

ANALYSIS OF MSME CUSTOMER PERCEPTION THROUGH SENTIMENT REVIEWS USING K-MEANS CLUSTERING

Sri Wulandari¹, M. Trihudyatmanto², Bahtiar Efendi³

^{1,2,3}Fakultas Ekonomi dan Bisnis, Universitas Sains Al – Qur'an

¹email: sriwulandari201102@gmail.com

²email: trihudyatmanto@unsiq.ac.id

³email: bahtiarefd@unsiq.ac.id

ABSTRACT

Digital technology continues to develop relentlessly and creates fundamental changes for businesses, thus opening up new opportunities but also creating more complex uncertainties. This study aims to analyse the perception of MSME customers using Google Maps reviews with the K-Means Clustering model. The dataset is retrieved through web scraping and processed with Python-based sentiment analysis. The results show customer perceptions in a more structured manner using the K-Means Model and sentiment analysis. Most reviews are positive, focusing on location, product quality, and price.

Keywords: customer perceived value, k-means, umkm, python, sentiment reviews.

ABSTRAK

Teknologi digital terus mengalami perkembangan tanpa henti dan menciptakan perubahan mendasar bagi bisnis, sehingga membuka peluang baru namun juga memunculkan ketidakpastian yang lebih kompleks. Penelitian ini bertujuan untuk menganalisis persepsi pelanggan UMKM menggunakan ulasan *Google Maps* dengan model *K-Means Clustering*. *Dataset* diambil melalui *web scraping* dan diproses dengan analisis sentimen berbasis *Python*. Hasil menunjukkan persepsi pelanggan secara lebih terstruktur dengan menggunakan Model *K-Means* dan analisis sentimen. Sebagian besar ulasan adalah positif, dengan fokus pada lokasi, kualitas produk, dan harga.

Kata kunci: *customer perceived value*, k-Means, umkm, python, ulasan sentimen.

1. Pendahuluan

Teknologi digital terus mengalami perkembangan tanpa henti dan menciptakan perubahan mendasar bagi bisnis, sehingga membuka peluang baru namun juga memunculkan ketidakpastian yang lebih kompleks (Nambisan et al., 2017). Meluasnya penggunaan teknologi digital telah membawa perubahan signifikan pada cara organisasi berdiri, beroperasi, dan berinteraksi dengan pelanggan (Mihu et al., 2023). Transformasi digital menawarkan peluang yang signifikan bagi UMKM, termasuk peningkatan efisiensi operasional, akses ke pasar yang lebih luas, dan kemampuan untuk menganalisis data untuk pengambilan keputusan yang lebih baik (Hendrawan et al., 2024). Namun demikian, banyak UMKM mengalami kendala secara signifikan dalam penerapan teknologi digital di Indonesia, sehingga menghadapi beberapa tantangan termasuk kurangnya pengetahuan keterampilan teknis, biaya implementasi yang tinggi, serta infrastruktur yang tidak memadai (Handayati et al., 2024).

Di era digital, ulasan pelanggan *online* memainkan peran penting dalam proses pengambilan keputusan konsumen (Sholeh et al., 2024). Peran UMKM *online* semakin meningkat setiap harinya pada setiap bisnis, dan oleh karena itu mampu menawarkan nilai kepada pelanggan yang pada gilirannya akan memberikan keuntungan bagi pebisnis dan diferensiasi dari para pesaing lain (Jaiswal & Singh, 2020). Di dunia yang saling terhubung secara digital saat ini, ulasan pelanggan memiliki pengaruh yang amat signifikan terhadap perilaku konsumen dan keputusan pembelian (Godara et al., 2024). Ulasan pelanggan *online* memainkan peran penting dalam mengevaluasi kinerja produk atau layanan dan berfungsi sebagai referensi penting untuk keputusan pembelian pelanggan lain (C. Huang et al., 2024). Meskipun ulasan pelanggan mendukung UMKM, mereka membutuhkan testimoni yang kuat untuk diferensiasi karena tingginya nilai ulasan di antara para pesaing lokal mereka.

