual dan juga *battery health* hp tersebut sudah mengalami penurunan yang cukup drastis selama masa penggunaannya.

5.3.2 Diagram *Fishbone*

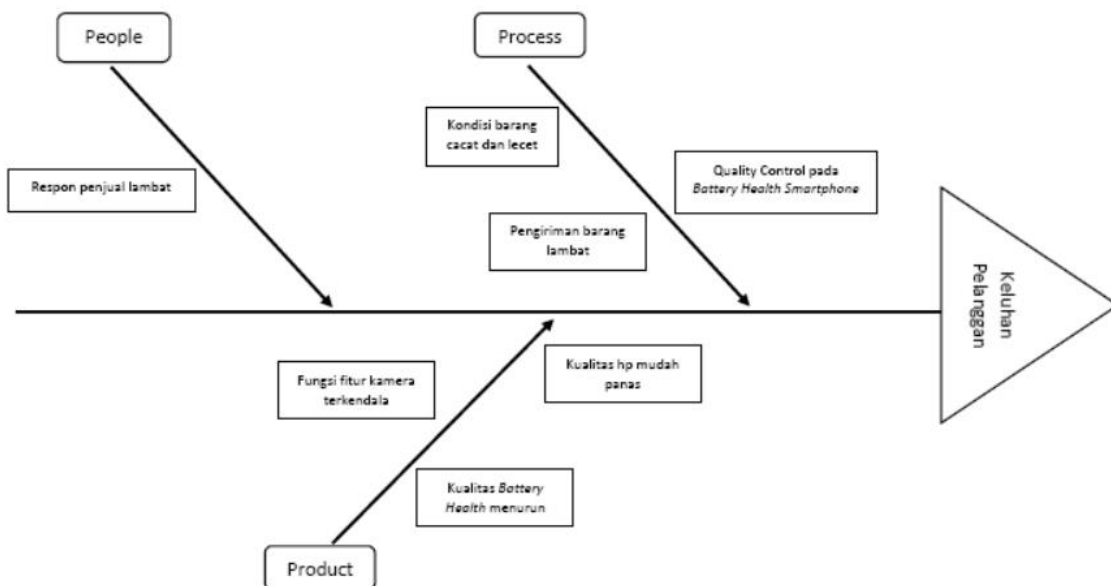
Dari hasil asosiasi kata tersebut maka dapat diketahui akar permasalahan atas ulasan negatif yang diberikan oleh pelanggan terhadap produk yang dijual oleh ketiga toko Appelhouse.store, Ifonespot, dan Jordan Corner.



Gambar 5.1 Diagram *Fishbone* Keluhan Pelanggan Toko Appelhouse.store



Gambar 5.2 Diagram *Fishbone* Keluhan Pelanggan Toko Iphonespot



Gambar 5.3 Diagram *Fishbone* Keluhan Pelanggan Toko Jordan Corner

Berdasarkan hasil asosiasi kata dari sentimen negatif tersebut maka dapat diperoleh informasi terkait masalah yang sering didapatkan pelanggan setelah membeli produk *smartphone* terhadap ketiga toko tersebut yang sebagian besar memiliki permasalahan yang serupa,

sehingga dapat diberikan saran perbaikan dari setiap pemecahan masalah yang ditemukan berdasarkan penelitian Praptiwi (2018) dan Gumilang (2018).

Tabel 5.11 Rencana Pemecahan Masalah Toko Appelhouse.store

No	Faktor	Permasalahan	Pemecahan Masalah
1	<i>Process</i>	Pengiriman barang lambat	Penyebab utama keterlambatan pengiriman sebenarnya terletak pada pihak jasa pengiriman. Oleh sebab itu perlu dimuat dalam keterangan di toko/ produk tips untuk mengatasi masalah / keterlambatan, dan beberapa masalah klasik yang sering terjadi dalam proses pengiriman. Selain itu dapat memilih layanan pemesanan barang yang tepat waktu dengan menggunakan jasa pengiriman yang profesional untuk memenuhi kebutuhan pelanggan sehingga tepat sasaran dan menjamin keamanan barang yang dikirimkan.
		Kurangnya <i>quality control</i> pada pengecekan bh (<i>battery health</i>) yang tidak sesuai dengan keterangan yang tertera	Meningkatkan <i>quality control</i> terhadap produk yang akan dikirimkan sebelum sampai di tangan pelanggan dengan cara memperhatikan kualitas produk yang tertera pada deskripsi dan pesanan pelanggan agar sesuai dengan produk yang diharapkan pelanggan.
		Kondisi barang cacat/ kurang baik	Lebih dijelaskan lagi di keterangan toko/produk mengenai garansi produk sampai apabila terjadi kerusakan dengan menyertakan bukti video <i>unboxing</i> produk dan melakukan pengajuan pengembalian pesanan (<i>return</i>) kepada admin setelah

