

**STRATEGI PENINGKATAN SERTIFIKASI HALAL
BERDASARKAN ANALISIS SENTIMEN DAN KLASSTER
KONSUMEN INDONESIA MENGGUNAKAN *TEXT MINING* DAN
K-MEANS CLUSTERING PADA MEDIA SOSIAL X**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-1
Program Studi Teknik Industri - Fakultas Teknologi Industri
Universitas Islam Indonesia**



Nama : Manda Aprilia
No. Mahasiswa : 22522058

**PROGRAM STUDI TEKNIK INDUSTRI PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2026**

PERNYATAAN KEASLIAN

PERNYATAAN KEASLIAN

Saya mengakui bahwa tugas akhir ini adalah hasil karya saya sendiri kecuali kutipan dan ringkasan yang seluruhnya sudah saya jelaskan sumbernya. Jika dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah maka saya bersedia ijazah yang telah saya terima ditarik kembali oleh Universitas Islam Indonesia.

Yogyakarta, 16 April 2026



(Manda Aprilia)
22522058

SURAT BUKTI PENELITIAN



FAKULTAS
TEKNOLOGI INDUSTRI

Gedung KH. Mas Mansur
Kampus Terpadu Universitas Islam Indonesia
Jl. Gallurang km 14,5 Yogyakarta 55584
T. (0274) 898444 ext. 4100, 4101
F. (0274) 895307
E. fti@uii.ac.id
W. fti.uui.ac.id

SURAT KETERANGAN PENELITIAN

Nomor: 012/Ka.Lab.Datmin/70/Lab.Datmin/V/2026

Assalamu'alaikum Warahmatullahi Wabarakatuh

Kami yang bertanda tangan dibawah ini, menerangkan bahwa mahasiswa dengan keterangan sebagai berikut :

Nama : Manda Aprilia
No. Mhs : 22522058
Dosen Pembimbing : Dr. Harwati, S.T., M.T.

Telah selesai melaksanakan penelitian yang berjudul "STRATEGI PENINGKATAN SERTIFIKASI HALAL BERDASARKAN ANALISIS SENTIMEN DAN KLASSTER KONSUMEN INDONESIA MENGGUNAKAN TEXT MINING DAN K-MEANS CLUSTERING PADA MEDIA SOSIAL X" di Laboratorium Data Mining, Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia tercatat mulai tanggal 20 November 2025 sampai dengan tanggal 5 Januari 2026.

Demikian surat keterangan kami keluarkan, agar dapat dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Yogyakarta, 18 Mei 2026

Kepala Laboratorium
Data Mining

Dr. Dwi Adi Purnama, S.T.

LEMBAR PENGESAHAN PEMBIMBING

**STRATEGI PENINGKATAN SERTIFIKASI HALAL BERDASARKAN
ANALISIS SENTIMEN DAN KLASSTER KONSUMEN INDONESIA
MENGUNAKAN *TEXT MINING* DAN *K-MEANS CLUSTERING* PADA
MEDIA SOSIAL X**



Yogyakarta, 13 Mei 2026

Dosen Pembimbing



(Dr. Harwati, S.T., M.T.)

LEMBAR PENGESAHAN DOSEN PENGUJI

**STRATEGI PENINGKATAN SERTIFIKASI HALAL BERDASARKAN
ANALISIS SENTIMEN DAN KLASSTER KONSUMEN INDONESIA
MENGUNAKAN *TEXT MINING* DAN *K-MEANS CLUSTERING* PADA
MEDIA SOSIAL X**

TUGAS AKHIR**Disusun Oleh :****Nama : Manda Aprilia****No. Mahasiswa : 22522058**

**Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk
memperoleh gelar Sarjana Strata-1 Teknik Industri Fakultas Tekonologi Industri
Universitas Islam Indonesia**

Yogyakarta, 25 Mei 2026**Tim Penguji**Dr. Harwati, S.T., M.T.

Ketua

Ir. Muchamad Sugarindra, S.T., M.T.I., IPM.

Anggota I

Didin Dwi Novianto, S.T., M.LSCM.

Anggota II

Mengetahui,

**Ketua Program Studi Teknik Industri Program Sarjana
Fakultas Teknologi Industri
Universitas Islam Indonesia**

(Ir. Muhammad Bilwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM)

015220101

HALAMAN PERSEMBAHAN

Alhamdulillahirabbilalamiin, segala puji dan Syukur penulis haturkan kepada Allah SWT atas segala nikmat, karunia, dan bantuannya akhirnya saya dapat menyelesaikan tugas akhir ini dengan baik. Tugas akhir ini penulis persembahkan kepada diri sendiri, yang telah berjuang dan berusaha selama masa perkuliahan yang tidak mudah ini. Penulis juga mengucapkan banyak terima kasih kepada ayah dan ibu, serta kakak-kakak penulis yang telah memberikan support dan doa sehingga penulis dapat bertahan selama masa perkuliahan. Tak lupa penulis juga mengucapkan terima kasih kepada Ibu Dr. Harwati, S.T., M.T. atas bimbingan dan ilmu yang diberikan selama proses penelitian ini.

MOTTO

“Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan. Maka apabila engkau telah selesai (dari suatu urusan), tetaplah bekerja keras (untuk urusan yang lain). Dan hanya kepada Tuhanmulah engkau berharap”
(Q.S Al Insyirah: 5-8)

“Orang lain gak akan paham *struggle* dan masa sulit nya kita yang mereka ingin tahu hanya bagian *success stories*. Berjuanglah untuk diri sendiri walaupun tidak ada tepuk tangan. Kelak diri kita dimasa depan akan sangat banga dengan apa yang kita perjuangkan hari ini”
(Fardi Yandi)

KATA PENGANTAR

Assalamu 'alaikum Warrahmatullahi Wabarakatuh.

Alhamdulillahirabbil'amin, puji dan syukur penulis panjatkan kepada Allah SWT, yang telah melimpahkan nikmat-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir dengan baik. Sholawat serta salam juga senantiasa tercurah kepada Nabi Muhammad SAW, keluarga, sahabat, dan para pengikut beliau yang telah berjuang membimbing umat manusia dari kegelapan menuju jalan terang demi meraih ridha Allah SWT. Berkat rahmat dan karunia-Nya, penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Strategi Peningkatan Sertifikasi Halal Berdasarkan Analisis Sentimen dan Klaster Konsumen Indonesia Menggunakan Text Mining dan K-Means Clustering pada Media Sosial X”** dengan baik.

Tugas Akhir ini merupakan salah satu syarat untuk meraih gelar Strata Satu (S1) di Jurusan Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia. Tujuan dari Tugas Akhir ini adalah untuk mengintegrasikan pengetahuan yang telah diperoleh selama perkuliahan. Selama proses pelaksanaan Tugas Akhir, penulis mendapatkan banyak bantuan, dukungan, dan kesempatan dari berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo, M.T., IPU., ASEAN.Eng. selaku Dekan Fakultas Teknologi Industri, Universitas Islam Indonesia.
2. Bapak Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM., selaku Ketua Program Studi Teknik Industri, Universitas Islam Indonesia.
3. Ibu Dr. Harwati, S.T., M.T. selaku dosen pembimbing laporan Tugas Akhir yang telah memberikan bimbingan serta arahan, motivasi, dan ilmu yang telah beliau berikan sehingga penulis dapat menyelesaikan Tugas Akhir ini.
4. Kepada seluruh keluarga besar, khususnya kedua orang tua tercinta, Bapak Sukisdi dan Ibu Sutini, yang senantiasa memberikan kasih sayang dan doa di setiap langkah. Selain itu, kepada kedua kakak saya, Wahyu Setiawan dan Popi Krisdiani, serta kakak ipar saya Shella Fauziah Swandani, yang terus memberikan dukungan dan semangat tanpa henti. Tidak lupa kepada kedua keponakan saya, Fabian Kenzo Setiawan dan Almeera Keysha Syahira, terima kasih atas keceriaan kalian yang selalu membangkitkan semangat dan kebahagiaan penulis sehingga mampu menyelesaikan skripsi ini hingga selesai.
5. Teman-teman seperjuangan Teknik Industri angkatan 2022, khususnya Arista Prabaningrum, Zafira Ratna Dewi Nastiti, Rifa Damayanti, Nasha Fazilatun Nisa, Insan Akhsani, dan Farah Nur Annisa yang telah menjadi rekan sejak awal perkuliahan. Selain itu, Zahra Alyssa Larasati, Annisa Nathaniela Syahda, dan Stevia Jauhara Zain senantiasa memberikan dukungan selama proses penyusunan Tugas Akhir.
6. Untuk diri saya, Manda Aprilia, terima kasih telah bertahan dan tetap kuat hingga saat ini. Terima kasih karena mampu mengendalikan diri di tengah berbagai tekanan, tidak menyerah dalam menghadapi kesulitan perkuliahan maupun proses penyusunan skripsi, serta tetap tegar saat dihadapkan pada berbagai permasalahan. Teruslah melangkah dengan rendah hati, karena ini baru permulaan dari perjalanan hidup. Tetap semangat, kamu pasti mampu melewati setiap tahap yang akan datang.
7. Semua pihak yang tidak dapat dituliskan satu persatu, yang telah terlibat dan memberikan dukungan dalam penyelesaian penulisan Tugas Akhir.

Penulis menyadari bahwa Tugas Akhir ini masih belum sempurna, sehingga penulis

menyampaikan permohonan maaf serta mengharapkan kritik dan saran dari semua pembaca guna penyempurnaan dimasa mendatang. Semoga Tugas Akhir ini dapat bermanfaat bagi semua pihak.

Wassalamu'alaikum Warrahmatullahi Wabarakatuh.

Yogyakarta, 16 April 2026

A handwritten signature in black ink, appearing to read 'Manda Aprilia', with a long horizontal stroke extending to the right.

Manda Aprilia
NIM 22522058

ABSTRAK

Pemberlakuan kewajiban sertifikasi halal di Indonesia telah mendorong peningkatan jumlah produk yang tersertifikasi halal. Namun demikian, implementasinya masih menghadapi berbagai kendala, antara lain kompleksitas prosedur, persepsi terhadap biaya, serta beragamnya respons konsumen. Persepsi tersebut tercermin dalam opini masyarakat yang disampaikan melalui media sosial X. Penelitian ini bertujuan untuk menganalisis sentimen, mengelompokkan karakteristik konsumen, serta merumuskan strategi peningkatan sertifikasi halal. Metode yang digunakan dalam penelitian ini meliputi text mining, analisis sentimen, dan *K-Means Clustering*. Hasil penelitian menunjukkan bahwa sentimen masyarakat didominasi oleh sentimen netral, dengan keberadaan sentimen positif dan negatif. Proses pengelompokan menghasilkan beberapa kluster dengan karakteristik yang berbeda, sedangkan faktor utama yang memengaruhi persepsi meliputi aspek prosedur, biaya, dan penyebaran informasi. Berdasarkan hasil tersebut, dirumuskan strategi peningkatan sertifikasi halal yang lebih efektif dan tepat sasaran.

Kata Kunci: Sertifikasi Halal, Analisis Sentimen, *K-Means Clustering*, *Text Mining*

DAFTAR ISI

PERNYATAAN KEASLIAN.....	i
SURAT BUKTI PENELITIAN	ii
LEMBAR PENGESAHAN PEMBIMBING	iii
LEMBAR PENGESAHAN DOSEN PENGUJI.....	iv
HALAMAN PERSEMBAHAN	v
MOTTO.....	vi
KATA PENGANTAR.....	vii
ABSTRAK.....	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
BAB I1	
PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	6
1.3 Tujuan Penelitian	7
1.4 Manfaat Penelitian	7
1.5 Batasan Penelitian.....	8
BAB II.....	9
TINJAUAN PUSTAKA.....	9
2.1 Kajian Literatur.....	9
2.2 Landasan Teori	16
2.2.1 Analisis Sentimen	16
2.2.2 Media Sosial X.....	16
2.2.3 Text Mining.....	17
2.2.4 Preprocessing.....	17
2.2.5 Sertifikasi Halal.....	19
2.2.6 Clustering.....	20
2.2.7 Algoritma K-Means	20
BAB III.....	22
METODE PENELITIAN.....	22
3.1 Objek dan Subjek Penelitian.....	22
3.2 Jenis Data.....	22
3.3 Metode Pengumpulan Data.....	22
3.4 Alur Penelitian	23
BAB IV.....	27
PENGUMPULAN DAN PENGOLAHAN DATA	27
4.1 Pengumpulan Data.....	27
4.2 Pengolahan Data	29
4.2.1. Analisis Deskriptif.....	30

4.2.2.	Processing Data	30
4.2.3.	Distribusi Sentimen	35
4.2.4.	Representasi Teks Menggunakan TF-IDF.....	36
4.2.5.	Penentuan Jumlah Cluster dengan Metode Elbow.....	37
4.2.6.	Visualisasi Cluster	38
BAB V		41
PEMBAHASAN		41
5.1	Analisis Hasil Sentimen.....	41
5.2	Analisis Hasil Cluster	44
5.3	Strategi Berdasarkan <i>Cluster</i>	45
5.4	Limitasi Penelitian	47
BAB VI		49
PENUTUP		49
6.1	Kesimpulan	49
6.2	Saran	50
DAFTAR PUSTAKA		52
LAMPIRAN		A

DAFTAR TABEL

Tabel 1. 1 Jumlah Produk Bersertifikasi Halal Tahun 2011-2025	2
Tabel 2. 1 Kajian Literatur.....	14
Tabel 4. 1 Proses <i>Case Folding</i>	31
Tabel 4. 2 Proses <i>Tokenizing</i>	32
Tabel 4. 3 Proses <i>Filtering</i>	33
Tabel 4. 4 Proses <i>Stemming</i>	34
Tabel 4. 5 Sampel Distribusi Fitur Sentimen.....	36
Tabel 5. 1 Strategi Peningkatan Sertifikasi Halal	46

DAFTAR GAMBAR

Gambar 1. 1 Jumlah Produk Bersertifikasi Halal Tahun 2011-2025.....	3
Gambar 3. 1 Alur Penelitian	23
Gambar 4. 1 Aplikasi X	27
Gambar 4. 2 Tanggapan Masyarakat Mengenai Sertifikasi Halal di X	28
Gambar 4. 3 Hasil Proses <i>Scraping Data</i>	29
Gambar 4. 4 Grafik <i>Clustering</i> Data Sertifikasi Halal.....	35
Gambar 4. 5 Metode <i>Elbow</i>	37
Gambar 4. 6 Visualisasi <i>Wordcloud Cluster 0</i>	38
Gambar 4. 7 Visualisasi <i>Wordcloud 1</i>	39
Gambar 4. 8 Visualisasi <i>Wordcloud</i>	40
Gambar 5. 1 Bigram Sentimen Negatif	41
Gambar 5. 2 Bigram Sentimen Netral	42
Gambar 5. 3 Bigram Sentimen Positif	43
Gambar 5. 4 Hasil <i>Clustering</i> dan Visualisai Diagram Lingkaran	44

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era globalisasi ekonomi yang semakin terintegrasi, kesadaran terhadap pentingnya produk halal mengalami peningkatan yang signifikan, termasuk di Indonesia. Konsep halal tidak hanya terbatas pada makanan dan minuman, tetapi juga mencakup berbagai sektor lain seperti kosmetik, farmasi, pakaian, pariwisata, keuangan, hingga industri fashion. Seluruh rangkaian proses produksi, mulai dari penyediaan bahan baku, proses pengolahan, hingga pendistribusian produk, dituntut untuk memenuhi ketentuan syariat Islam (B. shafira Hidayat, 2024). Selain sebagai kewajiban agama, konsep halal juga merepresentasikan aspek keamanan, kesehatan, dan keberlanjutan yang semakin menjadi perhatian konsumen global, terutama masyarakat Muslim yang menjadikan halal sebagai bagian dari gaya hidup.

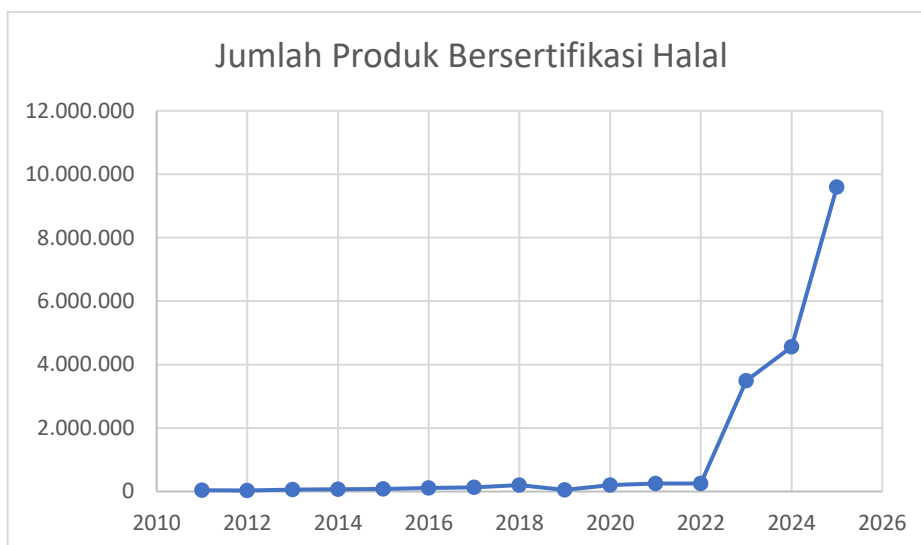
Indonesia merupakan negara dengan populasi Muslim terbesar di dunia, yaitu sekitar 229,62 juta jiwa atau 87,2% dari total penduduk, serta berkontribusi sebesar 13,1% terhadap populasi Muslim global (Humaida et al., 2024). Kondisi ini menempatkan Indonesia pada posisi strategis dalam pengembangan industri halal dunia, sehingga jaminan kehalalan produk menjadi kebutuhan esensial sekaligus bentuk perlindungan konsumen untuk memastikan keamanan dan kepastian produk yang dikonsumsi. Namun, sebelum diberlakukannya Undang-Undang Nomor 33 Tahun 2014 tentang Jaminan Produk Halal, pelaksanaan sertifikasi halal di Indonesia masih bersifat sukarela (*voluntary*) dan belum memberikan kepastian hukum yang kuat bagi konsumen Muslim. Pada masa tersebut, sertifikasi dan pelabelan halal belum memiliki dasar legitimasi hukum yang memadai (Legowati & Albab, 2019).