Dalam penelitian ini, peneliti menggunakan Ulasan *Google Maps* sebagai

sumber informasi, *Google Maps* adalah *platform web* yang memungkinkan pengguna mengulas bisnis dan layanan, termasuk perpustakaan (Borrego & Comalat Navarra, 2021). Penelitian ini akan membahas persepsi pelanggan pada ulasan *google maps*, Persepsi adalah proses pemilihan, pengorganisasian dan penginterpretasian masukan informasi, sensasi yang diterima melalui penglihatan, perasaan, pendengaran, penciuman dan sentuhan, untuk menghasilkan makna (Swanson & Davis, 2020). Dalam konteks belanja *online*, persepsi pelanggan dipengaruhi oleh faktor-faktor seperti manfaat, kemudahan penggunaan, efikasi diri, keamanan, kepercayaan, dan kegunaan yang dirasakan (Syahdan et al., 2021).

Penelitian ini bertujuan untuk mengisi kekosongan pengetahuan tentang bagaimana persepsi pelanggan pada ulasan *Google Maps* menggunakan metode *K-Means Clustering*. Algoritma *K-means clustering* adalah algoritma analisis *clustering* yang diselesaikan secara iteratif, langkahnya adalah memilih K objek secara acak sebagai pusat *cluster* awal, kemudian menghitung jarak antara setiap objek dengan setiap pusat *cluster seed*, dan menetapkan setiap objek ke pusat *cluster* terdekat (Z. Huang et al., 2024). Pengujian dilakukan menggunakan bahasa pemrograman *Python*, *Python* adalah bahasa pemrograman yang mudah dipelajari dan dipahami, bahkan bagi pemula, sintaksnya yang sederhana dan mirip bahasa Inggris membuatnya sangat populer. Selain itu, *Python* juga sangat fleksibel dan bisa digunakan untuk berbagai macam proyek (*Python*, 2021).

Penelitian terdahulu yang dilakukan oleh (Febriani & Putri, 2020) hasil pengolahan data, terbentuk 3 *cluster* dari 25 data pelanggan, Berdasarkan *cluster* yang terbentuk, dapat disimpulkan bahwa pembelian pelanggan memiliki pola yang berbeda. Sedangkan penelitian yang dilakukan oleh (Azmy & Furinto, 2008) hasil penelitian adalah setelah dilakukan pengulangan faktor analisis dengan rotasi, maka didapatkan 3 faktor yaitu: Faktor Atribut: Ekuitas Merek, Panduan, Referensi, Penawaran Kualitas dan Keyakinan Pencarian Manfaat Faktor Komunitas: Kesesuaian Religius dan Afiliasi

yang Dirasakan. Faktor Eksklusivitas : Perlakuan Khusus dan *Perceived Comfort* Variabel Ekuitas Merek, Referensi, Penawaran Kualitas, Manfaat Keyakinan dan Manfaat Perlakuan Khusus terbukti mempengaruhi variabel *Perceived Comfort*. Penerapan Algoritma *K-Means Clustering Analysis* Pada Penyakit Menular Manusia menginput sampel data sebanyak 32 data dan menghasilkan 6 *cluster* kelompok data (Bastian et al., 2018).

Dengan mengidentifikasi kelompok pelanggan berdasarkan persepsi mereka, UMKM dapat merancang strategi pemasaran yang lebih efektif, selain itu dapat membantu UMKM dalam mengidentifikasi area yang perlu diperbaiki dalam produk atau layanan mereka. Selain itu hasil dari penelitian ini bisa menjadi landasan bagi studi lanjutan yang lebih mendalam mengenai analisis sentimen dan persepsi konsumen pada UMKM. Selain itu, metodologi yang digunakan dapat diterapkan pada berbagai *platform* ulasan *online* lainnya, memperluas cakupan penelitian di masa depan.

2. Tinjauan Pustaka

Customer Perceived Value

Customer Perceived Value menempati posisi penting baik dalam penelitian akademis maupun praktik bisnis. *Customer Perceived Value* mengacu pada perbandingan antara manfaat yang diperoleh pelanggan dari suatu produk atau layanan dengan pengorbanan yang mereka keluarkan (Dodds et al., 1991). *Customer Perceived Value* dianggap sebagai faktor kunci dalam keberhasilan bisnis karena dapat meningkatkan kepuasan dan loyalitas pelanggan, yang pada akhirnya mendukung keuntungan jangka Panjang (Kumar & Reinartz, 2016).

Sebagai elemen penting dalam pengelolaan bisnis, *Customer Perceived Value* juga menjadi fokus utama dalam penelitian pemasaran dan layanan. Para ahli telah memberikan banyak wawasan tentang bagaimana *Customer Perceived Value* didefinisikan (Zeithaml et al., 2020), bagaimana cara mengukurnya (Leroi-Werelds et al., 2014) dan pengaruhnya terhadap keputusan pelanggan (Gallarza et al., 2017).