No	Faktor	Permasalahan	Pemecahan Masalah
			berdiskusi agar dapat mencari jalan tengah yang adil.
2	<i>Product</i>	Kualitas bh (<i>battery health</i>) mudah menurun dan tidak awet Fungsi fitur hp terkendala pada kamera	Meningkatkan kinerja karyawan untuk melakukan <i>quality control</i> terhadap produk yang akan dikirimkan sebelum sampai di tangan pelanggan dengan cara tetap memperhatikan kualitas produk sesuai pesanan pelanggan.
3	<i>People</i>	Respon penjual terhadap chat pelanggan lambat	Meningkatkan lagi kinerja <i>customer service</i> / admin agar pelanggan tetap merasa puas terutama dalam hal melayani keluhan dan masalah yang dirasakan oleh pelanggan

Tabel 5.12 Rencana Pemecahan Masalah Toko Ifonespot

No	Faktor	Permasalahan	Pemecahan Masalah
1	<i>Process</i>	Kurangnya <i>quality control</i> pada pengecekan bh (<i>battery health</i>) yang tidak sesuai dengan keterangan yang tertera	Meningkatkan <i>quality control</i> terhadap produk yang akan dikirimkan sebelum sampai di tangan pelanggan dengan cara memperhatikan kualitas produk yang tertera pada deskripsi dan pesanan pelanggan agar sesuai dengan produk yang diharapkan pelanggan.
		Kondisi barang <i>second</i> kurang baik	Lebih dijelaskan lagi di keterangan toko/produk mengenai garansi produk sampai apabila terjadi kerusakan dengan menyertakan bukti video <i>unboxing</i> produk dan melakukan pengajuan pengembalian pesanan (<i>return</i>) kepada admin setelah berdiskusi agar dapat mencari jalan tengah yang adil.

No	Faktor	Permasalahan	Pemecahan Masalah
2	<i>Product</i>	Kualitas bh (<i>battery health</i>) mudah menurun dalam beberapa hari	Meningkatkan <i>quality control</i> produk secara keseluruhan baik fitur secara <i>hardware</i> maupun <i>software</i> sebelum sampai di tangan pelanggan.
		Fungsi fitur <i>silent hp</i> terkendala	
3	<i>People</i>	Respon penjual membutuhkan waktu beberapa hari	Meningkatkan lagi kinerja <i>customer service</i> / admin agar pelanggan tetap merasa puas terutama dalam hal melayani keluhan dan masalah yang dirasakan oleh pelanggan

Tabel 5.13 Rencana Pemecahan Masalah Toko Jordan Corner

No	Faktor	Permasalahan	Pemecahan Masalah
1	<i>Process</i>	Kurangnya <i>quality control</i> pada pengecekan bh (<i>battery health</i>) yang tidak sesuai dengan keterangan yang tertera	Meningkatkan <i>quality control</i> terhadap produk yang akan dikirimkan sebelum sampai di tangan pelanggan dengan cara memperhatikan kualitas produk yang tertera pada deskripsi dan pesanan pelanggan agar sesuai dengan produk yang diharapkan pelanggan.
		Pengiriman barang lambat	Penyebab utama keterlambatan pengiriman sebenarnya terletak pada pihak jasa pengiriman. Oleh sebab itu perlu dimuat dalam keterangan di toko/ produk tips untuk mengatasi masalah / keterlambatan, dan beberapa masalah klasik yang sering terjadi dalam proses pengiriman. Selain itu dapat memilih layanan pemesanan barang yang tepat waktu dengan menggunakan jasa pengiriman yang profesional untuk memenuhi kebutuhan pelanggan sehingga tepat sasaran dan menjamin