Menurut Pusat Penelitian dan Pengembangan Bimas Agama dan Layanan Keagamaan tahun 2018 terhadap 1.200 responden menunjukkan bahwa intensi masyarakat Muslim Indonesia dalam menggunakan produk halal berada pada kategori sedang dengan jumlah responden sebanyak 843 atau hasil persentase 70,1 persen. Sebanyak 168 responden memiliki intensi membeli rendah yakni 14,0 persen dan sebanyak 191 orang berada pada kategorisasi intensi membeli produk halal tinggi atau sejumlah 15,9 persen. Bagi pelaku Usaha Mikro Kecil dan Menengah (UMKM), studi Badan Litbang dan Diklat Kementerian Agama RI tahun 2016 mencatat indeks kemauan

UMKM untuk melakukan sertifikasi halal berada pada kategori cukup (skor 67,1 dari 100), dengan kendala utama berupa biaya sertifikasi yang dianggap memberatkan. Meski demikian, data LPPOM MUI menunjukkan adanya tren peningkatan kesadaran pelaku usaha dalam mendaftarkan dan mensertifikasi produk halal (Badan Penyelenggara Jaminan Produk Halal, 2020). Berikut merupakan jumlah pelaku usaha yang meregistrasi jumlah produk bersertifikasi halal tahun 2011-2025 dapat dilihat pada tabel berikut:

Tabel 1. 1 Jumlah Produk Bersertifikasi Halal Tahun 2011-2025

Tahun	Jumlah Produk Bersertifikasi Halal
2011	39.002
2012	32.890
2013	64.121
2014	68.576
2015	77.256
2016	114.264
2017	127.286
2018	204.222
2019	50.000
2020	200.000
2021	250.000
2022	249.971
2023	3.494.693
2024	4.556.892
2025	9.600.000
TOTAL	19.129.173



Gambar 1. 1 Jumlah Produk Bersertifikasi Halal Tahun 2011-2025

Jumlah produk bersertifikat halal di Indonesia periode 2011–2025 menunjukkan peningkatan yang sangat signifikan, terutama setelah diberlakukannya kebijakan sertifikasi halal secara wajib (*mandatory*). Pada periode 2011–2018, jumlah produk bersertifikat halal masih relatif terbatas, yakni dari 39.002 produk (2011) menjadi 204.222 produk (2018). Kondisi ini terjadi karena sertifikasi halal masih bersifat sukarela sehingga belum mengikat seluruh pelaku usaha.

Perubahan besar terjadi setelah diberlakukannya Undang-Undang Nomor 33 Tahun 2014 tentang Jaminan Produk Halal dan implementasinya oleh BPJPH. Sejak 2023, jumlah produk bersertifikat halal melonjak drastis menjadi 3.494.693 produk, meningkat pada 2024 menjadi 4.556.892 produk (Irham, 2024), dan mencapai 9.600.000 produk pada 2025 (Hasan, 2025). Secara kumulatif, periode 2011–2025 mencatat 19.129.173 produk telah bersertifikat halal. Peningkatan ini sejalan dengan posisi Indonesia sebagai negara dengan populasi Muslim terbesar di dunia, sehingga kebutuhan akan jaminan kehalalan produk menjadi sangat penting, baik sebagai bentuk perlindungan konsumen maupun sebagai strategi penguatan daya saing industri halal nasional. Transformasi dari sistem sukarela ke sistem wajib menunjukkan komitmen pemerintah dalam memperkuat kepastian hukum, meningkatkan kesadaran pelaku usaha, serta mendorong pertumbuhan ekosistem halal secara nasional dan global.

Meskipun demikian, implementasi kewajiban sertifikasi halal di Indonesia masih menghadapi berbagai kendala. Sejumlah penelitian menunjukkan bahwa pelaku usaha, khususnya UMKM, masih mengalami hambatan dalam proses sertifikasi halal akibat keterbatasan informasi, kompleksitas prosedur, serta persepsi biaya yang tinggi

(Hikmah et al., 2025). Di sisi lain, tingkat pemahaman, persepsi, dan sikap konsumen terhadap kebijakan sertifikasi halal menunjukkan variasi yang cukup beragam. Sebagian konsumen mendukung kebijakan ini karena dinilai mampu meningkatkan jaminan keamanan dan kepercayaan terhadap produk, sementara sebagian lainnya bersikap kritis akibat kekhawatiran terhadap potensi kenaikan harga, tambahan beban administrasi, serta dampaknya terhadap keberlanjutan UMKM. Perbedaan persepsi tersebut tercermin dalam opini publik di media sosial, khususnya X, yang menjadi ruang bagi masyarakat untuk menyampaikan pandangan, kritik, serta diskursus terbuka terkait kebijakan sertifikasi halal.

Persepsi konsumen menjadi faktor yang penting karena memengaruhi sikap, yang pada akhirnya menentukan perilaku konsumsi, termasuk keputusan untuk membeli atau menghindari produk bersertifikat halal. Ketika sertifikasi halal dipandang sebagai jaminan kualitas dan nilai tambah produk, konsumen cenderung menunjukkan sikap positif dan meningkatkan tingkat kepercayaan terhadap produk UMKM. Sebaliknya, apabila kebijakan ini dipersepsikan sebagai kebijakan yang memberatkan pelaku usaha atau berpotensi meningkatkan harga produk, sikap konsumen dapat bergeser menjadi skeptis bahkan resistif. Pola ini terlihat dalam respons pengguna X, di mana sebagian warganet mendukung sertifikasi halal sebagai bentuk perlindungan konsumen dan kepastian hukum, sementara sebagian lainnya mempertanyakan efektivitas kebijakan, transparansi proses, serta kesiapan UMKM dalam pelaksanaannya.

Kondisi tersebut menunjukkan bahwa keberhasilan kebijakan sertifikasi halal tidak hanya ditentukan oleh aspek regulasi dan pencapaian target kuantitatif, tetapi juga sangat bergantung pada bagaimana kebijakan tersebut dipersepsikan dan diterima oleh konsumen sebagai pengguna akhir produk halal. Seiring meningkatnya penggunaan media sosial, khususnya X, masyarakat semakin terbuka dalam menyampaikan pandangan terkait kewajiban sertifikasi halal, sehingga platform ini menjadi sumber data yang strategis untuk memahami persepsi publik secara real time dan berskala luas. Penelitian (Timur et al., 2023) menunjukkan bahwa sentimen terhadap isu halal di X didominasi oleh sikap netral, namun tetap disertai sentimen positif dan negatif yang signifikan, terutama berkaitan dengan kemudahan proses sertifikasi dan tingkat kepercayaan terhadap lembaga halal, yang berpotensi memengaruhi sikap serta penerimaan konsumen. Kondisi tersebut sejalan dengan temuan bahwa konsumen memiliki persepsi yang beragam, mulai dari dukungan terhadap jaminan kehalalan

produk dan perlindungan konsumen hingga kekhawatiran mengenai biaya sertifikasi, kompleksitas prosedur, serta kesiapan pelaku usaha, khususnya UMKM, sehingga membentuk opini publik yang dinamis terhadap kebijakan sertifikasi halal di Indonesia (Azis, 2025). Hal ini menunjukkan bahwa peningkatan jumlah sertifikasi halal secara kuantitatif belum sepenuhnya menjamin efektivitas implementasi kebijakan, sehingga diperlukan strategi peningkatan yang berbasis pada segmentasi konsumen dan analisis akar permasalahan secara komprehensif.

Meskipun capaian sertifikasi halal terus meningkat secara kuantitatif, keberlanjutan dan efektivitas implementasinya belum sepenuhnya optimal karena masih adanya variasi persepsi dan respons publik terhadap kebijakan tersebut. Tanpa perumusan strategi yang berbasis pada karakteristik dan segmentasi konsumen serta identifikasi akar permasalahan yang memengaruhi persepsi masyarakat, peningkatan sertifikasi halal berpotensi tidak berjalan secara optimal, terarah, dan berkelanjutan. Namun demikian, kajian yang menganalisis respons publik di media sosial secara sistematis dengan pendekatan analisis data digital masih relatif terbatas. Sebagian penelitian sebelumnya masih bersifat deskriptif dan belum memanfaatkan potensi data teks dalam jumlah besar yang dihasilkan dari media sosial (Izza & Akbar, 2023). Padahal, metode text mining memungkinkan proses ekstraksi informasi dan identifikasi kecenderungan sentimen dari data teks secara otomatis dan efisien. Selain itu, pengelompokan opini publik berdasarkan kesamaan karakteristik dan sentimen dapat dilakukan menggunakan metode K-Means clustering, yang telah banyak digunakan dalam penelitian untuk mengelompokkan data teks berdasarkan pola kemiripan tertentu (Syaifudin & Irawan, 2018). Penerapan metode ini pada data X dapat membantu memetakan kelompok konsumen berdasarkan sikap, kekhawatiran, maupun dukungan terhadap implementasi sertifikasi halal. Namun, sebagian besar penelitian sebelumnya masih berfokus pada pengukuran sentimen secara umum dan belum mengintegrasikan analisis lanjutan untuk mengidentifikasi faktor penyebab utama yang memengaruhi persepsi masyarakat, sehingga hasil penelitian belum sepenuhnya mendukung perumusan strategi kebijakan yang spesifik dan berbasis segmentasi.

Berdasarkan kondisi tersebut, penelitian ini mengombinasikan metode *text mining*, analisis sentimen, dan *K-Means clustering* untuk menganalisis karakteristik konsumen Indonesia berdasarkan persepsi mereka terhadap implementasi kewajiban sertifikasi halal di media sosial X. Metode *K-Means Clustering* digunakan untuk

mengelompokkan konsumen ke dalam segmen-segmen berdasarkan kesamaan pola sentimen dan karakteristik persepsi. Melalui kombinasi metode tersebut, penelitian ini tidak hanya menghasilkan pemetaan karakteristik konsumen, tetapi juga memberikan pemahaman yang lebih mendalam mengenai penyebab munculnya persepsi negatif maupun positif terhadap kebijakan sertifikasi halal. Hasil penelitian ini diharapkan dapat menjadi dasar penyusunan rekomendasi strategis yang berbasis data bagi pemerintah dan pemangku kebijakan dalam merancang strategi peningkatan implementasi sertifikasi halal yang lebih efektif, tepat sasaran, dan berorientasi pada kebutuhan serta persepsi konsumen Indonesia. Penelitian ini juga diharapkan memberikan kontribusi teoritis dalam pengembangan model strategi berbasis analisis sentimen dan segmentasi konsumen, serta kontribusi praktis bagi perumusan kebijakan sertifikasi halal di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, Indonesia sebagai negara dengan jumlah penduduk Muslim terbesar di dunia, yaitu sekitar 87,2% dari total populasi, menjadikan jaminan kehalalan produk sebagai kebutuhan mendasar sehingga pemerintah menetapkan kebijakan kewajiban sertifikasi halal melalui Undang-Undang Nomor 33 Tahun 2014 yang mengubah sertifikasi halal dari bersifat sukarela (*voluntary*) menjadi wajib (*mandatory*). Namun, implementasi kebijakan ini masih menghadapi berbagai tantangan yang tercermin dari beragamnya respons dan persepsi konsumen, mulai dari dukungan terhadap jaminan kehalalan hingga kekhawatiran terkait kesiapan pelaku usaha, kompleksitas prosedur, dan potensi biaya. Persepsi tersebut banyak diekspresikan melalui media sosial, khususnya X, yang membentuk sentimen positif, netral, dan negatif terhadap kewajiban sertifikasi halal. Oleh karena itu, diperlukan strategi peningkatan sertifikasi halal yang disusun berdasarkan analisis sentimen dan karakteristik kelompok konsumen agar pemangku kepentingan dapat merumuskan kebijakan yang lebih efektif, tepat sasaran, dan berbasis data dalam meningkatkan keberhasilan implementasi sertifikasi halal di Indonesia. Berikut merupakan rumusan masalah pada penelitian ini:

1. Bagaimana hasil analisis sentimen konsumen Indonesia terhadap implementasi kewajiban sertifikasi halal berdasarkan data X menggunakan metode *text*

mining?

2. Bagaimana hasil pengelompokan (*clustering*) karakteristik konsumen Indonesia berdasarkan kesamaan pola sentimen dan persepsi terhadap kewajiban sertifikasi halal menggunakan metode *K-Means Clustering*?
3. Strategi apa yang dapat diusulkan pada setiap cluster konsumen berdasarkan hasil analisis sentimen dan karakteristik tersebut untuk mendukung peningkatan sertifikasi halal di Indonesia?

1.3 Tujuan Penelitian

Berikut merupakan tujuan penelitian dari penelitian ini:

1. Menganalisis sentimen konsumen Indonesia terhadap implementasi kewajiban sertifikasi halal berdasarkan data X menggunakan metode *text mining*.
2. Mengelompokkan konsumen Indonesia berdasarkan karakteristik dan kesamaan pola sentimen terhadap kewajiban sertifikasi halal menggunakan metode *K-Means clustering*.
3. Menyusun strategi peningkatan sertifikasi halal yang lebih efektif dan tepat sasaran berdasarkan hasil analisis sentimen dan karakteristik tiap *cluster* konsumen.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian yang dilakukan adalah sebagai berikut:

1. Bagi Masyarakat

Peneliti ini memberikan informasi kepada masyarakat mengenai penerapan kebijakan kewajiban sertifikasi halal serta gambaran sikap dan pandangan konsumen terhadap kebijakan tersebut. Hasil penelitian diharapkan dapat meningkatkan pemahaman dan kesadaran Masyarakat dalam memilih produk halal, serta memahami pentingnya sertifikasi halal sebagai bentuk perlindungan konsumen.

2. Bagi Pemerintah

Penelitian ini menggambarkan masukan bagi pemerintah dalam memahami pola persepsi dan karakteristik konsumen Indonesia terhadap implementasi kewajiban sertifikasi halal. Hasil penelitian ini dapat digunakan sebagai landasan dalam penyusunan kebijakan, perencanaan strategi sosialisasi, dan penguatan komunikasi publik yang lebih tepat guna dalam mendukung keberhasilan pelaksanaan program sertifikasi halal.

3. Bagi Peneliti Selanjutnya

Penelitian ini dapat menjadi referensi dan acuan bagi peneliti berikutnya dalam mengembangkan penelitian sejenis, khususnya yang berkaitan dengan analisis sentimen, *text mining*, dan metode *clustering* dalam kajian kebijakan publik dan industri halal, serta sebagai bahan perbandingan untuk penelitian lanjutan dengan objek dan metode yang berbeda.

1.5 Batasan Penelitian

Berikut merupakan batasan masalah dari penelitian ini:

1. Data penelitian diperoleh dari media sosial X yang membahas implementasi kewajiban sertifikasi halal di Indonesia.
2. Data yang dianalisis berupa *tweet* yang mengandung kata kunci terkait sertifikasi halal.
3. Data yang digunakan merupakan *tweet* yang dipublikasikan pada periode 2020–2025.
4. Data yang dianalisis dibatasi pada *tweet* yang menggunakan bahasa Indonesia dan bahasa Inggris.
5. Metode yang digunakan dalam penelitian ini meliputi *text mining*, analisis sentimen, dan *K-Means clustering* untuk mengidentifikasi karakteristik kelompok opini masyarakat terhadap sertifikasi halal.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Literatur

Kajian literatur adalah metode pengumpulan data yang dilakukan melalui proses penelusuran, penelaahan, dan analisis mendalam terhadap berbagai sumber akademik tertulis, seperti buku, jurnal ilmiah, dan laporan penelitian (Ramadhan et al., 2023). Dalam penelitian ini, referensi diperoleh dari sejumlah penelitian terdahulu yang disusun dalam bentuk kajian induktif, yaitu rangkuman hasil-hasil penelitian sebelumnya dengan topik dan metode yang relevan sebagai dasar pengembangan penelitian yang akan dilakukan. Berikut merupakan kajian induktif dengan topik dan metode yang relevan dengan penelitian ini:

Penelitian yang dilakukan oleh (Hajaroh et al., 2024) dengan judul **“Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Produk Makanan dan Minuman di Tokopedia”** bertujuan untuk menganalisis sentimen ulasan konsumen terhadap produk makanan dan minuman pada platform Tokopedia. Penelitian ini menggunakan 4.060 data ulasan yang diperoleh dari dataset terbuka *Kaggle* dan diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral, dengan menggunakan algoritma *Naive Bayes* melalui pendekatan *Knowledge Discovery in Databases* (KDD). Hasil pengujian model menunjukkan tingkat akurasi sebesar 75,90%, presisi sebesar 66,45%, dan *recall* sebesar 97,42%, yang menandakan bahwa model memiliki kemampuan yang cukup baik dalam mengidentifikasi sentimen positif dari ulasan konsumen. Temuan penelitian ini memberikan wawasan penting bagi pelaku bisnis dan penjual di *platform e-commerce* dalam memahami persepsi pelanggan, sehingga dapat digunakan sebagai dasar pengambilan keputusan strategis untuk meningkatkan kualitas produk dan layanan berdasarkan ulasan konsumen.

Selanjutnya penelitian dari (Hasanah & Agustiningsih, 2020) dengan judul **“Analysis of ‘Halal’ Word in Social Media Using Text Mining and Word Networking”** melakukan analisis terhadap percakapan pengguna X menggunakan metode *text mining* dan *word networking*. Hasil analisis terhadap 243 tweet terpilih menunjukkan bahwa kata “halal” memiliki keterkaitan yang kuat dengan sejumlah kata lain, terutama “*food*”, “*restaurant*”, “*certification*”, “*meat*”, dan “*animal*”. Berdasarkan pengukuran *Degree Centrality*, *Betweenness Centrality*, dan *Closeness Centrality*, kata “*food*” menjadi kata

yang paling dominan dengan nilai *degree centrality* sebesar 75, sehingga dapat disimpulkan bahwa persepsi publik terhadap konsep halal di media sosial paling kuat diasosiasikan dengan aspek makanan.