Sentiment Analysis

Dalam beberapa tahun terakhir, pesatnya perkembangan internet telah menghasilkan sejumlah besar komentar *online*, sehingga analisis sentimen dari komentar-komentar tersebut menjadi sangat penting untuk mendukung stabilitas dan pembangunan sosial (Birjali et al., 2021). Sebagai tanggapan terhadap tren ini, teknologi *Sentiment Analysis* (SA) mulai berkembang. SA, yang juga dikenal sebagai analisis opini atau penambahan opini, merupakan bidang penting dalam *Natural Language Processing* (NLP). Teknologi ini dirancang untuk secara otomatis mengekstrak dan menganalisis sentimen serta pandangan yang terdapat dalam teks (Chaturvedi et al., 2018). Menganalisis sikap dan emosi di balik komentar dapat memberikan bantuan yang kuat untuk bisnis, pemerintah, dan akademisi (Mao et al., 2024)

3. Metode Penelitian

Web Scraping

Pada penelitian ini sumber data diambil dari tautan web *Google Maps Review* Warung Enthok Bu Siti Wonosobo, *Web scraper* berfungsi mengambil tautan dari web, kemudian data diekstrak untuk mendapatkan data dari tautan sumber dan akhirnya menyimpan data tersebut ke dalam file *csv* (Thomas & Mathur, 2019). Proses ini memungkinkan data ulasan sentimen pada *google maps* yang tidak terstruktur menjadi data terstruktur sehingga dapat disimpan dan dianalisis dalam pusat *database* (Khder, 2021) yang bertujuan sama dengan visi web semantik, sebuah inisiasi ambisius yang dengan terbosan dalam pemrosesan teks, pemahaman semantik, kecerdasan buatan, serta interaksi manusia-komputer (Bogers et al., 2017). Pada penelitian ini data hasil *scraping* menghasilkan 312 *dataset* ulasan pelanggan yang akan dijadikan sumber data pada penelitian ini.

Data Cleaning

Data cleaning didefinisikan sebagai serangkaian proses yang dirancang untuk mendeteksi dan memperbaiki kesalahan dalam data, bertujuan meningkatkan kualitas data agar bisa diandalkan dalam analisis atau pengambilan keputusan (Chu, 2019). Mengidentifikasi dan membersihkan data yang

tidak rapi merupakan tantangan utama dalam analisis data. Jika proses ini diabaikan, hal tersebut dapat menyebabkan analisis yang keliru dan keputusan yang kurang dapat dipercaya (Chu et al., 2016). Pada Penelitian ini hasil dari data cleaning bisa dilihat pada gambar berikut :

Gambar 1. Hasil Data Cleaning

	Ulasan
0	rasa lumayan kalau di goreng kalau kuah agak k...
1	yang lebih shock lagi pas mbayar es teh es teh...
2	warungnya cuma menyediakan 1 menu yaitu opor e...
3	warung makan legend yang unik serba ambil send...
4	tiap mudik pasti mampir kesni opornya mantul p...
...	...
306	gemuk
307	legenda
308	tempat tidur enak
309	itu bagus
310	direkomendasikan

Sumber: hasil data olah phyton, 2024

Preprocessing

Rangkaian teknik yang digunakan sebelum penerapan metode data mining dinamakan sebagai data *preprocessing* untuk *data mining* (García et al., 2015) *Data Preprocessing* adalah langkah awal yang krusial dalam analisis data yang mencakup sejumlah teknik untuk membersihkan, mengubah, dan menyusun ulang data. Pemrosesan awal diperlukan agar properti yang diminati dapat diprediksi dengan benar (Mishra et al., 2020; Tri Jaka, 2015). Hasil preprocessing pada penelitian ini bisa dilihat pada gambar berikut :

Gambar 2. Preprocessing Data

Ulasan	english
lebih shock pas mbayar es teh es teh nya 5000 ...	It was even more of a shock when I paid for th...
warungnya cuma menyediakan 1 menu opor entok d...	the stall only provides 1 menu of opor duck he...
warung makan legend unik serba ambil sendiri s...	Unique legendary food stall, all pick your own...
tiap mudik mampir kesni opornya mantul potonga...	Every time I go home, I stop by here, Opor Man...
kuliner unik recommended nasi sambal opor entok...	Unique culinary delights, recommended for rice...