No	Faktor	Permasalahan	Pemecahan Masalah
			keamanan barang yang dikirimkan.
		Kondisi barang cacat dan lecet	Lebih dijelaskan lagi di keterangan toko/produk mengenai garansi produk sampai apabila terjadi kerusakan dengan menyertakan bukti video unboxing dan melakukan pengajuan pengembalian pesanan (<i>return</i>) kepada admin setelah berdiskusi untuk mencari jalan tengah yang adil.
2	<i>Product</i>	Kualitas bh (<i>battery health</i>) mudah menurun dalam beberapa hari Kualitas hp mudah panas dan lemah Fungsi fitur hp terkendala pada kamera	Meningkatkan <i>quality control</i> produk secara keseluruhan baik fitur secara <i>hardware</i> maupun <i>software</i> dan perlu adanya <i>testing</i> penggunaan hp sebelum sampai di tangan pelanggan.
3	<i>People</i>	Respon penjual terhadap chat pelanggan lambat	Meningkatkan lagi kinerja <i>customer service</i> / admin agar pelanggan tetap merasa puas terutama dalam hal melayani keluhan dan masalah yang dirasakan oleh pelanggan

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan dari penelitian tugas akhir yang telah dilakukan, maka diperoleh kesimpulan sebagai berikut.

1. Dari hasil analisis sentimen menggunakan kamus *SentiStrength*, diperoleh pelabelan *review* pelanggan pada toko Appelhouse.store dimana jumlah sentimen positif sebanyak 71,9%, sentimen netral 21,6%, dan sentimen negatif 6,5%. Sementara itu pada toko Iphonespot terdapat jumlah nilai sentimen sebesar 64% sentimen positif, 30% sentimen netral, dan 6% sentimen negatif. Sedangkan pada toko Jordan Corner memperoleh jumlah sentimen positif sebesar 70,8%, sentimen netral sebesar 21,1%, dan sentimen negatif sebesar 8,1%.
2. Hasil *word cloud* sentimen positif pada ketiga toko yaitu appelhouse.store, ifonespot, dan jordan corner memiliki inti ulasan yang serupa terkait kualitas produk, kualitas pelayanan dan kualitas pengiriman ('mulus', 'bagus', 'aman', dan 'ramah') serta sentimen netral memiliki inti ulasan terkait kualitas produk dan kualitas pelayanan ('bh/battery health', 'mulus', 'puas', dan 'awet'). Sedangkan sentimen negatif memiliki inti ulasan akan kualitas produk dan kualitas pengiriman ('bh/battery health', 'turun', dan 'kirim') yang diperoleh beberapa pelanggan kurang baik.
3. Untuk meningkatkan performa penjualan terutama pada *online customer review* toko *smartphone* Daerah Istimewa Yogyakarta di *e-marketplace* Shopee dapat diberikan beberapa rekomendasi kepada *seller* berdasarkan hasil *fishbone diagram* dengan meningkatkan *quality control* produk secara keseluruhan, menyertakan kebijakan untuk keterlambatan barang sampai dan kerusakan produk pada laman toko atau deskripsi produk karena kesalahan sebenarnya terletak pada jasa ekspedisi dan meningkatkan lagi kinerja *customer service* / admin agar pelanggan tetap merasa puas terutama dalam hal melayani keluhan dan masalah yang dirasakan oleh pelanggan.

6.2 Saran

Berdasarkan hasil analisis dan kesimpulan dari penelitian tugas akhir, adapun saran yang dapat diberikan sebagai berikut.

1. Pada penelitian ini masih memiliki batas jumlah data yang dikumpulkan dengan menggunakan *web scraping* pada *review* dan *rating* pelanggan Shopee di toko Appelhouse.store, Iphonespot, dan Jordan Corner. Sehingga diharapkan pada penelitian selanjutnya dapat mengumpulkan seluruh *review* yang terdapat pada toko yang akan diteliti di *website* Shopee, sehingga data yang diolah akan memberikan hasil yang lebih akurat dan representatif dari penelitian ini.
2. Diharapkan penelitian selanjutnya mampu menggunakan metode algoritma *lexicon based* lain dengan melakukan perbandingan antar metode algoritma tersebut, sehingga dapat mengetahui algoritma yang memiliki kinerja terbaik untuk melakukan pelabelan sentimen pada *review* pelanggan di *e-marketplace* Shopee.
3. Bagi *seller*, diharapkan dapat memperhatikan kembali hasil *review* dan *rating* dari pelanggan yang telah melakukan transaksi jual beli produk yang ditawarkan dengan mempertimbangkan dan mengimplementasikan rekomendasi yang telah diberikan.