Kemudian penelitian yang dilakukan oleh (H. H. Hidayat et al., 2021) dengan judul **“Pemetaan Kata Kunci dan Polaritas Sentimen Pengguna X Terhadap Kehalalan Produk”** menghasilkan analisis data X menggunakan metode *text mining* dan klasifikasi sentimen dengan *Naïve Bayes Classifier*. Hasil penelitian menunjukkan bahwa kata kunci utama yang sering muncul dalam pembahasan kehalalan produk adalah produk, makanan, minuman, garansi, dan lingkungan. Selain itu, analisis polaritas sentimen menunjukkan bahwa 70% *tweet* bersentimen positif, 16% netral, dan 14% negatif, yang menegaskan bahwa mayoritas pengguna X memberikan respons positif terhadap isu kehalalan produk.

Penelitian yang dilakukan oleh (Kumala & Dana, 2023) dengan judul penelitian **“Penerapan Algoritma K-Means Clustering Berdasarkan Hasil Penjualan Produk Inez Cosmetic di Toko Rosalia”** ini menghasilkan pengujian dari 224 item produk, menghasilkan 3 *cluster* yaitu: 1) *Cluster 0* yang menghasilkan 37 (16,5%) item yang memiliki karakteristik penjualan produk sedang namun dengan nilai transaksi yang tinggi. 2) *Cluster 1* yang menghasilkan 186 (83%) item yang memiliki karakteristik penjualan rendah namun dengan nilai transaksi yang sedang. 3) *Cluster 2* yang menghasilkan 1 (0,5%) item yang memiliki karakteristik penjualan produk tinggi namun dengan nilai transaksi yang rendah. Pengujian menghasilkan hasil analisis untuk membantu pemilik toko dalam menentukan produk *Inez Cosmetic* mana yang laku dan kurang laku di pasaran, karena pentingnya mengatasi kesulitan dalam pemanfaatan data penjualan untuk mencegah terjadinya penumpukan stok produk.

Penelitian yang dilakukan oleh (Maslahah & Destiarni, 2025) dengan judul penelitian **“Klasterisasi UMKM Terhadap Implementasi Sertifikasi Halal di Kabupaten Pamekasan”** ini menghasilkan pengujian dari 50 UMKM, menghasilkan 2 *cluster* yaitu: 1) *Cluster 1* (UMKM Prima) yang menghasilkan 22 (44%) UMKM yang memiliki kapasitas internal tinggi dan implementasi sertifikasi halal yang baik dikelompokkan menjadi 1 *cluster*. 2) *Cluster 2* (UMKM Potensial) yang dihasilkan 28 (56%) UMKM yang memiliki kendala tinggi sehingga membutuhkan pendampingan intensif dikelompokkan ke dalam 1 *cluster group*. Pengujian menghasilkan hasil analisis untuk membantu mengalokasikan sumber daya (seperti dana, pelatihan, atau

pendampingan) secara lebih efisien, karena pentingnya sertifikasi halal untuk meningkatkan kepercayaan konsumen dan daya saing produk.

Selanjutnya penelitian dari (Izza & Akbar, 2023) dengan judul penelitian *“Uncovering Halal Label Food Purchase Intention Behavior in Indonesia: A X Analytics and SEM-Based Investigation”* ini menghasilkan pengujian menggunakan metode campuran, menghasilkan 2 tahap analisis yaitu: 1) Analisis Kualitatif yang mengidentifikasi trending topik dari 2.665 percakapan label halal di X untuk mengungkap persepsi yang mendasarinya. 2) Analisis Kuantitatif yang mengumpulkan data dari 407 konsumen muslim yang menunjukkan bahwa variabel keamanan dan pengetahuan berpengaruh signifikan terhadap niat beli dan pengulangan pembelian. Pengujian menghasilkan hasil analisis untuk membantu memahami faktor pendorong minat beli konsumen, karena pentingnya meningkatnya kesadaran konsumen tentang kehalalan produk, kesehatan, dan nilai yang dirasakan untuk mendorong minat konsumen membeli produk makanan berlabel halal.

Kemudian penelitian yang dilakukan oleh (Nawangsih, 2023) dengan judul penelitian *“Analisa Penjualan Produk Kosmetik Dengan Metode Algoritma K-Means Di Toko Erremy”* ini menghasilkan pengujian dari 67 data produk, menghasilkan 2 *cluster* yaitu: 1) *Cluster 1* yang menghasilkan 54 (81%) produk yang memiliki karakteristik pergerakan stok lambat atau rendah (*Slow Moving Products*) dikelompokkan menjadi 1 *cluster*. 2) *Cluster 2* yang dihasilkan 13 (19%) produk yang memiliki karakteristik pergerakan stok cepat atau tinggi (*Fast Moving Products*) dikelompokkan ke dalam 1 *cluster group*. Pengujian menghasilkan hasil analisis untuk membantu mengetahui minat pembeli terhadap suatu produk agar dapat memprediksi penyediaan sesuai kebutuhan, karena pentingnya proses manajemen persediaan untuk menghindari penumpukan barang yang sama dan mencegah kehabisan stok produk yang laku terjual.

Penelitian yang dilakukan oleh (Puspitasari et al., 2025) dengan judul penelitian *“Clustering of Drug, Cosmetic, Traditional Medicine, and Food Crime Vulnerability in East Java Using the K-Medoids Algorithm”* ini menghasilkan pengujian dari 38 kabupaten/kota, menghasilkan 2 kategori status perkembangan yaitu: 1) Kelompok status semakin aman yang menghasilkan 26 (68%) kabupaten/kota yang memiliki tren penurunan tingkat kerawanan dikelompokkan menjadi 1 kelompok. 2) Kelompok status semakin rawan yang dihasilkan 12 (32%) kabupaten/kota yang memiliki tren peningkatan tingkat kerawanan dikelompokkan ke dalam 1 *cluster group*. Pengujian menghasilkan

hasil analisis untuk membantu pemerintah dalam merumuskan strategi pencegahan yang efektif, karena pentingnya pemberdayaan masyarakat dan intensifikasi komunikasi mengenai produk obat dan makanan yang legal dan aman untuk menekan angka kejahatan.

Penelitian dengan judul **“Produk Roti dalam Pola Konsumsi Pangan dan Keberadaan Label Halal dalam Keputusan Konsumsi Masyarakat (Kasus: Kota Bogor)”** yang ditulis oleh (Rizka et al., 2020) menghasilkan analisis pola konsumsi dan faktor keputusan pembelian menggunakan metode regresi logistik dan deskriptif kualitatif. Hasil penelitian menunjukkan bahwa produk roti merupakan sumber karbohidrat kedua setelah beras dengan tingkat konsumsi yang meningkat seiring tingginya golongan pendapatan, namun ketersediaan produk bersertifikat halal baru mencapai sekitar 11,6% dari total produk. Analisis regresi logistik menghasilkan temuan bahwa variabel halal *awareness*, lama pendidikan, *dummy* asal, dan *dummy* golongan pendapatan bawah berpengaruh signifikan terhadap keputusan konsumsi produk roti berlabel halal.

Selanjutnya penelitian dengan judul **“Hierarchical Clustering Bahan Menu di Kantin Universitas Untuk Menunjang Implementasi Sistem Jaminan Halal”** yang ditulis oleh (Sucipto et al., 2021) menghasilkan pengelompokan bahan menu menjadi dua kategori status kehalalan yang belum bersertifikat, yaitu bahan non-kritis dan bahan kritis. Hasil penelitian menunjukkan bahwa dari total 722 bahan, terdapat 487 bahan (67,45%) yang masuk dalam kategori bahan non-kritis, 26 bahan (3,55%) dalam kategori bahan kritis, dan 209 bahan (28,95%) lainnya telah bersertifikat halal. Implementasi metode *hierarchical clustering* ini dinilai efektif dalam membantu mengurai kompleksitas bahan menu dan menentukan penanganan yang tepat untuk mempercepat persiapan dokumen pengajuan sertifikasi halal.

Kemudian penelitian yang ditulis oleh (Susilawati et al., 2021) dengan judul **“Analisis Persepsi Warganet Mengenai Halal Menggunakan X”** melalui analisis terhadap 15.454 *tweet*, penelitian ini berhasil mengidentifikasi 10 kata kunci utama yang mendominasi diskursus warganet mengenai topik halal. Hasil studi menunjukkan bahwa kata 'Allah' memiliki frekuensi tertinggi dengan 3.032 kemunculan (30%), diikuti oleh kata 'Haram' sebanyak 1.696 (17%), dan kata 'Korea' sejumlah 1.057 (10%), serta ditemukan adanya pola hubungan segitiga yang menautkan kata halal, muslim, dan islam. Proses pengolahan data dilakukan menggunakan metode *Big Data Analysis* dengan

bantuan perangkat lunak R Studio yang mencakup tahapan *crawling*, *tokenizing*, dan *stopword removal* guna memvisualisasikan persepsi masyarakat secara akurat dalam bentuk *Word Cloud*.

Selanjutnya penelitian dengan judul “**Analisis Persepsi Publik di Media Sosial Terhadap Produk Halal di Indonesia Menggunakan *Text Mining***” yang ditulis oleh (Salima & Ma’arif, 2024) menghasilkan pengelompokan persepsi publik menjadi 5 klaster topik utama. Hasil penelitian menunjukkan bahwa *Cluster 1* didominasi oleh topik bahan dan paket halal, *Cluster 2* berfokus pada kualitas dan jasa, *Cluster 3* menyoroti aspek sertifikasi dan harga murah, *Cluster 4* berkaitan dengan inovasi, nutrisi, dan etika, serta *Cluster 5* menekankan pada kepuasan pelanggan dan produk ramah lingkungan. Validasi metode menggunakan algoritma *Latent Dirichlet Allocation* (LDA) dan *VADER Sentiment Analysis* menghasilkan pemetaan kata kunci dominan seperti 'terjangkau', 'kualitas', dan 'inovasi' yang menunjukkan bahwa konsumen tidak hanya peduli pada kepatuhan standar halal tetapi juga nilai ekonomi dan kualitas produk.

Kemudian penelitian dari (Santoti et al., 2024) dengan judul “**Implementasi Term Frequency - Inverse Document Frequency dan Cosine Similarity untuk Analisis Kemiripan Deskripsi Produk Halal**” menghasilkan prediksi klasifikasi deskripsi produk menjadi dua kategori utama, yaitu kategori produk halal dan produk haram. Hasil penelitian menunjukkan bahwa sistem berhasil mengenali produk halal dengan nilai *precision* 96%, *recall* 98%, dan *F1-score* 97%, sedangkan untuk produk haram mencapai *precision* 98%, *recall* 95%, dan *F1-score* 97%. Validasi metode menghasilkan nilai akurasi keseluruhan sebesar 97% yang menunjukkan tingkat kebaikan yang sangat efektif karena mengindikasikan keseimbangan yang baik antara *precision* dan *recall* dalam memprediksi kejelasan deskripsi produk.

Kemudian penelitian dengan judul “***Investigating Netizen Sentiment Toward Halal Certification in Indonesia Using Machine Learning***” yang ditulis oleh (Timur et al., 2023) menghasilkan pengelompokan respons netizen di X menjadi kategori sentimen dan topik diskusi utama menggunakan *Orange Data Mining*. Hasil penelitian menunjukkan bahwa mayoritas percakapan didominasi oleh sentimen netral, dengan emosi kegembiraan (*joy*) mencapai angka 68,22% dari total 1.000 tweet, sedangkan topik utama yang muncul meliputi 'halal', 'sertifikasi', 'makanan', dan 'MUI'. Analisis interaksi menunjukkan bahwa atensi tertinggi (*like* dan *retweet*) tertuju pada pembahasan

kemudahan perizinan dalam UU Cipta Kerja yang mengindikasikan penerimaan publik yang baik terhadap regulasi tersebut.

Selanjutnya penelitian yang ditulis oleh (Yuniar & Heikal, 2024) dengan judul **“Segmentasi sertifikat halal pada wilayah Indonesia menggunakan k-means clustering di badan penyelenggara jaminan produk halal”** menghasilkan pengelompokan wilayah di Indonesia menjadi 5 cluster strategi. Hasil penelitian menunjukkan bahwa *Cluster 1* merupakan kelompok wilayah dengan pertumbuhan sertifikat tertinggi namun kurang dalam kuantitas, *Cluster 2* merupakan wilayah dengan pertumbuhan selalu positif, *Cluster 3* memiliki pertumbuhan terendah dan cenderung menurun, *Cluster 4* didominasi usaha bisnis besar yang perlu dipertahankan, dan *Cluster 5* didominasi bisnis lokal berskala mikro dan kecil. Pengujian data menggunakan 12 variabel tolak ukur dan aplikasi SPSS V.25 ini menghasilkan analisis untuk membantu pemerintah membuat strategi yang cocok dan tepat sasaran bagi masing-masing kelompok wilayah.

Tabel 2. 1 Kajian Literatur

No	Penulis	Obyek (Industri)			Metode		
		Minuman	Makanan	Kosmetik	Text Mining	Cluster	Klasifikasi
1	(Hajaroh al., 2024)	√	√				√
2	(Hasanah & Agustiningsih, 2020)	√			√		
3	(H. H. Hidayat et al., 2021)	√	√		√		√
4	(Kumala & Dana, 2023)			√		√	
5	(Maslahah & Destiarni, 2025)		√			√	

No	Penulis	Obyek (Industri)			Metode		
		Minuman	Makanan	Kosmetik	Text Mining	Cluster	Klasifikasi
6	(Izza & Akbar, 2023)		√				√
7	(Nawangsih, 2023)			√		√	
8	(Puspitasari et al., 2025)		√	√		√	
9	(Rizka et al., 2020)		√				√
10	(Sucipto et al., 2021)		√			√	
11	(Susilawati et al., 2021)	√	√	√	√		
12	(Salima & Ma'arif, 2024)	√	√	√	√		
13	(Santoti, dkk, 2024)		√				√
14	(Timur et al., 2023)		√		√		
15	(Yuniar & Heikal, 2024)	√	√	√		√	
16	Peneliti	√	√	√	√	√	

Tabel tersebut menunjukkan hasil kajian literatur yang diperoleh dari database *Scopus* tahun 2020–2025 dengan kata kunci “sertifikasi halal”, "makanan", “cluster”, dan “text

mining” pada bagian judul. Dari 15 jurnal nasional dan internasional yang ditemukan, dipilih 7 jurnal utama yang paling relevan secara metodologis dan objek penelitian. Hasil analisis menunjukkan bahwa sebagian besar penelitian terdahulu cenderung menggunakan satu metode saja atau hanya mengombinasikan sebagian metode, seperti klasifikasi atau *text mining*, dengan objek industri yang masih bervariasi. Selain itu, belum ditemukan penelitian yang secara spesifik mengintegrasikan metode *clustering* dan *text mining* dengan fokus pada satu objek industri tertentu. Oleh karena itu, penelitian ini berupaya mengisi kesenjangan metodologis dengan mengombinasikan kedua metode tersebut secara terpadu serta memusatkan analisis pada satu objek penelitian, yaitu industri makanan, sehingga analisis yang dihasilkan menjadi lebih fokus, spesifik, dan kontekstual.

2.2 Landasan Teori

2.2.1 Analisis Sentimen

Analisis sentimen merupakan metode untuk mengidentifikasi kecenderungan makna dalam suatu teks dengan mengelompokkannya ke dalam kategori positif, negatif, atau netral. Pendekatan ini termasuk dalam ranah *text mining* dan *Natural Language Processing* (NLP) yang digunakan untuk menganalisis dan mengklasifikasikan opini, pandangan, serta persepsi konsumen terhadap kebijakan, produk, maupun layanan berdasarkan data teks tidak terstruktur yang bersumber dari media sosial (Nur et al., 2024). Dalam penelitian ini, analisis sentimen diterapkan untuk mengklasifikasikan tanggapan konsumen Indonesia terhadap implementasi kewajiban sertifikasi halal di X ke dalam sentimen positif, negatif, dan netral, sehingga hasilnya dapat dimanfaatkan sebagai dasar perumusan strategi peningkatan sertifikasi halal serta mendukung pengambilan keputusan berbasis data.

2.2.2 Media Sosial X

Media sosial merupakan media komunikasi berbasis daring yang memungkinkan individu untuk saling berinteraksi dan berbagi informasi dengan pengguna lain yang berada di lokasi berbeda melalui jaringan internet. Salah satu platform media sosial yang banyak digunakan adalah X, yang termasuk dalam jenis *microblogging* dan memungkinkan penggunanya menyampaikan pesan atau pendapat secara singkat dalam batasan jumlah karakter tertentu pada setiap unggahan. Secara umum, X berperan sebagai layanan komunikasi instan yang memungkinkan pengguna mempublikasikan konten kapan saja

dan di mana pun sehingga dapat diakses oleh publik secara luas. Melalui beragam fitur yang disediakan, X juga dimanfaatkan sebagai media penciptaan dan penyebaran konten publik. Fitur-fitur tersebut meliputi *followers* dan *following*, *direct message*, *X search*, *trending topics*, *latest news*, serta fitur interaksi seperti *thread*, *like*, *reply*, dan *comment* (Zaskya et al., 2020).

Dalam konteks penelitian, X menjadi sumber data yang relevan untuk menganalisis opini dan respons masyarakat, termasuk terkait implementasi kewajiban sertifikasi halal di Indonesia. Data opini pengguna yang tersebar dalam bentuk cuitan dapat diolah menggunakan metode *text mining* untuk mengetahui kecenderungan sentimen, serta metode *K-Means clustering* untuk mengelompokkan karakteristik konsumen. Hasil analisis tersebut dapat dimanfaatkan sebagai dasar dalam merumuskan strategi peningkatan sertifikasi halal yang lebih tepat sasaran sesuai dengan sentimen dan karakteristik konsumen Indonesia.

2.2.3 Text Mining

Text Mining adalah proses pengumpulan, pengolahan, dan pengambilan informasi serta pola tertentu dari data teks yang bersifat tidak terstruktur guna mendukung analisis lebih lanjut. Proses ini didasarkan pada pendekatan interdisipliner yang mengintegrasikan *information retrieval*, *data mining*, *machine learning*, statistika, dan linguistik komputasional guna mengolah, mengelompokkan, serta mengungkap pengetahuan tersembunyi dalam data teks sebagai dasar pendukung pengambilan keputusan (Firdaus et al., 2021).