Sumber: hasil data olah phyton, 2024

K-Means Clustering

Clustering adalah teknik yang digunakan untuk menganalisis data dengan cara yang efisien dan menghasilkan informasi yang

dibutuhkan (Xu & Wunsch, 2008). *K-Means* adalah metode pengelompokan yang bersifat tidak diawasi, *non-deterministik*, *numerik*, dan bekerja secara *interactive* (Yadav & Sharma, 2013). Dalam metode ini, setiap kluster direpresentasikan oleh nilai rata-rata dari objek-objek yang termasuk di dalamnya. Untuk mengelompokkan *dataset*, ada teknik bernama *k-means*, yang diterapkan yang didasarkan pada pemilihan titik pusat dan perhitungan Jarak *Euclidian*. Di sini, dalam *k-means*, *dataset* akan dimuat dan dari *dataset*. Titik pusat dipilih dengan menggunakan rumus jarak *Euclidian* dan berdasarkan jarak *Euclidian*, titik-titik tersebut dimasukkan ke dalam *cluster*.

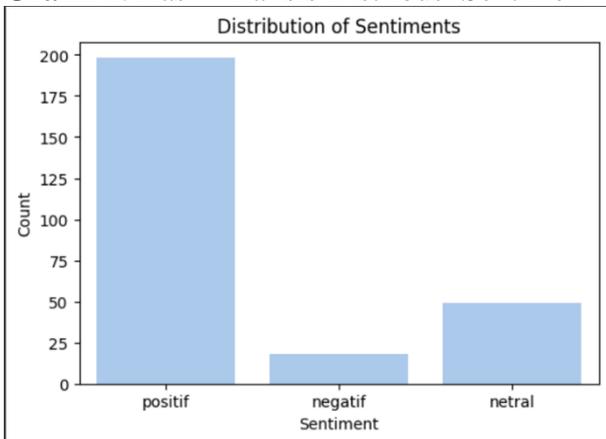
$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

- p, q = dua titik ruangan -n *Euclidean*
- q_i, p_i = vektor *Euclidean*, dimulai dari asal ruang (titik awal)
- n = ruang-n

4. Hasil dan Pembahasan

Hasil analisis sentimen ulasan pelanggan menunjukkan bahwa sebagian besar sentimen positif; hasil ulasan lain bersifat netral, dan sebagian besar sentimen negatif. Ini menunjukkan bahwa sebagian besar pelanggan memiliki persepsi positif tentang Warung Enthok Ibu Siti Wonosobo. Hasil data analisis sentimen dapat dilihat pada grafik berikut:

Grafik 1. Hasil Analisis Distribusi Sentimen

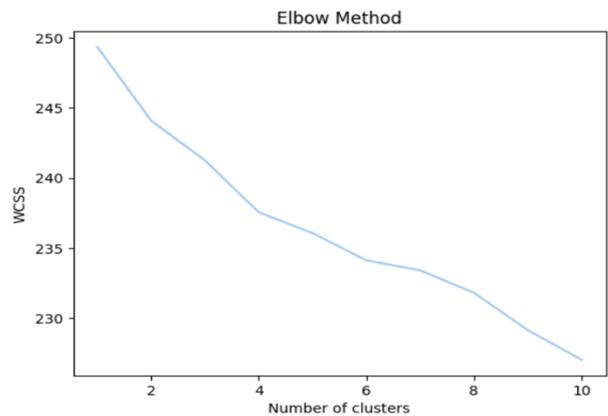


Sumber : hasil data olah python, 2024

Analisis sentimen yang dilakukan terhadap 265 ulasan pelanggan menunjukkan

bahwa sebagian besar pelanggan memberikan respons yang baik terhadap produk atau layanan, seperti yang ditunjukkan pada grafik di atas. Sebaliknya, 49 ulasan menunjukkan bahwa pelanggan memberikan ulasan tanpa banyak emosi, baik positif maupun negatif, dan paling sedikit ulasan negatif, 18 ulasan, menunjukkan bahwa sebagian kecil pelanggan yang tidak puas. Secara keseluruhan, distribusi ini menunjukkan persepsi pelanggan yang cenderung positif terhadap produk atau layanan yang dievaluasi, dengan beberapa ulasan netral dan sedikit ulasan negatif. Hasil penelitian ini dengan metode Elbow menunjukkan bahwa jumlah *cluster* optimal adalah 3.