DAFTAR PUSTAKA

- Abdillah, W. F., Premana, A., & Bhakti, R. M. H. (2021). Analisis Sentimen Penanganan Covid-19 dengan Support Vector Machine: Evaluasi Leksikon dan Metode Ekstraksi Fitur. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 3(02), 160–170. <https://doi.org/10.46772/intech.v3i02.556>
- Adiyana, I., & Hakim, F. (2015). Implementasi Teks Mining pada Mesin pencarian Twitter Untuk Menganalisis Topik-Topik Terkait KPK dan Jokowi. Prosiding Seminar Nasional Matematika dan Pendidikan Matematika UMS (pp. 570-581). Surakarta: UMS.
- Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. (2017). Machine Learning Techniques for Sentiment Analysis: A Review. *International Journal of Multidisciplinary Sciences and Engineering*, 8(3), 27–32.
- Ahmed, H., Korashy, H., Walaa, M. (2013). Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng. J.* 5(4), 1093–1113.
- Alamsyah, Andry & Faishal Nuruz Zuhri. (2017). “Measuring Public Sentiment Towards Services Level in Online Forum Using Naive Bayes Classifier and Word Cloud.” CRS-ForMIND International Conference and Workshop 2017 (October).
- Allahyari, M., Pouriye, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. <https://arxiv.org/abs/1707.02919>
- Amin, Fatkhul. (2012). Sistem Temu Kembali Informasi dengan Metode Vector Space Model. *Jurnal Sistem Informasi Bisnis* 02. Semarang: Universitas Stikubank.
- Astuti, W. et al. (2022). Analysis Sentiment on the Acceptance of Cpn 2021 on Twitter Social Media Using *SentiStrength*, 17(1), 1–6.
- Aziz, A., & Fauziah. (2022). Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1), 115–125.
- Baeza R.Y., dan Neto R., (1999). *Modern Information Retrieval*. Boston: USA.
- C. Mangles, “Understanding the Value Of Online Customer Reviews,” (2017). [Online]. Available: <https://www.smartinsights.com/online-pr/reputationmanagement->

online-pr/whats-review-understandingvalue-online-customer-reviews/. [Accessed: 21-Feb2018].

- Cindy Mutia Annur. (2022). 8 Produk yang Paling Diminati Konsumen Saat Belanja Online, Apa Saja? Retrieved June 12, 2022, from Katadata.co.id website: <https://databoks.katadata.co.id/datapublish/2022/03/28/8-produk-yang-paling-diminati-konsumen-saat-belanja-online-apa-saja>
- Coletta, L. F., da Silva, N. F., Hruschka, E. R., & Hruschka, E. R. (2014). Combining Classification and Clustering for Tweet Sentiment Analysis. *Brazilian Conference on Intelligent Systems*, 210 - 215.
- Darmalaksana, W., Hambali, R. Y. A., Masrur, A., & Muhlas. (2020). Analisis Pembelajaran Online Era WFH Pandemic Covid-19 sebagai Tantangan Pemimpin Digital Abad 21. *Karya Tulis Ilmiah (KTI) Era Work From Home (WFH) Covid-19 UIN Sunan Gunung Djati Bandung Tahun 2020*, 1(1), 1–12.
- David., (2014). *E-Marketplace Sebagai Penyedia Layanan Penjualan Barang*, ISSN: 2355-1941
- Devika, M. D., C. Sunitha, and Amal Ganesh. (2016). “Sentiment Analysis: A Comparative Study on Different Approaches.” *Procedia Computer Science* 87: 44–49. <https://doi.org/10.1016/j.procs.2016.05.124>.
- D’Andrea, A., Ferri, F., Grifoni, P., Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. *Int. J. Comput. Appl.* (0975-8887) 125(3)
- Erhaneli & Oki, I., (2015). Prediksi Perkembangan Beban Listrik Sektor Rumah Tangga di Kabupaten Sijunjung Tahun 2013-2022 dengan Simulasi SPSS. *Jurnal Momentum*, 17(2), pp. 14-25.
- Fadlisyah, B.D.A. (2014). *Statistika: Terapannya di Informatika, Edisi Pertama*. Yogyakarta: Graha Ilmu.
- Falentino Sembiring, Octaviana, Sudin Saepudin. (2020). “Implementasi Metode K-Means Dalam Pengklasteran Daerah Pungutan Liar Kabupaten Sukabumi (StudiKasus:Dinas Kependudukan dan Pencatatan Sipil)”, Vol.14, No.1.
- Fauziah, N. (2009). *Aplikasi Fishbone Analysis dalam Meningkatkan Kualitas Produksi Teh Pada PT Rumpun Sari Kemuning, Kabupaten Karanganyar*. Skripsi. Universitas Sebelas Maret.