2.2.4 Preprocessing

Preprocessing merupakan tahap awal dalam pengolahan data yang bertujuan mengubah data mentah dari berbagai sumber menjadi data yang lebih rapi, terstandar, dan siap digunakan untuk analisis lanjutan. Pada tahap ini, struktur teks pada tweet disesuaikan agar data menjadi lebih konsisten dan akurat, serta dilakukan penyesuaian format data (Salam et al., 2023). Data *preprocessing* terdiri atas beberapa tahapan yang dilakukan, yaitu:

1. Cleaning

Proses *cleaning* data dilakukan untuk mengidentifikasi dan memperbaiki atau menghilangkan data yang tidak valid maupun tidak relevan dalam suatu kumpulan data. Setelah tahap data *cleaning* dilaksanakan, data siap digunakan pada tahap *preprocessing* selanjutnya, yaitu proses pengolahan dan pengintegrasian data. Dalam

kondisi tertentu, proses data *cleaning* dapat dilakukan secara berulang sesuai dengan kebutuhan analisis atau pemodelan yang digunakan.

2. *Case Folding*

Dalam *text preprocessing*, proses *case folding* digunakan untuk menstandarkan penulisan huruf dengan mengubah seluruh karakter ke dalam satu bentuk yang sama, baik huruf kecil maupun huruf besar. Proses ini sangat membantu dalam analisis teks karena memudahkan perbandingan antar kata. Selain itu, *case folding* berfungsi mengatasi ketidakteraturan data, seperti perbedaan penggunaan huruf kapital yang tidak sesuai kaidah atau kesalahan penulisan, sehingga data menjadi lebih konsisten dan mudah diproses.

3. *Tokenizing*

Tokenizing merupakan suatu proses yang bertujuan untuk memecah suatu kumpulan teks menjadi sebuah kata. Proses ini memiliki peran penting dalam analisis teks karena membantu dalam memahami makna yang terkandung di dalamnya. Beberapa metode yang umum diterapkan dalam proses *tokenizing* antara lain pemisahan berbasis karakter, kata, dan frasa.

4. *Normalization*

Normalisasi merupakan metode yang digunakan untuk mengubah data ke dalam bentuk yang lebih baku atau seragam. Tahap ini bertujuan untuk memperjelas makna kata, meningkatkan konsistensi dan akurasi analisis, serta memudahkan sistem dalam melakukan klasifikasi dan pengolahan data secara lebih efektif.

5. *Stemming*

Stemming merupakan proses pengubahan kata ke dalam bentuk kata dasar melalui penghapusan imbuhan, sehingga dapat mengurangi variasi bentuk kata yang berbeda tetapi memiliki makna yang sama.

6. *Stopword*

Proses *stop-word removal* bertujuan untuk mengurangi jumlah kata yang memiliki nilai informasi rendah sehingga dapat meningkatkan kecepatan dan akurasi proses klasifikasi teks. Pada tahap ini, kata-kata yang tidak bermakna atau kurang relevan dihilangkan menggunakan daftar *stopwords*, karena dianggap tidak memberikan informasi penting dan hanya memperbesar ukuran data. Meskipun demikian, dalam konteks tertentu *stopwords* tetap dapat dipertahankan apabila memiliki peran penting dalam penyampaian makna.

2.2.5 Sertifikasi Halal

Berdasarkan Undang-Undang Nomor 33 Tahun 2014 tentang Jaminan Produk Halal yang ketentuannya telah disempurnakan melalui Undang-Undang Nomor 11 Tahun 2020 tentang Cipta Kerja, pemerintah berperan aktif dalam mendorong produsen untuk menjamin kehalalan produknya, di mana sertifikasi halal yang semula bersifat sukarela (*voluntary*) kini ditetapkan sebagai kewajiban (*mandatory*) (MKRI., 2024). Ketentuan tersebut diperkuat dengan pemberlakuan Undang-Undang Nomor 7 Tahun 1996 tentang Pangan, Undang-Undang Nomor 8 Tahun 1999 tentang Perlindungan Konsumen, serta Peraturan Pemerintah Nomor 69 Tahun 1999 menegaskan bahwa jaminan kehalalan produk pangan, minuman, obat-obatan, kosmetik, dan produk lainnya tidak hanya berada pada individu atau tokoh agama, melainkan juga menjadi kewenangan dan tanggung jawab pemerintah. Bagi umat Islam, mengonsumsi produk halal merupakan kewajiban agama sekaligus hak warga negara yang dijamin oleh Undang-Undang Dasar 1945 dan Undang-Undang Perlindungan Konsumen (Lutfika, 2024). Oleh karena itu, sebagai negara dengan mayoritas penduduk Muslim, Indonesia memerlukan sistem dan lembaga yang mampu menjamin kepastian status kehalalan produk yang beredar di masyarakat.

Dalam konteks tersebut, Majelis Ulama Indonesia (MUI) memiliki peran penting sebagai lembaga yang menghimpun ulama, zuama, dan cendekiawan Muslim dalam menjaga kepentingan umat Islam. Salah satu bentuk nyata dari peran tersebut adalah pendirian Lembaga Pengkajian Pangan, Obat-obatan, dan Kosmetik (LPPOM MUI) pada 6 Januari 1989 yang bertugas melakukan pemeriksaan dan sertifikasi halal. Keberadaan LPPOM MUI menjadi semakin penting mengingat Indonesia tidak hanya memiliki populasi Muslim terbesar di dunia, tetapi juga merupakan pasar besar bagi produk impor, sehingga jaminan kehalalan produk menjadi kebutuhan yang mendesak baik bagi konsumen maupun pemerintah (Farhan, 2018).

Sertifikasi halal merupakan aspek penting dalam pemasaran produk makanan dan minuman karena berpengaruh terhadap perilaku serta preferensi konsumen. Sertifikasi ini dapat dimanfaatkan sebagai *Unique Selling Point* (USP) yang membedakan produk halal dari produk lain di pasar, sehingga menjadi bagian strategis dalam (Hartini & Malahayatie, 2024). Produk yang telah bersertifikat halal memiliki nilai tambah dan keunggulan kompetitif, yang tercermin melalui pencantuman logo halal pada kemasan sebagai identitas dengan nilai jual serta pembentuk citra positif di mata konsumen.

Selain sebagai pemenuhan standar kehalalan dan wujud etika bisnis produsen dalam memberikan jaminan halal kepada konsumen, sertifikasi halal juga berperan dalam meningkatkan tingkat kepercayaan konsumen, khususnya konsumen Muslim. Dari perspektif ekonomi, label halal memberikan manfaat bagi produsen, antara lain meningkatkan kepercayaan pasar, memperkuat USP produk, memperluas akses ke pasar halal global, meningkatkan daya saing dan pemasaran produk, serta menjadi investasi yang relatif rendah dibandingkan dengan potensi peningkatan pendapatan yang dihasilkan (Faridah, 2019).

2.2.6 Clustering

Clustering merupakan metode pengelompokan data yang dilakukan berdasarkan tingkat kedekatan atau kemiripan antar objek, di mana pembentukan setiap kelompok ditentukan oleh kesamaan karakteristik data dan tidak harus memiliki bentuk yang sama. Tingkat kedekatan tersebut salah satunya dapat diukur menggunakan perhitungan jarak *Euclidean*. Metode ini bertujuan untuk menemukan serta mengelompokkan data yang memiliki karakteristik serupa dan termasuk dalam pendekatan *data mining* tanpa pengawasan (*unsupervised*), karena proses pengelompokan dilakukan tanpa melalui tahap pelatihan, tanpa ada guru (*teacher*), serta tidak memerlukan penentuan target *output* (Wakhidah, 2024).

2.2.7 Algoritma K-Means

Algoritma *K-Means* merupakan metode *clustering* non-hirarki yang bersifat sederhana dan umum diterapkan untuk mengelompokkan data dalam jumlah besar berdasarkan atribut tertentu ke dalam sejumlah kelompok (*cluster*) yang jumlahnya telah ditetapkan sebelumnya, yaitu K (Ramdani, 2022). Algoritma ini melakukan pembagian data dengan cara mengelompokkan objek yang memiliki karakteristik serupa ke dalam *cluster* yang sama, sementara data dengan karakteristik berbeda ditempatkan pada *cluster* yang lain. Melalui mekanisme tersebut, *K-Means* bertujuan untuk menghasilkan variasi yang rendah di dalam satu *cluster* dan perbedaan yang tinggi antar *cluster* (Singh, 2011). Menurut (Wakhidah, 2014), berikut merupakan proses untuk melakukan metode algoritma *K-Means*.

1. Menentukan jumlah *cluster* yang akan dibentuk
2. Menentukan nilai *cluster* (centroid)

Pada tahap awal atau iterasi pertama, penentuan nilai centroid dilakukan secara acak. Selanjutnya, pada iterasi berikutnya, nilai centroid diperbarui dengan menggunakan rumus perhitungan tertentu.

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj}$$

Dimana:

v_{ij} = centroid atau rata-rata cluster ke-i untuk variabel ke-j

N_i = jumlah data yang menjadi anggota cluster ke-i

i, k = indeks cluster

j = koordinat centroid

x_{kj} = nilai data ke-k yang ada didalam *cluster* tersebut untuk variabel ke-j

3. Menghitung jarak antara centroid dan setiap titik data

Perhitungan jarak antara centroid dengan masing-masing objek dapat dilakukan menggunakan rumus *Euclidean distance*.

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Dimana:

D_e = *Euclidean distance*

i = banyaknya objek

x, y = koordinat objek

s, t = koordinat centroid

4. Proses pengelompokan data

Penentuan keanggotaan *cluster* dilakukan dengan membandingkan jarak minimum setiap objek terhadap centroid. Pada matriks jarak, nilai keanggotaan data dinyatakan dalam bentuk biner, yaitu nilai 1 untuk data yang termasuk ke dalam *cluster* dengan karakteristik paling sesuai, dan nilai 0 untuk data yang dialokasikan ke *cluster* lainnya.

5. Mengulangi proses ke dua (langkah penentuan centroid)

Tahapan perhitungan diulang hingga diperoleh nilai centroid yang stabil dan tidak terjadi lagi perpindahan anggota data antar *cluster*.

BAB III

METODE PENELITIAN

3.1 Objek dan Subjek Penelitian

Objek pada penelitian ini yaitu komentar masyarakat Indonesia terkait implementasi kewajiban sertifikasi halal yang disampaikan melalui media sosial X. Sedangkan subjek pada penelitian ini yaitu konsumen Indonesia pengguna X yang menyampaikan opini, tanggapan, dan pengalaman mereka mengenai kebijakan kewajiban sertifikasi halal. X dipilih sebagai objek penelitian karena merupakan *platform* media sosial yang aktif digunakan untuk mengekspresikan pendapat secara terbuka dan real time terhadap isu-isu publik, termasuk sertifikasi halal. Keberagaman komentar yang muncul mencerminkan perbedaan sentimen serta karakteristik konsumen terhadap penerapan kewajiban sertifikasi halal. Oleh karena itu, penelitian ini menggunakan metode *text mining* dan *K-Means clustering* untuk menganalisis sentimen serta mengelompokkan karakteristik konsumen Indonesia sebagai bahan evaluasi dalam mendukung implementasi kebijakan sertifikasi halal yang lebih efektif.

3.2 Jenis Data

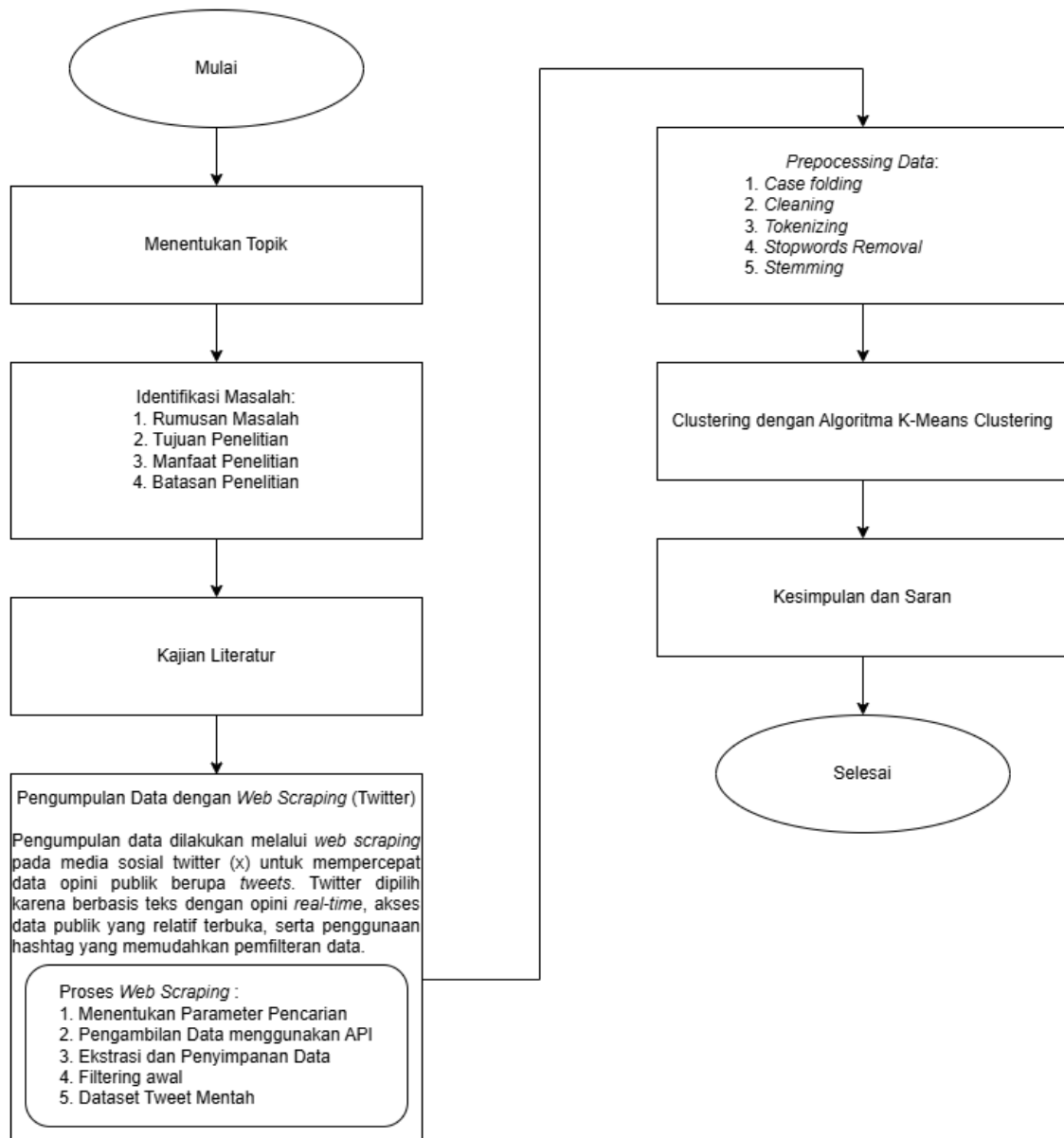
Data yang digunakan dalam penelitian ini merupakan data sekunder, yaitu data yang diperoleh melalui media sosial X. Data tersebut berupa komentar masyarakat Indonesia yang berkaitan dengan kewajiban sertifikasi halal. Proses pengumpulan data dilakukan dengan memanfaatkan X API melalui penerapan teknik *web scraping*, sehingga memungkinkan peneliti untuk mengumpulkan data secara otomatis, sistematis, dan terstruktur sesuai dengan kata kunci serta kriteria yang telah ditentukan guna mendukung proses analisis dalam penelitian ini.

3.3 Metode Pengumpulan Data

Metode pengumpulan data pada penelitian ini dilakukan dengan menggunakan teknik *web scraping*, yang bertujuan untuk memperoleh data berupa komentar masyarakat di media sosial X yang berkaitan dengan kewajiban sertifikasi halal. Proses pengambilan data dilakukan dengan memanfaatkan X API serta perangkat lunak pendukung lainnya yang memungkinkan peneliti mengumpulkan data secara otomatis dan terstruktur sesuai dengan kata kunci yang telah ditentukan.

3.4 Alur Penelitian

Berikut merupakan alur pada penelitian ini:



Gambar 3. 1 Alur Penelitian

Berikut merupakan penjelasan alur pada penelitian ini:

1. Mulai

Peneliti mempersiapkan dan memulai penelitian.

2. Menentukan Topik

Tahap kedua yang dilakukan pada penelitian ini yaitu menentukan topik penelitian. Berdasarkan hasil pengamatan awal, terdapat berbagai komentar masyarakat di media sosial X terkait kewajiban sertifikasi halal yang menunjukkan sentimen

positif, netral, dan negatif. Oleh karena itu, penelitian ini difokuskan pada analisis sentimen dan karakteristik konsumen terhadap pelaksanaan kewajiban sertifikasi halal dengan memanfaatkan data komentar masyarakat di X melalui penerapan metode *text mining*.

3. Identifikasi Masalah dan Rumusan Masalah

Tahap ketiga dalam penelitian ini yaitu melakukan identifikasi dan perumusan masalah. Berdasarkan hasil pengamatan terhadap komentar masyarakat di media sosial X, ditemukan beragam opini yang disampaikan masyarakat terkait kewajiban sertifikasi halal, baik yang bersifat positif, netral, maupun negatif. Beberapa komentar negatif menunjukkan adanya kekhawatiran masyarakat terhadap kesiapan pelaku usaha, kompleksitas prosedur, serta persepsi biaya yang tinggi dalam pelaksanaan kewajiban sertifikasi halal. Berdasarkan permasalahan tersebut, perumusan masalah dalam penelitian ini meliputi bagaimana hasil analisis sentimen masyarakat terhadap kewajiban sertifikasi halal, bagaimana hasil pengelompokan karakteristik konsumen berdasarkan pola sentimen menggunakan metode *K-Means Clustering*, serta strategi apa yang dapat diusulkan untuk mendukung peningkatan dan keberhasilan implementasi kewajiban sertifikasi halal. Pada tahap ini juga ditetapkan tujuan penelitian, batasan penelitian, serta metode yang digunakan dalam penelitian.