Grafik 2. Hasil Analisis Metode Elbow

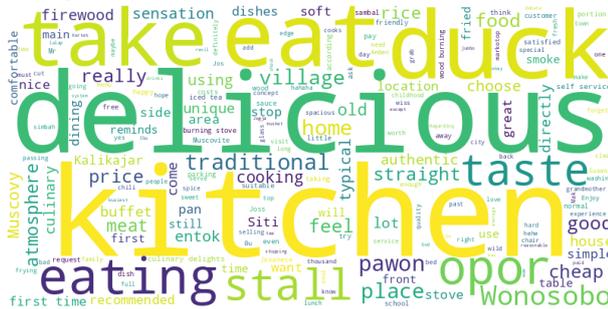


Sumber : hasil olah data metode elbow, 2024

Grafik di atas menunjukkan hasil analisis yang dilakukan menggunakan metode *Elbow* untuk menentukan jumlah kluster yang ideal untuk algoritma *K-Means*. Sumbu x menunjukkan jumlah kluster yang diuji, dan sumbu y menunjukkan nilai WCSS, yang merupakan jumlah variansi data dalam kluster. Nilai WCSS menurun seiring dengan jumlah kluster, yang menunjukkan bahwa data semakin terkonsentrasi di dalam kluster. Namun, setelah jumlah kluster menjadi 3, penurunan nilai WCSS menjadi kurang signifikan. Karena grafik membentuk "siku", titik ini sering disebut sebagai titik siku. Berdasarkan temuan ini, jumlah kluster ideal untuk penelitian ini adalah 3, karena pada titik ini tercapai keseimbangan antara jumlah kluster dan variasi data yang dijelaskan. Pada tahap berikutnya, jumlah kluster ini digunakan untuk proses *clustering*.

Hasil dari tabel konten, yang mengambil data dari tahap *stemming*, diproses ke dalam perhitungan *K-Means Clustering* untuk menentukan kumpulan kalimat dari data yang diolah masuk ke dalam *cluster* mana. Kumpulan data pengujian ditampilkan dengan masing-masing kalimat memasuki *cluster* yang sesuai dengan hasil algoritma *K-Means Clustering*. Selanjutnya, dibuat sebuah *wordcloud* untuk menampilkan sebaran kata dari data komentar sesuai dengan masing-masing *cluster*.

Gambar 3. Hasil Visualisasi Kluster 0



Sumber : hasil data olah phyton, 2024

Kluster 0, Fokus pada tempat: Kata-kata seperti *kitchen*, *delicious*, *duck*, dan *traditional* mendominasi kluster ini, yang menunjukkan bahwa ulasan dalam kluster ini cenderung berfokus pada apresiasi pelanggan terhadap suasana tradisional dan pengalaman makan yang unik.

Gambar 4. Hasil Visualisasi Kluster 1



Sumber : hasil data olah phyton, 2024

Kluster 1 Fokus pada kualitas produk: Kata "*delicious*" sangat dominan dalam kluster ini, diikuti oleh "*sauce*", "*chili*", dan "*duck*". Ini menunjukkan bahwa ulasan pelanggan dalam kluster ini terutama berfokus pada rasa makanan, kualitas bahan, dan kelezatan menu, terutama yang berkaitan dengan saus dan

sambal. Kata-kata seperti benar-benar dan makanan menunjukkan rasa yang menyenangkan.

Gambar 5. Hasil Visualisasi Kluster 2



Sumber : hasil data olah phyton, 2024

Kluster 2 Fokus pada Harga dan Produk: Kluster ini mencakup kata-kata seperti *price*, *rice*, *muscovy*, dan *duck*. Ini menunjukkan bahwa ulasan dalam kluster ini lebih cenderung membahas harga dan porsi, serta komponen tertentu dari menu, seperti nasi dan lauk-pauk. Perhatian konsumen terhadap nilai yang mereka dapatkan saat membeli barang atau layanan juga ditunjukkan dengan kata-kata seperti *ribu* dan *satu*.