- Fauziyyah, A. K. (2020). Analisis Sentimen Pandemi Covid19 Pada Streaming Twitter Dengan Text Mining Python. *Jurnal Ilmiah SINUS*, 18(2), 31. <https://doi.org/10.30646/sinus.v18i2.491>
- Ghozali, Imam. (2006). Aplikasi Analisis Multivariate dengan Program SPSS (Edisi Ke 4). Semarang:Badan Penerbit Universitas Diponegoro.
- Ghozali, Imam. (2013). Aplikasi Analisis Multivariat dengan Program SPSS. Semarang: Badan Penerbit Universitas Diponegoro.
- H. N. P. S. R. Michael Levi, (2020). "Perbandingan Performa Tools Web Scraping pada Website dengan Data Statis dan Dinamis," *Jurnal Infra*, vol. 8.
- H. S. P. H. M. Dhita Deviacita Ayani. (2019). "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," *Jurnal Sistem dan Teknologi Informasi*, vol. 7.
- Han, J., & Kamber, M. (2006). Classification and prediction. *Data mining: Concepts and techniques*, 347-350.
- Hernikawati Balai Pengembangan Sumber Daya Manusia dan Penelitian Kominfo Jakarta, D., Pegangsaan Timur No, J., & Pusat, J. (2021). Kecenderungan Tanggapan Masyarakat Terhadap Vaksin Sinovac Berdasarkan Lexicon Based Sentiment Analysis The Trend of Public Response to Sinovac Vaccine Based on Lexicon Based Sentiment Analysis. *Jurnal Ilmu Pengetahuan Dan Teknologi Komunikasi*, 23(1), 21–31. Retrieved from <http://dx.doi.org/10.33169/iptekkom.23.1.2021.21-31>
- Hikmah, J. (2017). STATISTIK DESKRIPTIF Leni Masnidar Nasution, 14(1), 49–55.
- Hootsuite, (2018). "2018 Digital Yearbook : Headline Internet, Social Media, and Mobile Use Data for Every Country in the World,".
- Iprice.co.id. (2022). [Laporan] Perusahaan E-Commerce Mana yang Paling Berpengaruh di Asia Tenggara pada Q1 2022?. Retrieved June 6, 2022, from Iprice.co.id website: <https://iprice.co.id/trend/insights/laporan-perusahaan-e-commerce-mana-yang-paling-berpengaruh-di-asia-tenggara-pada-q1-2022/>
- Išoraitė, M., & Miniotienė, N. (2018). Electronic Commerce: Theory and Practice. *IJBE (Integrated Journal of Business and Economics)*, 2(2), 194–200. <https://doi.org/10.33019/ijbe.v2i2.78>

- J. Reynolds, (2004). "Chapter 1: The E-Commerce Phenomenon," in *The Complete E-Commerce Book: Design, Build, & Maintain a Successful Web-based Business*, Second Edition.,CMP Media LLC.
- J. W. G. Putra, (2019). "Pengenalan Konsep Pembelajaran Mesin dan Deep Learning," no. August, pp. 1–235, [Online]. Available: <https://www.researchgate.net/publication/323700644>.
- Junaidi. (2014). *Statistik Deskriptif dengan Microsoft Office Excel*, 1–8.
- K. Spärck Jones, (2004). "IDF term weighting and IR research lessons," *J. Doc.*, vol. 60, no. 5, pp. 521–523, doi: 10.1108/00220410410560591.
- Kang, J.W.; Namkung, Y. (2019). The information quality and source credibility matter in customers' evaluation toward food O2O commerce. *Int. J. Hosp. Manag.* 78, 189–198
- Kemertian Perindustrian. (2019). *Laporan Kinerja Kementerian Perindustrian Tahun 2015-2019*. Kementerian Perindustrian, 1–128. <https://kemenperin.go.id/download/21250/Laporan-Kinerja-KementerianPerindustrian-2018>
- Khaira, U., Johanda, R., Utomo, P. E. P., & Suratno, T. (2020). Sentiment Analysis Of Cyberbullying On Twitter Using SentiStrength. *Indonesian Journal of Artificial Intelligence and Data Mining*, 3(1), 21. <https://doi.org/10.24014/ijaidm.v3i1.9145>
- Kunal, Sourav, Arijit Saha, Aman Varma, and Vivek Tiwari. (2018). "Textual Dissection of Live Twitter Reviews Using Naive Bayes." *Procedia Computer Science* 132 (Iccids): 307–13. <https://doi.org/10.1016/j.procs.2018.05.182>.
- Kuncoro, A., (2017). Korelasi Penguasaan Kosakata dengan Keterampilan Berbicara Siswa dalam Bahasa Inggris. *Jurnal Susunan Artikel Pendidikan*, 1(3), pp. 302-311.
- Lackermair, G,Kailer, D. & Kanmaz, K., (2013). Importance of Online Product Reviews from a Consumer's Perspective., 1(1),pp1-5.
- Li, N., & Zhang, P. (2002). Consumer online shopping attitudes and behavior: An assessment of research. *Information Systems proceedings of Eighth Americas Conference*.
- Loria, P. Keen,M. Honnibal, R. Yankovsky, D. Karesh and E. Dempsey. (2014). "Simplified Text Processing".
- M. G. Pradana. (2020). "Penggunaan Fitur Wordcloud Dan Document Term Matrix Dalam