4. Kajian Literatur

Tahap keempat dalam penelitian ini adalah melakukan kajian pustaka sebagai upaya untuk memperkuat landasan penelitian, yang dilakukan setelah tahap perumusan masalah. Kajian pustaka ini mencakup kajian induktif berupa perangkuman penelitian-penelitian terdahulu yang memiliki keterkaitan dan relevansi dengan topik penelitian, serta kajian deduktif berupa teori-teori pendukung. Adapun literatur yang dikaji meliputi topik *text mining*, analisis sentimen, *association rules*, serta penggunaan *fishbone diagram* dalam analisis permasalahan.

5. Pengumpulan Data

Tahap ini bertujuan untuk memperoleh data mentah berupa opini masyarakat terkait kewajiban sertifikasi halal dari X dalam bentuk teks (*tweet*) yang merepresentasikan respons publik secara real-time. Proses pengumpulan data diawali dengan penetapan parameter pencarian seperti kata kunci, rentang waktu, bahasa, dan jenis *tweet* agar data yang diperoleh relevan dengan penelitian. Selanjutnya, pengambilan data

dilakukan menggunakan X API untuk mengakses data publik secara terstruktur, meliputi teks *tweet*, tanggal unggahan, serta informasi pendukung lainnya. Data yang terkumpul kemudian diekstraksi dan disimpan dalam format terstruktur, disertai penyaringan awal untuk menghapus duplikasi dan data yang tidak relevan, sehingga menghasilkan dataset *tweet* mentah yang siap diproses pada tahap *preprocessing*. Metode *web scraping* dipilih karena X berbasis teks, bersifat *real-time*, memiliki akses data publik yang relatif terbuka, serta penggunaan hashtag yang memudahkan pemfilteran topik.

6. *Preprocessing Data*

Tahap *preprocessing* data dilakukan untuk membersihkan dan menyiapkan data sebelum dilakukan pengolahan lanjutan, sehingga data tersebut dapat lebih mudah dikenali dan diproses oleh sistem. Adapun tahapan *preprocessing* data yang dilakukan pada penelitian ini adalah sebagai berikut:

a. *Case Folding*

Case folding merupakan tahapan untuk menyeragamkan seluruh huruf, sehingga huruf yang semula kapital menjadi huruf kecil.

b. *Cleaning*

Tahapan *cleaning* data merupakan proses menghapus tanda baca, karakter non-ASCII, dan emoji yang terdapat dalam teks agar data lebih bersih dan siap untuk dianalisis.

c. *Tokenizing*

Tokenizing adalah tahap *preprocessing* data yang bertujuan untuk memecah suatu kalimat menjadi unit-unit kata, sehingga hanya menyisakan kata-kata yang relevan dan dibutuhkan saja.

d. *Stopwords Removal*

Proses *stopwords removal* dilakukan untuk menghapus kata-kata yang tidak memiliki makna penting, seperti *the*, *and*, *but*, *or*, dan kata-kata umum lainnya.

7. *Clustering dengan Algoritma K-Means Clustering*

Data komentar masyarakat yang telah melalui tahap *pre-processing* selanjutnya diproses pada tahap pengelompokan data. Pada proses pengelompokan komentar masyarakat tersebut, metode machine learning yang digunakan adalah algoritma *K-Means Clustering*.

8. Kesimpulan dan Saran

Tahap akhir dalam penelitian ini adalah merangkum seluruh proses dan hasil yang diperoleh, dengan menyajikan kesimpulan yang disusun berdasarkan tujuan dan rumusan masalah penelitian sesuai hasil analisis yang telah dilakukan, serta memberikan saran berupa rekomendasi dan masukan sebagai langkah konkret yang dapat dijadikan perbaikan dan pengembangan pada penelitian selanjutnya.

9. Selesai

BAB IV

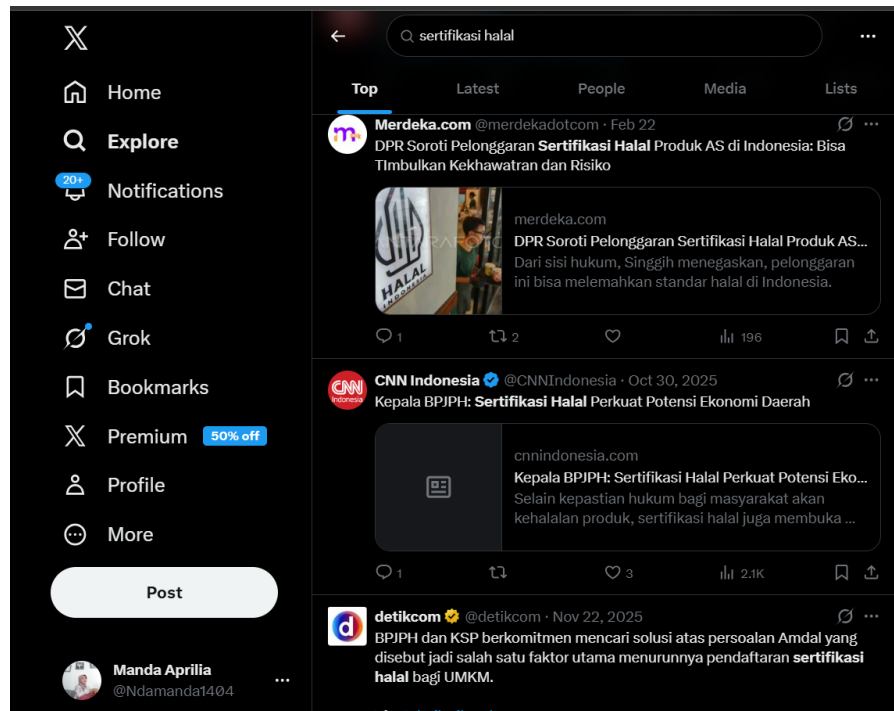
PENGUMPULAN DAN PENGOLAHAN DATA

4.1 Pengumpulan Data

Data penelitian ini berupa komentar pengguna X yang berkaitan dengan kebijakan sertifikasi halal. Pengumpulan data dilakukan menggunakan teknik *web scraping* dengan bantuan aplikasi X untuk mengekstraksi komentar, yang selanjutnya disimpan dalam format *comma separated value* (.csv). Adapun tahapan pengumpulan data komentar X terkait sertifikasi halal dijelaskan sebagai berikut.

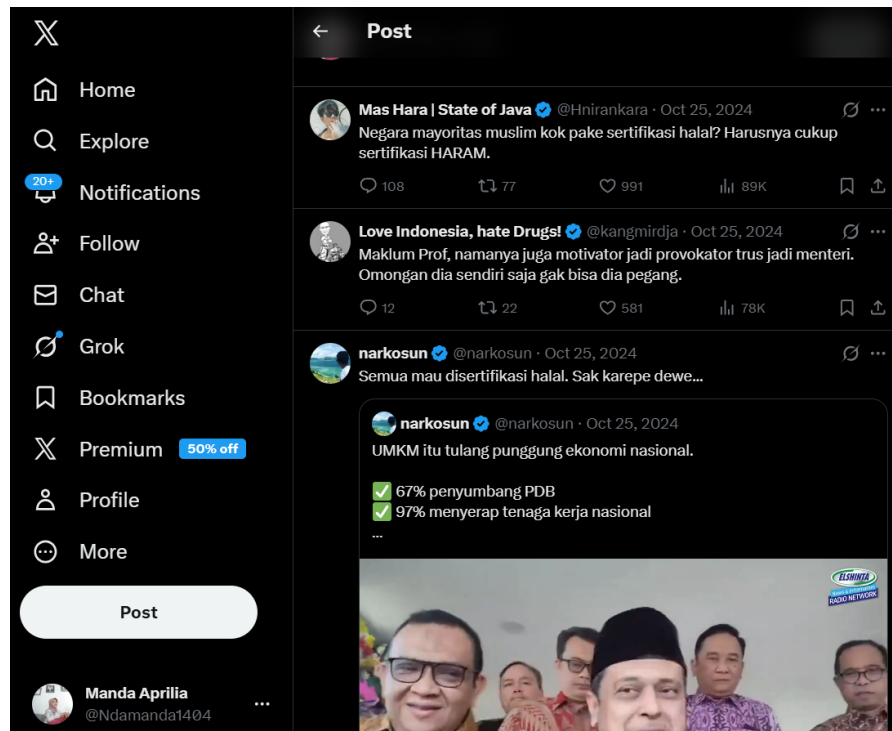
1. Membuka aplikasi X

Tahap pertama dalam melakukan pengambilan data tanggapan masyarakat yaitu membuka aplikasi X melalui *browser* atau aplikasi yang tersedia. Kemudian, masukkan kata kunci yang akan ditelusuri pada kolom pencarian. Pada penelitian ini, kata kunci yang digunakan untuk memperoleh data yaitu “sertifikasi halal” dan “halal” yang berkaitan dengan tanggapan masyarakat terhadap kebijakan sertifikasi halal.



Gambar 4. 1 Aplikasi X

2. Melakukan proses pendataan terhadap tanggapan masyarakat mengenai sertifikasi halal di X



Gambar 4. 2 Tanggapan Masyarakat Mengenai Sertifikasi Halal di X

Pada tahap pengumpulan data, peneliti melakukan proses pendataan terhadap tanggapan masyarakat mengenai sertifikasi halal melalui media sosial X. Platform ini dipilih karena bersifat terbuka, sehingga memungkinkan pengguna menyampaikan opini, komentar, dan respons secara langsung terhadap isu-isu yang sedang berkembang, termasuk kebijakan sertifikasi halal. Proses pengumpulan data dilakukan dengan metode web scraping menggunakan kata kunci tertentu seperti “sertifikasi halal”, “halal”, dan istilah lain yang relevan. Data yang diperoleh berupa unggahan, komentar, jumlah interaksi (*likes*, *repost*, dan *balasan*), serta waktu publikasi. Gambar di atas menunjukkan contoh hasil penelusuran unggahan yang memuat berbagai tanggapan masyarakat, baik yang bersifat mendukung, mengkritik, maupun mempertanyakan kebijakan sertifikasi halal. Melalui tahapan ini, peneliti memperoleh data berupa opini publik yang kemudian diolah dan dianalisis menggunakan metode yang telah ditentukan dalam penelitian. Data yang terkumpul selanjutnya dibersihkan (*data cleaning*) untuk menghapus duplikasi, spam, atau konten yang tidak relevan agar hasil analisis lebih akurat dan representatif terhadap persepsi masyarakat.

3. Melakukan proses *scrapping data*

Tahap akhir dalam pengumpulan data adalah melaksanakan proses *scraping*. Pada tahap ini, peneliti menjalankan proses pengambilan data sesuai dengan pengaturan yang telah ditetapkan agar seluruh data yang relevan dapat terekam secara menyeluruh. Selanjutnya, sistem akan mengekstraksi data unggahan sesuai dengan kata kunci yang digunakan. Setelah proses pengambilan data selesai, data yang diperoleh disimpan dalam format yang dibutuhkan untuk tahap analisis selanjutnya. Dalam penelitian ini, data tanggapan masyarakat terkait sertifikasi halal berhasil dikumpulkan dan disimpan dalam format *csv*. Berikut merupakan hasil dari proses *scraping* data yang telah dilakukan.

conversation_id_str	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name	location	quote_count	reply_count	retweet_count	tweet_url	user_id_str	username
0_200206085482657363	Fri Dec 19 16:57:58 +0000	4	Dirak karna btm lolos BPOM me	200206085482657363		in		0	1	0	https://t.co/145607744886910212	145607744886910212	
1_158860753432091228	Wed Dec 10 09:33:36 +0000	0	CUJ-SSSS... umil kltan yang	158860753432091228		in		0	1	0	https://t.co/126984172544332190	126984172544332190	
2_1583331232091666513	Tue Dec 09 13:56:27 +0000	0	URKMM WAJIB TAHU! Batas Ser	1583331232091666513	https://pbs.twimg.com/media/	in		0	1	0	https://t.co/158861438403171712	158861438403171712	
3_1597570408436914646	Sun Dec 07 03:36:07 +0000	13	minimal kultural audience ind	1597570408436914646		in		0	0	0	https://t.co/155503593230832768	155503593230832768	
4_1596516889042620067	Thu Dec 04 10:05:58 +0000	0	@Sentjoko @CharGPT app Jap	1596516889042620067		Sentjoko	in	0	1	0	https://t.co/1720665183188922368	1720665183188922368	
5_159561608205147330	Tue Dec 02 11:23:30 +0000	0	URUS IZIN HALAL ME INSTAN!	159561608205147330		in		0	0	0	https://t.co/161529552800073672	161529552800073672	
6_1594601310263278051	Sat Nov 29 03:20:15 +0000	0	Sertifikasi halal kelompok da	1594601310263278051	https://pbs.twimg.com/media/	in		0	0	0	https://t.co/165224020405943200	165224020405943200	
7_1594601426437245450	Sat Nov 29 02:56:52 +0000	0	Memiliki sertifikasi halal ant	1594601426437245450	https://pbs.twimg.com/media/	in		0	0	0	https://t.co/17254449750517427488	17254449750517427488	
8_1593882263037131419	Thu Nov 27 13:55:03 +0000	0	@grok @yudhomanesha @Val	1593882263037131419	grok	in		0	1	0	https://t.co/196604462	196604462	
9_1593254343330661425	Tue Nov 25 09:44:04 +0000	0	info yang jual uraha makanan	1593254343330661425		in		0	0	0	https://t.co/1394389675715266306	1394389675715266306	
10_1593205523968618843	Tue Nov 25 09:28:44 +0000	0	@halalindonesia... Dari makan	1593205523968618843		halalindonesia	in	0	0	0	https://t.co/1551081511730188288	1551081511730188288	
11_15932075985171474658	Mon Nov 24 10:46:04 +0000	0	Euy infokan pentul makanan	15932075985171474658		in		4	27	0	https://t.co/1248335962158631041	1248335962158631041	
12_1591524257571046460	Fri Nov 21 14:36:27 +0000	1	@gulfurion @arscompepted	1591524257571046460	gulfurion	in		0	1	0	https://t.co/151027978693047255	151027978693047255	
13_159133503227757615	Fri Nov 21 04:16:04 +0000	0	@repaaidoo @AryaDwipangga	159133503227757615	repaaidoo	in		0	0	0	https://t.co/1156746464	1156746464	
14_1591558781363585257	Fri Nov 21 00:07:49 +0000	0	Adakah sertifikasi untuk mak	1591558781363585257		in		0	0	0	https://t.co/138964660023804544	138964660023804544	
15_1591558781363585257	Fri Nov 21 00:07:49 +0000	0	@MingJi51 Indonesia itu berpe	1591558781363585257	MingJi51	in		0	0	0	https://t.co/2207668320	2207668320	
16_159133503227757615	Thu Nov 20 12:54:50 +0000	0	@GNFI Logistik kebal di may	159133503227757615	GNFI	in		0	0	0	https://t.co/1227589468417573635	1227589468417573635	
17_159133503227757615	Thu Nov 20 12:23:18 +0000	0	@GNFI kalo jual makanan kus	159133503227757615	GNFI	in		0	1	0	https://t.co/1500780546	1500780546	
18_159133503227757615	Thu Nov 20 09:18:31 +0000	0	Ada aja alai bulunya buat mal	159133503227757615	in			0	0	0	https://t.co/1706348328	1706348328	
19_159133503227757615	Thu Nov 20 08:43:19 +0000	55	@GNFI Negara mayoritas musl	159133503227757615	GNFI	in		1	3	6	https://t.co/1647063936367665664	1647063936367665664	
20_159133503227757615	Thu Nov 20 08:28:17 +0000	0	URUS IZIN HALAL CLAHAN ES!	159133503227757615	in			0	0	0	https://t.co/161529552800073672	161529552800073672	
21_159133503227757615	Thu Nov 20 08:28:00 +0000	3	@GNFI Good news dari manan	159133503227757615	GNFI	in		0	0	1	https://t.co/239436314	239436314	
22_159133503227757615	Thu Nov 20 06:07:53 +0000	38	@GNFI Indonesia negara musul	159133503227757615	in			0	3	6	https://t.co/161529552800073672	161529552800073672	
23_159133503227757615	Thu Nov 20 06:04:55 +0000	0	@tdipakaigula @GNFI Btw mal	159133503227757615	tdipakaigula	in		0	0	0	https://t.co/1520364949	1520364949	
24_159034742738827915	Mon Nov 17 09:13:00 +0000	0	Pemerintah memberikan vaku	159034742738827915	in			0	0	0	https://t.co/135735460	135735460	
25_158839572556827756	Fri Nov 14 21:12:17 +0000	1	@kompascom Bagi konoha lebi	158839572556827756	kompascom	in		0	0	0	https://t.co/184328163210682368	184328163210682368	
26_1589190723984609412	Fri Nov 14 04:36:40 +0000	0	5 Bahan Baku Makanan yang	1589190723984609412		in		0	0	0	https://t.co/10082325	10082325	
27_15883094455322333	Thu Nov 13 14:05:16 +0000	0	@halalindonesia... Kolaborasi	15883094455322333	halalindonesia	in		0	0	0	https://t.co/132780765755181916	132780765755181916	
28_158704536724506239	Thu Nov 11 23:20:05 +0000	0	@halalindonesia... Penting ban	158704536724506239	halalindonesia	in		0	0	0	https://t.co/1361385719155333620	1361385719155333620	
29_1586740587870760238	Fri Nov 07 10:22:10 +0000	0	URUS IZIN HALAL COKLAT CUL	1586740587870760238	in			0	0	0	https://t.co/161529552800073672	161529552800073672	
30_1585681388281147467	Thu Nov 06 04:12:43 +0000	0	@Fashionlin @janudin53 @yan	1585681388281147467	Fashionlin	in		0	1	0	https://t.co/1720665183188922368	1720665183188922368	
31_158558183833991024	Wed Nov 05 04:16:01 +0000	1	@Taka_93_77 @ask1825 @E	158558183833991024	Taka_93_77	in		0	0	0	https://t.co/1720665183188922368	1720665183188922368	
32_158558183833991024	Wed Nov 05 03:57:52 +0000	0	@ask1825 @Taka_93_77 @E	158558183833991024	ask1825	in		0	1	0	https://t.co/1720665183188922368	1720665183188922368	
33_1585877734405682846	Wed Nov 05 03:42:50 +0000	2	@hapuulargipol @ajaylalack	1585877734405682846	hapuulargipol	in		0	0	0	https://t.co/1720665183188922368	1720665183188922368	
34_158550428451378619	Tue Nov 04 05:40:20 +0000	2	@kanlelegen @tigon_nao70 @	158550428451378619	kanlelegen	in		0	3	0	https://t.co/243398004	243398004	
35_158555840842937344	Tue Nov 04 04:08:08 +0000	0	@tutdmedia Lho zalahnya dm	158555840842937344	tutdmedia	in		0	0	0	https://t.co/176743174	176743174	
36_158555744285168089	Tue Nov 04 03:58:07 +0000	0	Lho zalahnya dmana? Stop her	158555744285168089	in			0	0	0	https://t.co/176743174	176743174	
37_158443301532031449	Mon Nov 03 02:33:11 +0000	0	@halalindonesia... Mantrap nih	158443301532031449	halalindonesia	in		0	0	0	https://t.co/119572398	119572398	
38_15841440360201801	Mon Nov 03 02:25:14 +0000	0	@halalindonesia... Ikarna deng	15841440360201801	halalindonesia	in		0	0	0	https://t.co/1680245089508945320	1680245089508945320	
39_1584581770141823194	Sat Nov 01 20:39:38 +0000	4	@suelokjro @GoodRecom Syp	1584581770141823194	suelokjro	in		0	3	0	https://t.co/3224001421	3224001421	
40_1583383339777667564	Fri Nov 31 07:20:59 +0000	0	@dhebransen @CalKumHali	1583383339777667564	dhebransen	in		0	0	0	https://t.co/61579502	61579502	

Gambar 4. 3 Hasil Proses Scraping Data

4.2 Pengolahan Data

Data yang diperoleh dari proses scraping X selanjutnya diolah untuk menghasilkan informasi yang bermakna dan relevan dengan tujuan penelitian. Pengolahan diawali dengan analisis deskriptif untuk memberikan gambaran mengenai karakteristik data yang terkumpul. Setelah itu, proses dilanjutkan dengan pendekatan *text mining* yang dimulai dari tahap *preprocessing*, meliputi *case folding*, *tokenizing*, *filtering*, *stemming*, dan *term weighting*. Setelah seluruh tahapan *preprocessing* selesai, data dianalisis menggunakan metode *K-Means Clustering* untuk mengelompokkan tanggapan masyarakat berdasarkan tingkat kemiripan teks. Adapun tahapan pengolahan data dalam penelitian ini dijabarkan sebagai berikut.