Hasil analisis ini menunjukkan bahwa persepsi pelanggan dapat dikelompokkan ke dalam tiga tema utama: pengalaman budaya dan tradisional, kualitas rasa makanan, serta aspek harga dan porsi. Temuan ini memberikan gambaran bahwa strategi pemasaran dapat diarahkan sesuai dengan fokus utama di setiap kluster. Misalnya, aspek tradisional dapat diperkuat dalam materi promosi, sementara kualitas rasa dapat menjadi elemen utama dalam menjaga kepuasan pelanggan. Selain itu, aspek harga dapat disesuaikan untuk memenuhi ekspektasi segmen pelanggan yang lebih sensitif terhadap biaya.

6.Kesimpulan

Studi ini dapat menemukan persepsi pelanggan secara lebih terstruktur dengan menggunakan algoritma *K-Means* dan analisis sentimen. Sebagian besar ulasan adalah positif, dengan fokus pada lokasi, kualitas produk, dan harga. Hasil menunjukkan bahwa Warung Enthok Ibu Siti Wonosobo telah berhasil memenuhi harapan mayoritas pelanggan,

tetapi masih ada beberapa hal yang perlu diperbaiki.

Saran untuk UMKM Hasil penelitian ini menunjukkan bahwa umkm dapat menggunakan analisis sentimen untuk menilai kualitas produk dan layanan mereka. Fokusnya pada elemen yang disoroti konsumen seperti kualitas produk, pengalaman suasana tradisional, dan keseimbangan antara harga dan nilai yang diterima konsumen diyakini dapat membantu umkm dalam memperkuat pasar. Untuk meningkatkan citra merk, Umkm harus memaksimalkan platform digital untuk melacak dan merespon ulasan pelanggan. Umkm harus lebih kompetitif di tengah persaingan pasar yang dinamis dengan melakukan evaluasi berkala pada sentimen pelanggan untuk mengidentifikasi perubahan tren dan preferensi.

Kami menyarankan peneliti yang akan dapat memperluas cakupan data dari berbagai *platform digital*, bereksperimen dengan algoritma lain seperti DBSCAN, serta menambahkan analisis emosi guna mendapatkan hasil yang lebih komprehensif. Studi longitudinal juga penting untuk memahami dinamika perubahan persepsi konsumen.

DAFTAR PUSTAKA

- Azmy, M. S., & Furinto, A. (2008). Analisa Faktor-Faktor Yang Mempengaruhi Persepsi Konsumen Dalam Pemilihan Biro Perjalanan Umroh Dan Haji Khusus. *Journal of Business Strategy and Execution*, 1(1), 171–193.
- Bastian, A., Sujadi, H., & Febrianto, G. (2018). *Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)*.
- Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134. <https://doi.org/https://doi.org/10.1016/j.knosys.2021.107134>
- Bogers, M., Zobel, A.-K., Afuah, A., Almirall, E., Brunswicker, S., Dahlander, L., Frederiksen, L., Gawer, A., Gruber, M., Haefliger, S., Hagedoorn, J., Hilgers, D., Laursen, K., Magnusson, M. G., Majchrzak, A., McCarthy, I. P., Moeslein, K. M., Nambisan, S., Piller, F. T., ... Ter Wal, A. L. J. (2017). The open innovation research landscape: established perspectives and emerging themes across different levels of analysis. *Industry and Innovation*, 24(1), 8–40. <https://doi.org/10.1080/13662716.2016.1240068>
- Borrego, Á., & Comalat Navarra, M. (2021). What users say about public libraries: an analysis of Google Maps reviews. *Online Information Review*, 45(1), 84–98.
- Chaturvedi, I., Cambria, E., Welsch, R. E., & Herrera, F. (2018). Distinguishing between facts and opinions for sentiment analysis: Survey and challenges. *Information Fusion*, 44, 65–77.
- Chu, X. (2019). *Data Cleaning*.
- Chu, X., Ilyas, I. F., Krishnan, S., & Wang, J. (2016). Data cleaning: Overview and emerging challenges. *Proceedings of the 2016 International Conference on Management of Data*, 2201–2206.
- Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of price, brand, and store information on buyers' product evaluations. *Journal of Marketing Research*, 28(3), 307–319.
- Febriani, A., & Putri, S. A. (2020). Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means. *Jiems (Journal of Industrial Engineering and Management Systems)*, 13(2).
- Gallarza, M. G., Arteaga, F., Del Chiappa, G., Gil-Saura, I., & Holbrook, M. B. (2017). A multidimensional service-value scale based on Holbrook's typology of customer value: Bridging the gap between the concept and its measurement. *Journal of Service Management*, 28(4), 724–762.
- García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining* (Vol. 72). Springer.
- Godara, R. S., Yadav, D., Sagar, M., & Disari, R. (2024). Impact of customer reviews on purchase decision of a brand: A study of online shopping. *International Research*