- Text Mining,” *J. Ilm. Inform.*, vol. 8, no. 1, pp. 38–43.
- Mahendrajaya, R., Buntoro, G. A., & Setyawan, M. B. (2019). Analisis Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based Dan Support Vector Machine. *Komputek*, 3(2), 52. <https://doi.org/10.24269/jkt.v3i2.270>
- Manning, Christopher D, Prabhakar Raghavan, dan Hinrich Schütze. (2008). *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Mas Diyasa, I. G. S., Marini Mandenni, N. M. I., Fachrurrozi, M. I., Pradika, S. I., Nur Manab, K. R., & Sasmita, N. R. (2021). Twitter Sentiment Analysis as an Evaluation and Service Base On Python *SentiStrength*. *IOP Conference Series: Materials Science and Engineering*, 1125(1), 012034. <https://doi.org/10.1088/1757-899x/1125/1/012034>
- Moe, Wendy W. and David A. Schweidel (2012), “Online Product Opinions: Incidence, Evaluation and Evolution,” *Marketing Science*, 31, 3, 372–86.
- Murnawan, M. (2017). Pemanfaatan Analisis Sentimen Untuk Peningkatan Popularitas Tujuan Wisata. *Jurnal Penelitian Pos Dan Informatika*, 7(2), 109–120. <https://doi.org/10.17933/jppi.2017.070203>
- Nasukawa, T., & Yi, J. (2003). Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In *Proceedings of the 2nd International Conference on Knowledge Capture*, 70 - 77.
- Nuraeni, Y., Martasya, M., & Aufa, A. (2022). Analisis Produktivitas pada Pendapatan PT BMT OPPO Manufacturing Indonesia Terhadap Penjualan Smartphone OPPO pada Tahun 2019-2021. *Owner*, 6(1), 701–704. <https://doi.org/10.33395/owner.v6i1.643>
- Praneeth Sai, J. V., & Balachander, B. (2020). Sentimental analysis of twitter data using tweepy and *SentiStrength*. *International Journal of Advanced Science and Technology*, 29(3), 6537–6544
- Pratiwi, A., Taruna, R. M., ... Wibowo, D. S. A. (2020). Prediksi Curah Hujan Bulanan Di Wilayah Lombok Barat Menggunakan Principal Component Regression (PCR). *Jurnal Teori Dan Aplikasi Fisika*, 8(2), 175–182. <https://doi.org/10.23960/jtaf.v8i2.2597>
- R. Habibi, D. B. Setyohadi and E. Wati, (2016). "Analisis sentimen pada Twitter mahasiswa menggunakan metode backpropagation," *Jurnal Informatika*, vol. 12.