4.2.1. Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan gambaran umum mengenai tanggapan masyarakat terhadap isu sertifikasi halal yang dibahas di media sosial X. Media sosial X merupakan salah satu *platform* yang banyak digunakan masyarakat untuk menyampaikan opini, komentar, maupun pandangan terhadap suatu isu yang sedang berkembang secara *real-time* melalui cuitan (*tweet*). Oleh karena itu, data yang berasal dari X dapat dimanfaatkan sebagai sumber informasi untuk mengetahui persepsi dan respons masyarakat terhadap suatu topik tertentu.

Dalam penelitian ini, analisis deskriptif digunakan untuk menggambarkan karakteristik data tanggapan masyarakat yang diperoleh dari hasil *web scraping* pada X dengan kata kunci yang berkaitan dengan sertifikasi halal. Analisis ini menampilkan informasi awal mengenai distribusi data, seperti jumlah data yang diperoleh, frekuensi kemunculan kata, serta kecenderungan tanggapan masyarakat terhadap isu sertifikasi halal. Melalui analisis deskriptif, peneliti dapat mengetahui gambaran umum percakapan masyarakat di X terkait sertifikasi halal sebelum dilakukan analisis lebih lanjut menggunakan metode *text mining* dan *clustering*.

Selain itu, analisis deskriptif juga dapat digunakan untuk melihat pola umum dalam data percakapan, misalnya topik yang paling sering dibicarakan, kata-kata yang dominan muncul dalam *tweet*, serta kecenderungan opini masyarakat terhadap kebijakan atau pelaksanaan sertifikasi halal. Informasi ini menjadi tahap awal yang penting karena dapat memberikan pemahaman awal mengenai bagaimana masyarakat merespons isu sertifikasi halal di media sosial. Dengan demikian, analisis deskriptif membantu peneliti dalam memahami karakteristik data dan menjadi dasar dalam proses analisis lanjutan seperti *pre-processing*, *clustering*, dan visualisasi data.

4.2.2. Processing Data

Data yang digunakan dalam penelitian ini berupa tanggapan masyarakat yang diperoleh melalui proses *scraping* dari media sosial X. Data yang terkumpul masih memiliki berbagai variasi penulisan serta format yang belum terstruktur, sehingga perlu dilakukan tahap *preprocessing* untuk membersihkan dan menyusun data agar lebih sistematis serta siap untuk dianalisis. Proses *preprocessing* dilakukan menggunakan *Jupyter Notebook* dengan bahasa pemrograman *Python*. Beberapa *library* yang dimanfaatkan antara lain Sastrawi, *Natural Language Toolkit* (NLTK), *Regular Expression* (RegEx), *String*, *NumPy*, dan *Pandas*. NLTK, RegEx, dan *String* digunakan pada tahap awal pengolahan

teks, khususnya dalam proses *case folding* dan *tokenizing*, sedangkan *filtering* dan *stemming* dilakukan menggunakan Sastrawi karena lebih sesuai untuk pengolahan teks berbahasa Indonesia. Berikut merupakan tahapan *preprocessing* data dalam penelitian ini.

4.2.2.1. *Case Folding*

Tahap awal dalam proses *preprocessing* data adalah *case folding*. Pada tahap ini dilakukan beberapa proses, yaitu mengubah seluruh huruf menjadi huruf kecil (*lowercase*), kemudian menghapus angka, serta menghilangkan tanda baca yang tidak relevan. Tujuan dari tahap ini adalah untuk menyamakan format teks sehingga lebih mudah diproses pada tahap analisis berikutnya. Proses *case folding* dilakukan dengan bantuan *library Natural Language Toolkit* (NLTK), *String*, dan *Regular Expression* (RegEx). Berikut merupakan contoh penerapan tahap *case folding*.

Tabel 4. 1 Proses *Case Folding*

No	Input	Output
1	Ditarik karna blm lolos BPOM menurut sertifikasi halal dan kesesuaian pangan indo Kenapa dikatakan merusak ginjal? Karna ada ingredients yg over ga sesuai standarisasi makanan indo Jangan apa2 mikirnya pengen persenan doang	ditarik karna blm lolos bpom menurut sertifikasi halal dan kesesuaian pangan indo kenapa dikatakan merusak ginjal karna ada ingredients yg over ga sesuai standarisasi makanan indo jangan apa mikirnya pengen persenan doang
2	GUYSSSS... untuk kalian yang punya usaha makanan/minuman yang mau didaftarkan SERTIFIKASI HALAL boleh bgt pc aku yaaa FREE YAAAA GA DIPUNGUT BIAYAAA!!!!	guyssss untuk kalian yang punya usaha makanan minuman yang mau didaftarkan sertifikasi halal boleh bgt pc aku yaaa free yaaaa ga dipungut biayaaa
3	Tidak Dinas Ketahanan Pangan tidak memiliki wewenang khusus untuk menyelidiki atau menyatakan status halal/non-halal makanan yang dijual ke masyarakat. Sertifikasi halal diatur secara terpusat oleh Badan	tidak dinas ketahanan pangan tidak memiliki wewenang khusus untuk menyelidiki atau menyatakan status halal non-halal makanan yang dijual ke masyarakat sertifikasi halal diatur secara terpusat oleh badan

No	Input	Output
	Penyelenggara Jaminan Produk Halal (BPJPH) dan Majelis Ulama Indonesia	penyelenggara jaminan produk halal (bpjph) dan majelis ulama indonesia
4	Yang ada di Indonesia itu sertifikasi halal bukan informasi haram. Jadi tidak ada kewajiban tempat makan memberitahu makanan tidak halal kalau enggak ditanya. Tidak memberitahu kalau makanan di tempatnya haram ≠ menipu pelanggan.	yang ada di indonesia itu sertifikasi halal bukan informasi haram jadi tidak ada kewajiban tempat makan memberitahu makanan tidak halal kalau enggak ditanya tidak memberitahu kalau makanan di tempatnya haram menipu pelanggan
5	sertifikasi keamanan pangan kan udah ada BPOM? pangan segar minim/belum diolah udah ada OKKP daerah dan pusat??? makanan halal udah ada BPJPH????	sertifikasi keamanan pangan kan udah ada bpom pangan segar minim belum diolah udah ada okkp daerah dan pusat makanan halal udah ada bpjph

4.2.2.2. Tokenizing

Tahap kedua dalam preprocessing data adalah proses tokenizing. Tahap ini merupakan proses memecah teks atau kalimat menjadi bagian-bagian kata secara terpisah. Proses ini bertujuan untuk memudahkan analisis pada setiap kata yang terkandung dalam teks. Proses tokenizing ini dilakukan dengan bantuan library Natural Language Toolkit (NLTK). Berikut merupakan contoh penerapan dari proses tokenizing.

Tabel 4. 2 Proses *Tokenizing*

No	Input	Output
1	ditarik karna blm lolos bpom menurut sertifikasi halal dan kesesuaian pangan indo kenapa dikatakan merusak ginjal karna ada ingredients yg over ga sesuai standarisasi makanan indo jangan apa mikirnya pengen persenan doang	['ditarik', 'karna', 'blm', 'lolos', 'bpom', 'menurut', 'sertifikasi', 'halal', 'dan', 'kesesuaian', 'pangan', 'indo', 'kenapa', 'dikatakan', 'merusak', 'ginjal', 'karna', 'ada', 'ingredients', 'yg', 'over', 'ga', 'sesuai', 'standarisasi', 'makana', 'indo', 'jangan', 'apa', 'mikirnya', 'pengen', 'persenan', 'doang']

No	Input	Output
2	guyssss untuk kalian yang punya usaha makanan minuman yang mau didaftarkan sertifikasi halal boleh bgt pc aku yaaa free yaaaa ga dipungut biayaaa	['guyssss', 'untuk', 'kalian', 'yang', 'punya', 'usaha', 'makanan', 'minuman', 'yang', 'mau', 'didaftarkan', 'sertifikasi', 'halal', 'boleh', 'bgt', 'pc', 'aku', 'yaa', 'free', 'yaaaa', 'ga', 'dipungut', 'biayaaa']
3	sertifikasi keamanan pangan kan udah ada bpom pangan segar minim belum diolah udah ada okkp daerah dan pusat makanan halal udah ada bpjph	['sertifikasi', 'keamanan', 'pangan', 'kan', 'udah', 'ada', 'bpom', 'pangan', 'seger', 'minim', 'belum', 'diolah', 'udah', 'ada', 'okkp', 'daerah', 'dan', 'pusat', 'makanan', 'halal', 'udah', 'ada', 'bpjph']

4.2.2.3. *Filtering*

Tahap ketiga dalam proses *preprocessing* adalah *filtering*. Tahapan ini dilakukan setelah proses *tokenizing*, yang bertujuan untuk menyaring kata-kata hasil pemecahan teks agar hanya kata yang bermakna dan relevan dengan pembahasan sertifikasi halal di media sosial X yang digunakan dalam analisis. Pada tahap ini dilakukan penghapusan kata-kata umum yang tidak memiliki makna khusus, seperti kata hubung atau kata depan, yang dikenal sebagai *stopword removal*. Proses *filtering* dilakukan dengan memanfaatkan *library* Sastrawi karena sesuai untuk pengolahan teks berbahasa Indonesia. Berikut merupakan contoh penerapan proses *filtering*.

Tabel 4. 3 Proses *Filtering*

No	Input	Output
1	['ditarik', 'karna', 'blm', 'lolos', 'bpom', 'menurut', 'sertifikasi', 'halal', 'dan', 'sertifikasi', 'halal', 'kesesuaian', 'kesesuaian', 'pangan', 'indo', 'kenapa', 'pangan', 'indo', 'merusak', 'ginjal', 'dikatakan', 'merusak', 'ginjal', 'karna', 'karna', 'ingredients', 'over', 'sesuai', 'ada', 'ingredients', 'yg', 'over', 'ga', 'standarisasi', 'makana', 'indo', 'sesuai', 'standarisasi', 'makana', 'indo', 'mikirnya', 'pengen', 'persenan', 'jangan', 'apa', 'mikirnya', 'pengen', 'doang']	['ditarik', 'karna', 'blm', 'lolos', 'bpom', 'kesesuaian', 'pangan', 'indo', 'merusak', 'ginjal', 'karna', 'karna', 'ingredients', 'over', 'sesuai', 'ada', 'ingredients', 'yg', 'over', 'ga', 'standarisasi', 'makana', 'indo', 'mikirnya', 'pengen', 'persenan', 'doang']

No	Input	Output
2	['guyssss', 'untuk', 'kalian', 'yang', 'punya', 'usaha', 'makanan', 'minuman', 'minuman', 'didaftarin', 'sertifikasi', 'yang', 'mau', 'didaftarin', 'sertifikasi', 'halal', 'bgt', 'yaa', 'free', 'yaaaa', 'halal', 'boleh', 'bgt', 'pc', 'aku', 'yaa', 'dipungut', 'biayaaa', 'free', 'yaaaa', 'ga', 'dipungut', 'biayaaa']	['guyssss', 'usaha', 'makanan', 'minuman', 'didaftarin', 'sertifikasi', 'halal', 'bgt', 'yaa', 'free', 'yaaaa', 'dipungut', 'biayaaa']
3	['sertifikasi', 'keamanan', 'pangan', 'kan', 'udah', 'ada', 'bpom', 'pangan', 'bpom', 'pangan', 'seger', 'minim', 'seger', 'minim', 'belum', 'diolah', 'udah', 'belum', 'diolah', 'okkp', 'daerah', 'ada', 'okkp', 'daerah', 'dan', 'pusat', 'pusat', 'makanan', 'halal', 'bpjph', 'makanan', 'halal', 'udah', 'ada', 'bpjph']	['sertifikasi', 'keamanan', 'pangan', 'bpom', 'pangan', 'seger', 'minim', 'diolah', 'okkp', 'pusat', 'makanan', 'halal', 'bpjph']

4.2.2.4. Stemming

Tahap keempat dalam proses *preprocessing* adalah *stemming*, yaitu proses mengubah kata-kata hasil filtering ke dalam bentuk dasar atau kata baku agar memiliki keseragaman makna dalam analisis. Proses ini dilakukan dengan menggunakan *library* Sastrawi karena sesuai untuk pengolahan teks berbahasa Indonesia. Dalam penelitian mengenai tanggapan masyarakat terhadap sertifikasi halal di media sosial X, *stemming* diterapkan karena dinilai efektif dan efisien dalam menyederhanakan berbagai variasi kata sehingga menghasilkan data yang lebih relevan dalam waktu yang relatif singkat. Dengan begitu, proses analisis sentimen dapat dilakukan secara lebih optimal. Berikut merupakan contoh penerapan proses *stemming*.

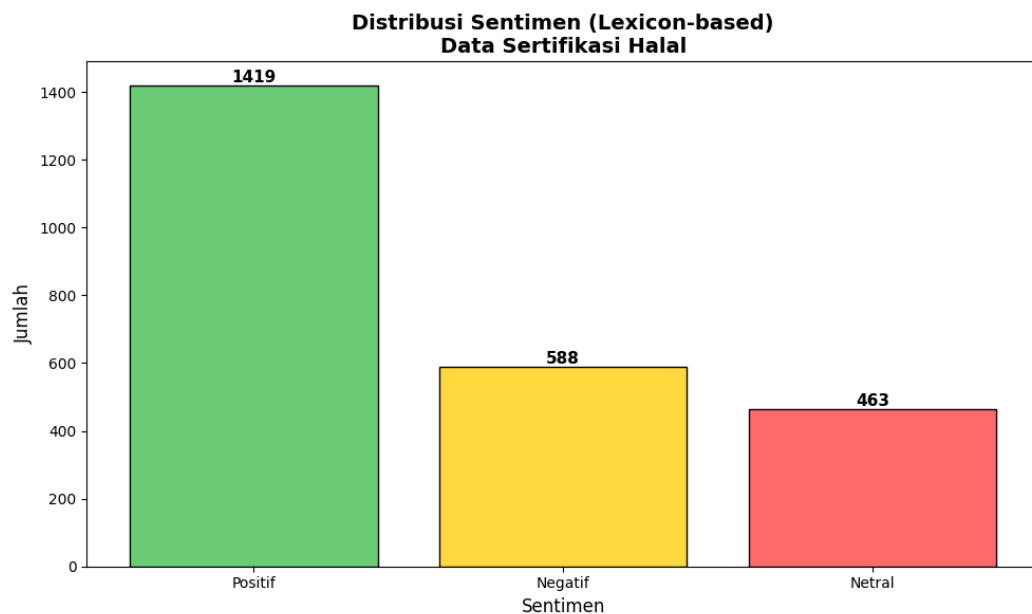
Tabel 4. 4 Proses *Stemming*

No	Input	Output
1	['ditarik', 'karna', 'blm', 'lolos', 'bpom', 'sertifikasi', 'halal', 'kesesuaian', 'pangan', 'pangan', 'indo', 'merusak', 'ginjal', 'indo', 'rusak', 'ginjal', 'karena', 'karna', 'ingredients', 'over', 'sesuai', 'ingredients', 'over', 'sesuai', 'standarisasi', 'makana', 'indo', 'standarisasi', 'makana', 'indo', 'mikirnya', 'pengen', 'persenan', 'doang']	['tarik', 'karena', 'blm', 'lolos', 'bpom', 'sertifikasi', 'halal', 'sesuai', 'pangan', 'pangan', 'indo', 'rusak', 'ginjal', 'karena', 'ingredients', 'over', 'sesuai', 'standarisasi', 'makana', 'indo', 'mikirnya', 'ingin', 'persen', 'doang']
2	['guyssss', 'usaha', 'makanan', 'minuman', 'didaftarin', 'sertifikasi', 'minuman', 'didaftarin', 'sertifikasi', 'minuman', 'didaftarin', 'sertifikasi']	['guyssss', 'usaha', 'makanan', 'minuman', 'didaftarin', 'sertifikasi']

No	Input	Output
	'halal', 'bgt', 'yaa', 'free', 'yaaaa', 'dipungut', 'biayaaa']	'halal', 'bgt', 'yaa', 'free', 'yaaaa', 'dipungut', 'biayaaa']
3	['sertifikasi', 'keamanan', 'pangan', 'bpom', 'pangan', 'seger', 'minim', 'belum', 'diolah', 'okkp', 'daerah', 'pusat', 'makanan', 'halal', 'bpjph']	['sertifikasi', 'keamanan', 'pangan', 'bpom', 'pangan', 'seger', 'minim', 'belum', 'diolah', 'okkp', 'daerah', 'pusat', 'makanan', 'halal', 'bpjph']

4.2.3. Distribusi Sentimen

Tahap selanjutnya adalah melakukan proses *clustering*. Pada tahap ini, data tanggapan masyarakat terhadap sertifikasi halal di media sosial X dikelompokkan menggunakan algoritma *K-Means Clustering* dengan jumlah klaster yang telah ditentukan sebelumnya, yaitu 3 klaster. Proses clustering dilakukan melalui aplikasi *Jupyter Notebook* dengan bahasa pemrograman *Python*. Berikut merupakan grafik hasil pengelompokan data tanggapan masyarakat terkait sertifikasi halal di media sosial X.