- Journal of Economics and Management Studies IRJEMS*, 3(4).
- Handayati, R., Rozaq, A. S., As' adi, F. N., Saputra, M. Z. D., & Ubaidillah, M. M. (2024). Challenges and Solutions in the Implementation of Digital Technology in MSMEs: Perspectives from Entrepreneurs and Consumers. *Technology and Society Perspectives (TACIT)*, 2(2), 244–251.
- Hendrawan, S. A., Chatra, A., Iman, N., Hidayatullah, S., & Suprayitno, D. (2024). Digital transformation in MSMEs: Challenges and opportunities in technology management. *Jurnal Informasi Dan Teknologi*, 141–149.
- Huang, C., Chen, C., & Wang, H. (2024). Effects of online customer reviews on sustainable clothing purchase intentions: The mediating role of perceived diagnosticity. *Journal of Consumer Behaviour*.
- Huang, Z., Zheng, H., Li, C., & Che, C. (2024). Application of Machine Learning-Based K-means Clustering for Financial Fraud Detection. In *Academic Journal of Science and Technology* (Vol. 10, Issue 1).
- Jaiswal, S., & Singh, A. (2020). Influence of the determinants of online customer experience on online customer satisfaction. *Paradigm*, 24(1), 41–55.
- Kumar, V., & Reinartz, W. (2016). Creating enduring customer value. *Journal of Marketing*, 80(6), 36–68.
- Leroi-Werelds, S., Streukens, S., Brady, M. K., & Swinnen, G. (2014). Assessing the value of commonly used methods for measuring customer value: A multi-setting empirical study. *Journal of the Academy of Marketing Science*, 42, 430–451.
- Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal of King Saud University - Computer and Information Sciences*, 36(4), 102048. <https://doi.org/https://doi.org/10.1016/j.ksuci.2024.102048>
- Mihu, C., Pitic, A., & Bayraktar, D. (2023). Drivers of Digital Transformation and their Impact on Organizational Management. *Studies in Business and Economics*, 18, 149–170. <https://doi.org/10.2478/sbe-2023-0009>
- Mishra, P., Biancolillo, A., Roger, J. M., Marini, F., & Rutledge, D. N. (2020). New data preprocessing trends based on ensemble of multiple preprocessing techniques. *TrAC Trends in Analytical Chemistry*, 132, 116045. <https://doi.org/https://doi.org/10.1016/j.trac.2020.116045>
- Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management. *MIS Quarterly*, 41(1), 223–238.
- Python, W. (2021). Python. *Python Releases for Windows*, 24.
- Sholeh, B., Indarwati, T. A., Kautsar, A., Paramita, R. A. S., Indriastuty, N., Nugrahanti, R. B., & Oktavia, W. D. (2024). Scientometrics Approach: Online Review Customer and Purchase Decision. *International Journal Of Economics, Management, Business And Social Science(IJEMBIS)*, 4(2), 841–857.
- Swanson, S. R., & Davis, J. C. (2020). GEN Z's Conception of Marketing. *Journal for Advancement of Marketing Education*, 28(1).
- Syahdan, R., Khairun, U., Pertamina Kampus, J. I., & Selatan, T. (2021). Factors influencing Gen-Z customer perception online shopping in Indonesia. In *International Journal of Communication and Society* (Vol. 3, Issue 2). <http://pubs2.ascee.org/index.php/IJCSIJCS@ascee.org>
- Thomas, D. M., & Mathur, S. (2019). Data analysis by web scraping using python. *2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 450–454.
- Tri Jaka, A. H. (2015). *Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining*.
- Xu, R., & Wunsch, D. (2008). *Clustering*. John Wiley & Sons.

- Yadav, J., & Sharma, M. (2013). A Review of K-mean Algorithm. *Int. J. Eng. Trends Technol*, 4(7), 2972–2976.
- Zeithaml, V. A., Verleye, K., Hatak, I., Koller, M., & Zauner, A. (2020). Three decades of customer value research: paradigmatic roots and future research avenues. *Journal of Service Research*, 23(4), 409–432.