- R. Kinanti Ambarsari, (2018). “Analisis Kanal Penjualan Produk Smartphone pada E-Marketplace di Indonesia,”.
- R. Parlita, S. I. Pradika, A. M. Hakim, and K. R. N. M, (2020). “Analisis Sentimen Twitter Terhadap Bitcoin dan Cryptocurrency Berbasis Python *SentiStrength*,” *J. Ilm. Teknol. Inf. dan Robot.*, vol. 2, pp. 33–37.
- Rahmadi, (2016). *Sistem Informasi Manajemen*, Jakarta: Rajawali.
- Ratu Aghnia Raffaidy Wiguna, Andri Irfan Rifai. (2021). “Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining”, Vol. 3, No. 1.
- S. W. Iriananda, M. A. Muslim, and H. S. Dachlan, (2019). “Identifikasi Kemiripan Teks Menggunakan Class Indexing Based dan Cosine Similarity Untuk Klasifikasi Dokumen Pengaduan,” *MATICS*, vol. 10, no. 2, p. 30, doi: 10.18860/mat.v10i2.5327.
- Santoso, E. B., & Nugroho, A. (2019). Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook. *Eksplora Informatika*, 9(1), 60–69. <https://doi.org/10.30864/eksplora.v9i1.254>
- Setiawan, D. F., Tristiyanto, T., & Hijriani, A. (2020). Aplikasi web scraping deskripsi produk. *Jurnal Teknoinfo*, 14(1), 41. <https://doi.org/10.33365/jti.v14i1.498>
- Sindhvani, V, Melville, P. (2008). Document-word co-regularization for semi-supervised sentiment analysis. In: Eight IEEE International Conference on Data Mining, 2008.ICDM 08, pp. 1025–1030. IEEE.
- Thakur, R. (2018). Customer engagement and online reviews. *Journal of Retailing and Consumer Services*, 41(November 2017), 48–59.
- Ulwan, M.N. (2016). *Pattern Recognition pada Unstructured Data Teks Menggunakan Support Vector Machine dan Association*. Skripsi. Program Studi Statistika FMIPA UII Yogyakarta.
- Valatehan, Lucky, Muhammad Fachrurrozi, dan Osvari Arsalan. (2016). Identifikasi Kalimat Pemborosan Menggunakan Rule Based Reasoning. *Annual Research Seminar Vol 2 No. 1*. Palembang: Universitas Sriwijaya.

- Wahid, D. H., & Azhari, S. N. (2016). Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 10(2), 207-218.
- Xie, K. L., Chen, C., & Wu, S. (2016). Online Consumer Review Factors Affecting Offline Hotel Popularity: Evidence from Tripadvisor. *Journal of Travel and Tourism Marketing*, 33(2), 211–223
- Y. A. B. Raharjo, F. E. N. Saputro, M. E. Pratiwi, T. A. Lorosae, R. Ardhani, F. Sumanto, F. Fitriadi dan E. Utami, (2017). “Rekayasa Kebutuhan Perangkat Lunak E-Marketplace Gerobak Kopi,” *Jurnal Sistem Informasi dan Teknologi Informasi*, vol. 6, no. 1, pp. 21 – 30.
- Yosepha Pusparisa. (2020). Pengguna Smartphone diperkirakan Mencapai 89% Populasi pada 2025. Retrieved June 12, 2022, from Katadata.co.id website: <https://databoks.katadata.co.id/datapublish/2020/09/15/pengguna-smartphone-diperkirakan-mencapai-89-populasi-pada-2025>

LAMPIRAN

Lampiran 1. Script Python Web Scraping

```

1 import requests
2 import csv
3 from datetime import datetime
4
5 def shopee(url):
6     shop_name = url.split('/')[1]
7     url = f'https://shopee.co.id/api/v4/shop/get_shop_detail?sort_sold_out=0&username={shop_name}'
8     req = requests.get(url)
9     data_shop = req.json()
10    shop_id = data_shop['data']['shopid']
11    user_id = data_shop['data']['userid']
12    count = 0
13    result = []
14    while True:
15        try:
16            count += 6
17            print(count)
18            url = f'https://shopee.co.id/api/v4/seller_operation/get_shop_ratings?limit=6&offset={count}&shop_id={shop_id}&user_id=
{user_id}'
19            req = requests.get(url)
20            data_req = req.json()
21            for value in data_req['data']:
22                data_result = {
23                    'nama pengguna': value['author_username'],
24                    'produk': value['product_items'][0]['name'],
25                    'review': value['comment'],
26                    'rating': value['rating_star'],
27                    'waktu transaksi': datetime.utcfromtimestamp(value['ctime']).strftime('%Y-%m-%d %H:%M')
28                }
29                result.append(data_result)
30            except KeyError:
31                break

```

Lampiran 2. A-Script Python Preprocessing Data Case Folding

```

In [8]: # Proses Case Folding
import re
def casefolding(Review):
    Review = Review.lower()
    Review = Review.strip(" ")
    Review = re.sub(r'[?|!|@|_|!|?|:|"](-+|,|'|',Review)
    return Review
data['Review'] = data['Review'].apply(casefolding)
data.head()