Gambar 4. 4 Grafik *Clustering* Data Sertifikasi Halal

Berdasarkan gambar di atas, dapat dilihat bahwa grafik data tanggapan masyarakat terkait sertifikasi halal yang dianalisis menggunakan metode *lexicon-based*, di mana hasil klasifikasi membagi data ke dalam tiga kategori, yaitu positif sebanyak 1.419 data, negatif sebanyak 558 data, dan netral sebanyak 463 data. Hasil tersebut menunjukkan bahwa sentimen positif mendominasi dibandingkan sentimen negatif dan netral,

sehingga setelah proses klasifikasi dilakukan, tahap selanjutnya adalah menganalisis lebih lanjut masing-masing kategori untuk mengidentifikasi pola opini masyarakat serta memahami kecenderungan persepsi publik terhadap kebijakan sertifikasi halal.

4.2.4. Representasi Teks Menggunakan TF-IDF

Metode TF-IDF merupakan metode yang digunakan untuk mengukur tingkat kepentingan suatu kata dalam kumpulan dokumen dengan mempertimbangkan frekuensi kata pada dokumen tertentu maupun pada seluruh koleksi data. Metode ini mampu mengurangi pengaruh kata yang sering muncul dan meningkatkan bobot kata yang jarang digunakan melalui skala logaritmik dalam proses perhitungannya (Afda, 2024). Berikut merupakan representasi teks menggunakan TF-IDF:

Tabel 4. 5 Sampel Distribusi Fitur Sentimen

	skor_pos	skor_neg	skor_sentimen	sentimen_rasio
count	2.470.000	2.470.000	2.470.000	2.470.000
mean	1.185	0.677	0.507	0.398
std	0.857	0.802	1.267	0.219
min	0.000	0.000	-4.000	0.000
25%	1.000	0.000	0.000	0.333
50%	1.000	1.000	1.000	0.500
75%	2.000	1.000	1.000	0.500
max	6.000	5.000	6.000	0.857

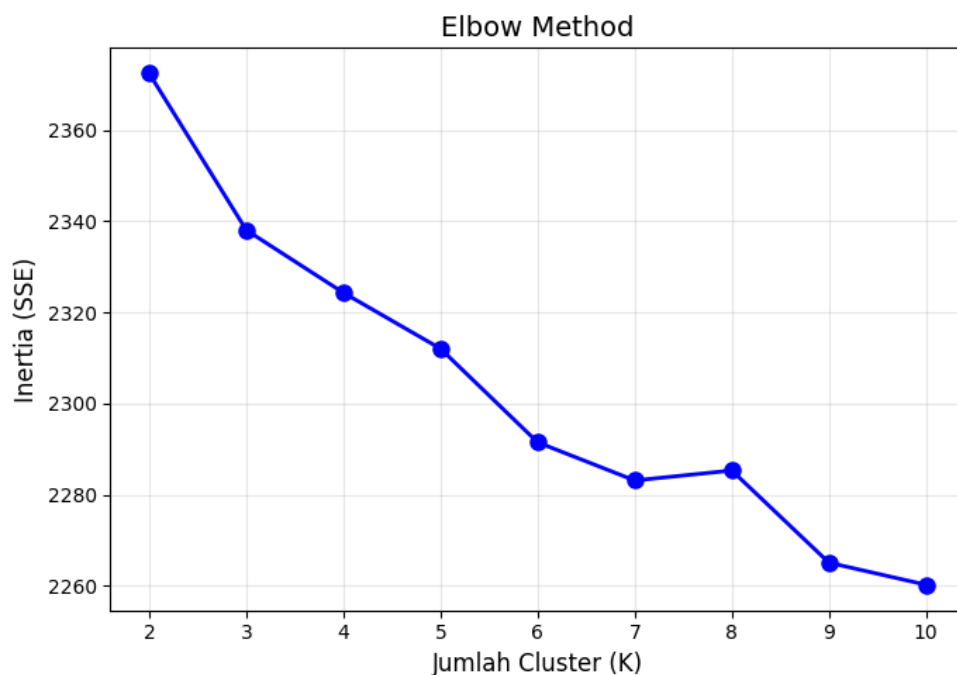
Berdasarkan tabel hasil representasi teks menggunakan TF-IDF dan fitur sentiment, diperoleh statistik deskriptif dari 2.600 data komentar yang dianalisis. Variabel yang digunakan meliputi skor_pos, skor_neg, skor_sentimen, dan sentimen_rasio sebagai fitur tambahan dalam proses clustering. Pada variabel skor_pos diperoleh nilai rata-rata sebesar 1.185 dengan nilai maksimum 6, yang menunjukkan bahwa beberapa komentar memiliki kecenderungan sentimen positif yang cukup tinggi. Sementara itu, skor_neg memiliki rata-rata sebesar 0,677 dengan nilai maksimum 5, yang menunjukkan adanya komentar dengan unsur negatif meskipun intensitasnya relatif rendah. Nilai median kedua variabel sebesar 0 menunjukkan bahwa sebagian besar komentar tidak terlalu didominasi oleh kata positif maupun negatif.

Selanjutnya, variabel skor_sentimen yang merupakan selisih antara skor positif dan negatif memiliki rata-rata sebesar 0,507 sehingga secara umum komentar cenderung

sedikit lebih positif. Nilai minimum sebesar -4 menunjukkan adanya komentar yang sangat negatif, sedangkan nilai maksimum sebesar 6 menunjukkan adanya komentar yang sangat positif. Pada variabel *sentimen_rasio* diperoleh rata-rata sebesar 0,398 dengan nilai maksimum 0,857 yang menunjukkan adanya beberapa komentar dengan dominasi sentimen positif yang cukup tinggi. Secara keseluruhan, hasil ini menunjukkan bahwa data komentar memiliki variasi sentimen yang beragam, mulai dari negatif, netral, hingga positif. Kombinasi TF-IDF dan fitur sentimen membantu proses *clustering* menjadi lebih optimal karena tidak hanya mempertimbangkan kemunculan kata, tetapi juga kecenderungan opini dalam komentar.

4.2.5. Penentuan Jumlah *Cluster* dengan Metode *Elbow*

Tahap awal dalam proses *clustering* tanggapan masyarakat terhadap sertifikasi halal di media sosial X menggunakan algoritma *K-Means Clustering* adalah menentukan jumlah kluster (k) yang akan digunakan. Dalam penelitian ini, penentuan nilai k dilakukan dengan metode *Elbow*. Metode *Elbow* digunakan untuk memperoleh jumlah kuadrat selisih data terhadap nilai prediksi atau pusat kluster untuk menentukan jumlah kluster yang optimal. Metrik yang digunakan dalam perhitungan metode ini adalah *Sum of Squared Errors* (SSE). Proses perhitungan metode *Elbow* dilakukan menggunakan aplikasi *Jupyter Notebook* dengan bahasa pemrograman *Python*. Berikut disajikan grafik hasil perhitungan metode *Elbow*.



Gambar 4. 5 Metode *Elbow*

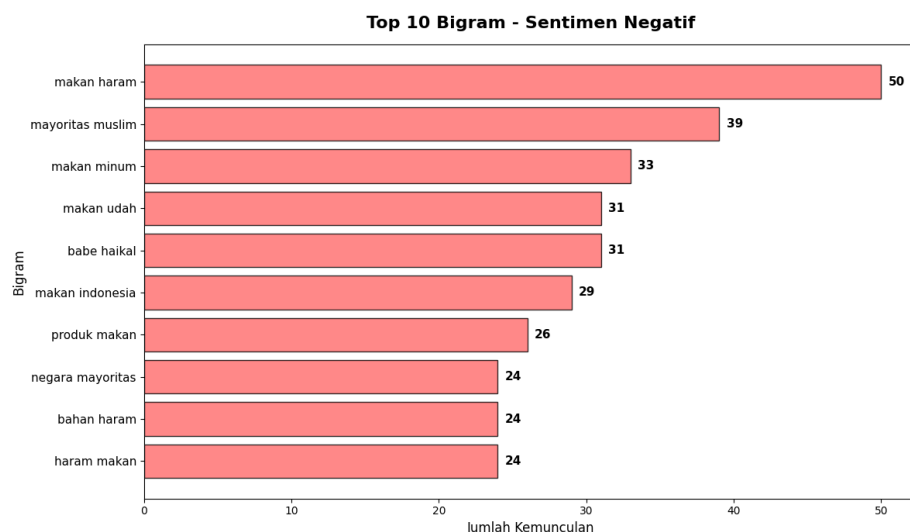
BAB V

PEMBAHASAN

5.1 Analisis Hasil Sentimen

Analisis sentimen digunakan untuk mengekstraksi dan mengidentifikasi kecenderungan opini masyarakat yang tercermin dalam data teks. Metode ini memungkinkan peneliti memperoleh gambaran mengenai persepsi konsumen terhadap suatu topik tertentu (Rivaro et al., 2026). Pada penelitian ini, analisis sentimen diterapkan untuk mengkaji tanggapan masyarakat Indonesia terhadap sertifikasi halal melalui data media sosial X, yang selanjutnya digunakan sebagai dasar dalam pengelompokan konsumen dan perumusan strategi peningkatan sertifikasi halal. Berikut merupakan hasil sentimen dari bigram pada tiap *cluster*:

a. Cluster 0

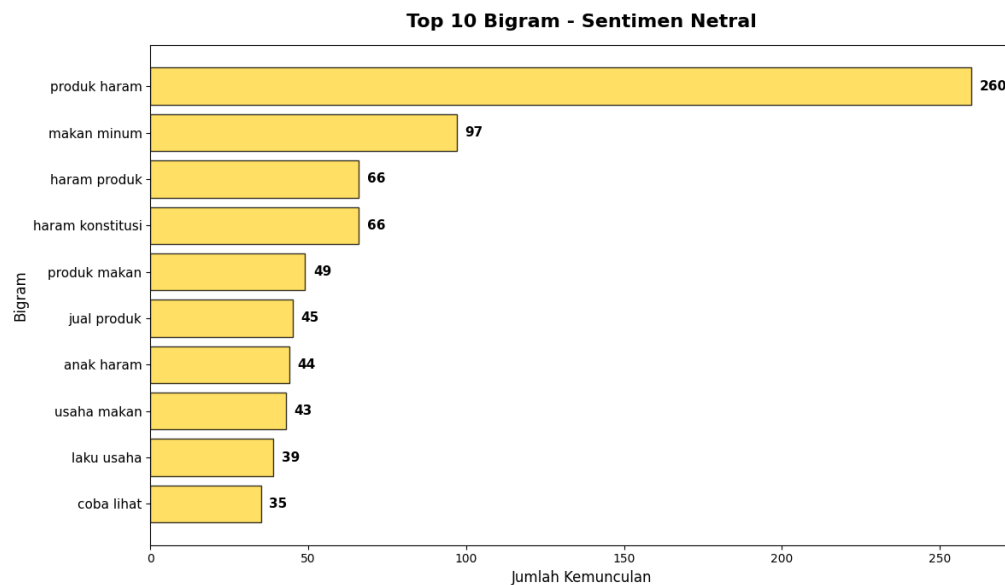


Gambar 5. 1 Bigram Sentimen Negatif

Gambar diatas menunjukkan 10 bigram (kombinasi dua kata) yang paling sering muncul pada sentimen negatif terkait sertifikasi halal di media sosial X. Bigram yang paling dominan adalah “makan haram” dengan frekuensi 50 kali, diikuti oleh “mayoritas muslim” sebanyak 39 kali, dan “makan minum” sebanyak 33 kali. Selain itu, muncul kata bigram seperti “makan udah” dan “babe haikal” yang masing-masing memiliki frekuensi 31 kali, serta “makan indonesia” sebanyak 29 kali. Bigram lainnya yang juga sering muncul adalah “produk makan” dengan frekuensi 26 kali, serta “negara mayoritas”, “bahan haram”, dan “haram makan” yang masing-masing muncul sebanyak 24 kali. Secara umum, hasil ini

menunjukkan bahwa sentimen negatif masyarakat banyak berfokus pada isu kehalalan makanan dan minuman, terutama terkait penggunaan istilah haram dalam konteks konsumsi produk. Selain itu, kemunculan bigram seperti “mayoritas muslim” dan “negara mayoritas” mengindikasikan adanya diskusi mengenai penerapan sertifikasi halal dalam konteks masyarakat Indonesia yang mayoritas beragama Islam. Hal ini menunjukkan bahwa sentimen negatif lebih banyak dipengaruhi oleh perdebatan mengenai status kehalalan produk, persepsi masyarakat terhadap aturan halal, serta berbagai opini dan kritik yang berkembang di ruang publik terkait implementasi sertifikasi halal.

b. Cluster 1

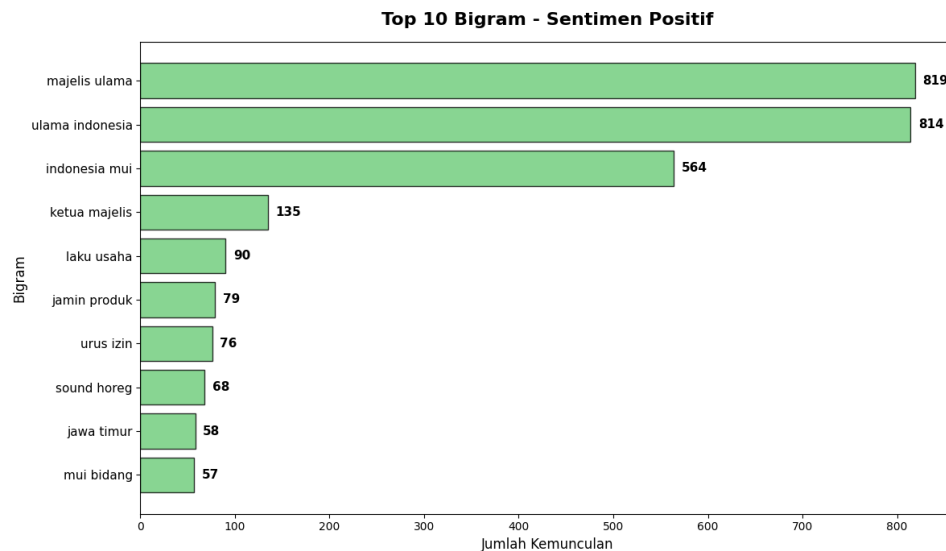


Gambar 5. 2 Bigram Sentimen Netral

Gambar di atas menunjukkan bahwa terdapat 10 bigram (kombinasi dua kata) yang paling sering muncul pada sentimen netral terkait sertifikasi halal di media sosial X. Bigram yang paling dominan adalah “produk haram” dengan frekuensi 260 kali, diikuti oleh “makan minum” sebanyak 97 kali, untuk “haram produk” dan “haram konstitusi” yang masing-masing muncul sebanyak 66 kali. Selain itu, terdapat bigram lain seperti “produk makan” (49 kali), “jual produk” (45 kali), “anak haram” (44 kali), “usaha makan” (43 kali), “laku usaha” (39 kali), dan “coba lihat” (35 kali). Kemunculan bigram-bigram tersebut menunjukkan bahwa *tweet* netral banyak berfokus pada produk, makanan dan minuman, pelaku usaha, serta pembahasan mengenai status kehalalan produk. Kata-kata yang muncul cenderung bersifat informatif dan deskriptif, tanpa menunjukkan ekspresi dukungan maupun

penolakan yang kuat terhadap sertifikasi halal. Secara umum, sentimen netral didominasi oleh penyampaian informasi, diskusi, dan penyebutan istilah-istilah yang berkaitan dengan produk halal, aktivitas usaha, serta isu-isu yang sedang dibahas di ruang publik terkait sertifikasi halal.

c. Cluster 2

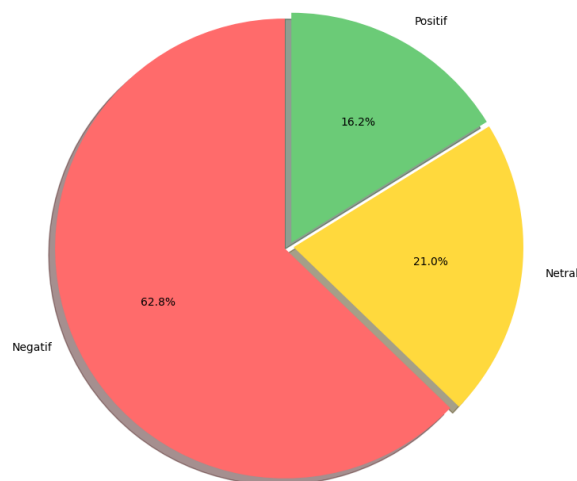


Gambar 5. 3 Bigram Sentimen Positif

Gambar di atas menunjukkan bahwa terdapat 10 bigram (kombinasi dua kata) yang paling sering muncul pada sentimen positif terkait sertifikasi halal di media sosial X. Bigram yang paling dominan adalah “majelis ulama” dengan frekuensi 819 kali, diikuti oleh “ulama indonesia” sebanyak 814 kali, dan “indonesia mui” sebanyak 564 kali. Selain itu, terdapat bigram lain seperti “ketua majelis” (135 kali), “laku usaha” (90 kali), “jamin produk” (79 kali), dan “urus izin” (76 kali) yang berkaitan dengan proses penyelenggaraan dan implementasi sertifikasi halal. Bigram lainnya yang juga sering muncul adalah “sound horeg” (68 kali), “jawa timur” (58 kali), dan “mui bidang” (57 kali). Kemunculan bigram-bigram tersebut menunjukkan bahwa percakapan positif banyak berfokus pada peran Majelis Ulama Indonesia (MUI), jaminan produk halal, dan pelaku usaha yang mengurus perizinan dan sertifikasi halal. Secara umum, sentimen positif mencerminkan adanya kepercayaan, dukungan, dan apresiasi masyarakat terhadap lembaga serta proses sertifikasi halal, yang dianggap mampu memberikan kepastian dan jaminan kehalalan produk bagi konsumen.