```

Lampiran 3. B- Script Python Preprocessing Data Case Folding

```
In [11]: # emoji removal
import re
import emoji

def remove_emoji(string):
    emoji_pattern = re.compile("[
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        u"\U00002702-\U000027B0"
        u"\U000024C2-\U0001F251"
    ]+", flags=re.UNICODE)
    return emoji_pattern.sub(r'', string)
return data
data['Review'] = data['Review'].apply(remove_emoji)
data.head()
```

Lampiran 4. Script Python Preprocessing Data Tokenizing

```
In [12]: # Proses Tokenizing
def token(Review):
    nstr = Review.split(' ')
    dat= []
    a = -1
    for hu in nstr:
        a = a + 1
    if hu == '':
        dat.append(a)
    p = 0
    b = 0
    for q in dat:
        b = q - p
        del nstr [b]
        p = p + 1
    return nstr
data['Review'] = data['Review'].apply(token)
data.head()
```

Lampiran 5. Script Python Preprocessing Data Filtering

```
In [15]: # Proses Filtering
import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords

def stopword_removal(Review):
    filtering = stopwords.words('indonesian','english')
    x = []
    data = []
    def myFunc(x):
        if x in filtering:
            return False
        else:
            return True
    fit = filter(myFunc, Review)
    for x in fit:
        data.append(x)
    return data
data['Review'] = data['Review'].apply(stopword_removal)
data.head()

[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\Asus\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
```

Lampiran 6. Script Python Preprocessing Data Stemming

```
In [17]: # Proses Stemming
! pip install Sastrawi
from sklearn.pipeline import Pipeline
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

def stemming(Review):
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    do = []
    for w in Review:
        dt = stemmer.stem(w)
        do.append(dt)
    d_clean=[]
    d_clean=" ".join(do)
    print(d_clean)
    return d_clean
data['Review'] = data['Review'].apply(stemming)

data.to_csv('data_clean.csv', index=False)
data_clean = pd.read_csv('data_clean.csv', encoding='latin1')
data_clean.head()
```

Lampiran 7. Script Python Pelabelan Data *SentiStrength*

```
import string
import pandas as pd
import numpy as np
import os
from nltk.sentiment.vader import SentimentIntensityAnalyzer
import json
import replib
import nltk
```

```
data = pd.read_csv("data3_clean.csv", sep=',', encoding="ISO-8859-1")
print(data.head())
```

```
sia2 = SentimentIntensityAnalyzer()
sia2.lexicon.clear()
with open('_json_sentiwords_id.txt') as f:
    data2 = f.read()
```

```
senti = json.loads(data2)
sia2.lexicon.update(senti)
```

```
def is_positive_senti(data: str) -> bool:
    """True if data has positive compound sentiment, False otherwise."""
    return sia2.polarity_scores(data)["compound"] > 0
data['is_positive_senti'] = data['Review'].apply(is_positive_senti)
data.head()
```

```
data['senti'] = data['Review'].apply(sia2.polarity_scores)
data.head()
```

الجمعة ١٤٤١هـ
الجمعة ١٤٤١هـ
الجمعة ١٤٤١هـ

Lampiran 8. *Script Python Word Cloud*

```
In [5]: import matplotlib.pyplot as plt
import numpy as np
from wordcloud import WordCloud, STOPWORDS
import sys, os
```

```
In [7]: os.chdir(sys.path[0])
```

```
In [20]: text = open('Neu_JC.txt', mode='r', encoding='utf-8').read()
stopwords = STOPWORDS

wc = WordCloud(
background_color='white',
stopwords=stopwords,
height = 600,
width=400
)

wc.generate(text)

wc.to_file('wordcloud_output.png')
```

```
Out[20]: <wordcloud.wordcloud.WordCloud at 0x2806a28fee0>
```

Lampiran 9. *Script Software R Asosiasi Kata*

```
library("tm")
library("SnowballC")
library("wordcloud")
library("RColorBrewer")
library("stringr")

setwd("D:/Tugas/smt 7/Yok Bisa TA")
docs<-readLines("Neg_AS.csv")
# Load the data as a corpus
```



```
docs<- Corpus(VectorSource(docs))

#Build a term-document matrix
dtm<- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
head(d, 25)

#asosiasi kata
v<-as.list(findAssocs(dtm, terms
=c("barang","hp", "kirim", "moga" dan "bh"),
corlimit = c(0.15,0.15,0.15,0.15,0.15)))
```