5.2 Analisis Hasil Cluster

Clustering merupakan proses pengelompokan data ke dalam beberapa *cluster* berdasarkan karakteristik, sehingga data dalam satu kelompok memiliki tingkat kesamaan yang lebih tinggi dibandingkan dengan data pada kelompok lain (Awaludin Abid, 2023). Proses ini bertujuan membagi data menjadi kelompok-kelompok yang homogen, di mana setiap *cluster* berisi data dengan karakteristik yang paling mirip serta memiliki tingkat perbedaan yang lebih besar terhadap *cluster* lainnya. Dalam penelitian ini, proses *clustering* dilakukan terhadap tanggapan masyarakat mengenai sertifikasi halal di media sosial X berdasarkan pola kata, persepsi, dan kecenderungan sentimen yang muncul pada data *tweet*.



Gambar 5. 4 Hasil *Clustering* dan Visualisai Diagram Lingkaran

Berdasarkan hasil *clustering* dan visualisasi diagram lingkaran di atas, diketahui bahwa sentimen negatif mendominasi percakapan masyarakat dengan persentase sebesar 62,8%, diikuti sentimen netral sebesar 21,0%, dan sentimen positif sebesar 16,2%. Hasil ini menunjukkan bahwa sebagian besar masyarakat masih memberikan tanggapan negatif terkait sertifikasi halal, seperti kendala biaya, prosedur yang rumit, serta kesiapan pelaku usaha dalam memenuhi sertifikasi halal. Sementara itu, *cluster* netral lebih banyak berisi informasi dan diskusi umum terkait sertifikasi halal, sedangkan *cluster* positif menunjukkan adanya dukungan dan kepercayaan masyarakat terhadap pentingnya sertifikasi halal. Dengan adanya pengelompokan ini, karakteristik tiap cluster dapat

dipahami secara lebih rinci dan dapat digunakan sebagai dasar dalam penyusunan strategi peningkatan implementasi sertifikasi halal yang lebih tepat sasaran.

5.3 Strategi Berdasarkan *Cluster*

Berdasarkan hasil analisis pada tahap yang telah diuraikan pada subbab sebelumnya menunjukkan bahwa percakapan masyarakat di media sosial X terkait sertifikasi halal dipengaruhi oleh berbagai faktor yang memengaruhi implementasi serta peningkatan sertifikasi halal di Indonesia. Analisis dilakukan melalui tahapan *text mining* dan metode *K-Means Clustering* yang menghasilkan beberapa kelompok percakapan (*cluster*) berdasarkan kesamaan kata kunci dominan. Hasil pengelompokan tersebut mengungkap berbagai topik yang mencerminkan persepsi masyarakat, seperti aspek biaya, proses pengurusan, kualitas produk, dan tingkat pemahaman pelaku usaha terhadap sertifikasi halal. Untuk mengidentifikasi akar permasalahan yang muncul, dilakukan analisis terhadap faktor penyebab dan faktor pendorong pada masing-masing *cluster*. Pendekatan ini bertujuan mengelompokkan faktor-faktor yang berpengaruh secara sistematis sehingga hubungan antara permasalahan utama dan penyebabnya dapat dipahami secara lebih komprehensif, sekaligus menjadi dasar dalam perumusan strategi peningkatan sertifikasi halal di Indonesia.

Hasil *clustering* menunjukkan bahwa percakapan masyarakat mengenai sertifikasi halal terbagi menjadi tiga *cluster* utama, yaitu cluster negatif, netral, dan positif. *Cluster* 0 termasuk ke dalam *cluster* negatif yang menunjukkan adanya kritik, kekhawatiran, dan ketidakpuasan masyarakat terhadap implementasi sertifikasi halal. Hal ini terlihat dari dominasi kata-kata seperti “sertifikasi halal”, “halal”, “haram”, “produk”, dan “makan”, yang mencerminkan pembahasan mengenai kehalalan produk serta kewajiban sertifikasi halal. Selain itu, kemunculan kata-kata seperti “biaya”, “proses”, “wajib”, dan “pelaku usaha” menunjukkan adanya keluhan terkait prosedur dan beban yang harus ditanggung dalam proses sertifikasi. Kondisi ini mengindikasikan bahwa sebagian masyarakat masih memandang sertifikasi halal sebagai kebijakan yang memerlukan biaya dan proses yang cukup kompleks. Oleh karena itu, strategi yang diperlukan adalah penyederhanaan prosedur sertifikasi, peningkatan transparansi layanan, pemberian pendampingan bagi pelaku usaha, serta sosialisasi yang lebih intensif mengenai manfaat sertifikasi halal.

Selanjutnya, *cluster* 1 dikategorikan sebagai *cluster* netral karena percakapan masyarakat lebih bersifat informatif dan deskriptif tanpa menunjukkan dukungan maupun

penolakan secara jelas terhadap sertifikasi halal. Kata-kata dominan seperti “majelis ulama”, “ulama”, “lembaga”, “MUI”, “bidang”, “dewan pimpinan”, dan “jamin produk” menunjukkan bahwa masyarakat lebih banyak membahas organisasi, kepengurusan, kegiatan, dan program yang berkaitan dengan sertifikasi halal. Fokus pembahasan pada aspek kelembagaan menunjukkan bahwa opini masyarakat dalam *cluster* ini belum mengarah pada penilaian positif maupun negatif terhadap kebijakan sertifikasi halal. Oleh karena itu, strategi yang dapat dilakukan adalah memperkuat komunikasi publik, meningkatkan penyebaran informasi yang mudah dipahami masyarakat, serta memanfaatkan media digital untuk meningkatkan pemahaman mengenai peran lembaga dan manfaat sertifikasi halal.

Sementara itu, *cluster* 2 merupakan *cluster* positif yang menunjukkan adanya dukungan dan kepercayaan masyarakat terhadap sertifikasi halal. Kata-kata seperti “barokah”, “kompetitif”, “amanah”, “keren”, “solusi”, “mantap”, “selamat”, “cepat”, dan “mutu” menggambarkan respons yang bersifat apresiatif terhadap produk maupun kebijakan yang telah memenuhi standar halal. *Cluster* ini menunjukkan bahwa sertifikasi halal dipandang mampu meningkatkan kualitas produk, menjamin keamanan konsumsi, serta memperkuat daya saing produk di pasar. Oleh karena itu, strategi yang perlu dilakukan adalah menjaga kualitas layanan sertifikasi halal, memperkuat promosi dan branding produk halal, meningkatkan kepercayaan konsumen melalui transparansi informasi, serta mendorong kolaborasi antara pemerintah, pelaku usaha, dan lembaga terkait untuk memperluas implementasi sertifikasi halal secara berkelanjutan.

Tabel 5. 1 Strategi Peningkatan Sertifikasi Halal

No	Cluster	Label Cluster	Karakteristik Cluster	Strategi	Referensi
1	Cluster 0	Cluster Negatif	Kritik terhadap biaya, prosedur, kewajiban sertifikasi, dan dampaknya bagi pelaku usaha.	Penyederhanaan proses sertifikasi, peningkatan transparansi, bantuan biaya bagi UMKM, sosialisasi dan literasi halal yang berkelanjutan.	(Mustajab, 2025) dan (Wulan Ariby, Naila Deswita, 2025)

No	Cluster	Label Cluster	Karakteristik Cluster	Strategi	Referensi
2	Cluster 1	Cluster Netral	Fokus pada informasi kelembagaan dan program sertifikasi halal, dengan opini masyarakat yang cenderung netral.	Edukasi publik, penyebaran informasi yang jelas, dan penguatan komunikasi melalui media digital.	(Nasya Tanasya, 2022), (Mahardi & Dadang Karya Bakti, 2025), dan (Prastiwi, 2025).
3	Cluster 2	Cluster Positif	Dukungan terhadap produk halal dan kepercayaan pada kualitas produk.	Penguatan branding halal, peningkatan kualitas layanan, berkolaborasi dengan industri dan pemerintah.	(Mardianto et al., 2025) dan (Zuraidah et al., 2023)

5.4 Limitasi Penelitian

Penelitian ini masih memiliki beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil analisis mengenai strategi peningkatan sertifikasi halal berdasarkan analisis sentimen dan *K-Means Clustering* pada media sosial X. Adapun keterbatasan penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data yang berasal dari media sosial X, sehingga hasil analisis belum dapat merepresentasikan keseluruhan persepsi masyarakat Indonesia terhadap implementasi sertifikasi halal. Hal ini disebabkan karena pengguna X memiliki karakteristik pengguna tertentu dan tidak seluruh masyarakat aktif menyampaikan pendapat melalui *platform* tersebut.
2. Data yang digunakan dalam penelitian berupa teks *tweet* memiliki beragam bentuk bahasa, seperti singkatan, bahasa tidak baku, bahasa gaul, dan penggunaan sarkasme yang sulit dipahami secara otomatis oleh sistem. Kondisi tersebut dapat memengaruhi hasil *preprocessing*, analisis sentimen, maupun proses pengelompokan *cluster* menggunakan metode *K-Means Clustering*.

3. Penelitian ini menggunakan algoritma *K-Means Clustering* yang proses pengelompokannya dipengaruhi oleh penentuan jumlah *cluster* dan pemilihan *centroid* awal. Oleh karena itu, hasil *cluster* yang terbentuk memungkinkan mengalami perubahan apabila menggunakan parameter atau metode *clustering* yang berbeda.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil penelitian mengenai analisis sentimen konsumen Indonesia terhadap implementasi kewajiban sertifikasi halal dengan memanfaatkan data dari media sosial X serta menggunakan pendekatan *text mining*, *K-Means Clustering*, dan analisis *Fishbone*, maka dapat diperoleh beberapa kesimpulan sebagai berikut:

1. Hasil analisis sentimen konsumen Indonesia terhadap implementasi kewajiban sertifikasi halal berdasarkan data X melalui metode *text mining*, diperoleh tiga kelompok sentimen, yaitu negatif, netral, dan positif. Sentimen negatif didominasi oleh pembahasan mengenai status kehalalan produk, aturan halal, serta kritik terhadap implementasi sertifikasi halal. Sentimen netral lebih banyak berisi informasi dan diskusi terkait produk halal, makanan dan minuman, serta aktivitas pelaku usaha tanpa menunjukkan kecenderungan opini yang kuat. Sementara itu, sentimen positif mencerminkan dukungan dan kepercayaan masyarakat terhadap sertifikasi halal, khususnya terhadap peran MUI, jaminan produk halal, dan proses sertifikasi. Secara keseluruhan, hasil penelitian menunjukkan bahwa persepsi masyarakat terhadap sertifikasi halal masih beragam, namun sentimen negatif lebih dominan dibandingkan sentimen netral dan positif. Hal ini mengindikasikan bahwa masih diperlukan peningkatan literasi halal, transparansi informasi, dan sosialisasi yang lebih efektif untuk meningkatkan kepercayaan masyarakat terhadap implementasi kewajiban sertifikasi halal.
2. Hasil pengelompokan karakteristik konsumen Indonesia menggunakan metode *K-Means Clustering* menunjukkan bahwa data percakapan masyarakat di X dapat diklasifikasikan ke dalam beberapa *cluster* yang memiliki karakteristik sentimen dan persepsi yang berbeda terhadap kebijakan sertifikasi halal. Proses pengelompokan diawali dengan penentuan jumlah *cluster* optimal menggunakan metode *elbow*, dan diperoleh nilai $k = 3$ dengan nilai metrik *Sum of Squared Error* (SSE) sekitar 2.600 data *tweet*. *Cluster* 0 didominasi kata seperti “sertifikasi halal”, “halal”, “haram”,

“produk”, dan “makan”, yang mencerminkan kritik dan kekhawatiran terhadap biaya serta proses sertifikasi halal. *Cluster* 1 didominasi kata “majelis ulama”, “MUI”, “lembaga”, dan “jamin produk”, yang menunjukkan pembahasan bersifat informatif terkait organisasi dan program sertifikasi halal. Sementara itu, *cluster* 2 didominasi kata “barokah”, “amanah”, “kompetitif”, dan “mantap”, yang menunjukkan dukungan dan kepercayaan masyarakat terhadap sertifikasi halal. Secara keseluruhan, sentimen negatif mendominasi percakapan sebesar 62,8%, diikuti sentimen netral 21,0%, dan sentimen positif 16,2%, sehingga menunjukkan bahwa persepsi masyarakat terhadap sertifikasi halal masih beragam.

3. Strategi yang dapat diusulkan berdasarkan hasil analisis sentimen dan karakteristik *cluster* menunjukkan bahwa upaya peningkatan sertifikasi halal di Indonesia perlu dilakukan dengan mempertimbangkan berbagai faktor yang memengaruhi persepsi masyarakat. Berdasarkan hasil analisis sentimen, percakapan masyarakat di X terkait sertifikasi halal terbagi ke dalam tiga *cluster*, yaitu negatif, netral, dan positif, yang masing-masing mencerminkan perbedaan persepsi serta faktor penyebab dan pendorong. *Cluster* negatif menunjukkan adanya kritik terhadap biaya, prosedur, dan kewajiban sertifikasi, sehingga diperlukan penyederhanaan proses, peningkatan transparansi, pendampingan, dan literasi halal. *Cluster* netral didominasi oleh informasi terkait lembaga dan program sertifikasi halal, sehingga perlu dilakukan edukasi publik, penguatan komunikasi, dan optimalisasi media digital. Sementara itu, *cluster* positif menunjukkan dukungan dan kepercayaan masyarakat terhadap sertifikasi halal, sehingga strategi yang diperlukan adalah memperkuat branding produk halal, menjaga kualitas layanan, serta meningkatkan kolaborasi antara pemerintah, pelaku usaha, dan lembaga terkait. Strategi tersebut diharapkan dapat meningkatkan jumlah produk bersertifikasi halal dan memperkuat kepercayaan masyarakat terhadap sistem jaminan produk halal di Indonesia.

6.2 Saran

Berdasarkan hasil penelitian mengenai analisis sentimen dan pengelompokan karakteristik konsumen terhadap implementasi sertifikasi halal di Indonesia,

terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian dan implementasi kebijakan ke depan, yaitu sebagai berikut:

1. Penelitian selanjutnya disarankan untuk menggunakan data dari berbagai *platform* media sosial lain, seperti Instagram, TikTok, Facebook, atau YouTube agar hasil analisis persepsi masyarakat terhadap sertifikasi halal menjadi lebih luas, beragam, dan representatif.
2. Penelitian selanjutnya dapat mengembangkan metode analisis dengan menerapkan algoritma *machine learning* atau *deep learning* yang lebih kompleks agar proses klasifikasi sentimen mampu memahami konteks bahasa, ironi, dan penggunaan bahasa informal secara lebih akurat.
3. Pemerintah dan lembaga terkait diharapkan dapat memanfaatkan hasil segmentasi *cluster* konsumen dalam penelitian ini sebagai dasar penyusunan strategi sosialisasi, edukasi, dan pendampingan sertifikasi halal yang lebih efektif dan tepat sasaran, khususnya bagi kelompok masyarakat dan pelaku UMKM yang masih memiliki persepsi negatif terhadap implementasi sertifikasi halal.

DAFTAR PUSTAKA

- Afda, A. (2024). *Analisa Dokumen Menggunakan Metode TF-IDF*. 5(5), 466–470.
- Awaludin Abid, R. P. S. (2023). *PEMANFAATAN METODE CLUSTERING UNTUK MEMILIKI KETERAMPILAN TEKNOLOGI INFORMASI DAN KOMUNIKASI (TIK)*. 02(02), 36–41.
- Azis, A. F. (2025). *URGENSI SERTIFIKASI HALAL PRODUK MAKANAN UMKM DAN DAMPAKNYA TERHADAP KEPERCAYAAN KONSUMEN PERSPEKTIF HUKUM EKONOMI SYARIAH*. 3(01), 60–69.
- Badan Penyelenggara Jaminan Produk Halal. (2020). *RENCANA STRATEGIS BADAN PENYELENGGARA JAMINAN PRODUK HALAL KEMENTERIAN AGAMA TAHUN 2020 - 2024*. BPJPH Kementerian Agama RI. https://cmsbl.halal.go.id/uploads/Rencana_Strategis_BPJPH_tahun_2020_2024_2a96a649e3.pdf
- Farhan, A. (2018). *PELAKSANAAN SERTIFIKASI HALAL LPPOM MUI TERHADAP PRODUK USAHA MIKRO, KECIL DAN MENENGAH (UMKM) (STUDI LPPOM MUI PROVINSI BENGKULU)*. 3, 1–16.
- Faridah, H. D. (2019). *SERTIFIKASI HALAL DI INDONESIA: SEJARAH, PERKEMBANGAN, DAN IMPLEMENTASI*. 68–78.
- Firdaus, A., Firdaus, W. I., Studi, P., Informatika, T., Digital, M., & Sriwijaya, P. N. (2021). *Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan)*. 13(1), 66–78.
- Hajaroh, Suprpti, T., & Narasati, R. (2024). *IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN ULASAN PRODUK MAKANAN DAN MINUMAN DI TOKOPEDIA*. 8(1), 111–118.
- Hartini, & Malahayatie. (2024). *Implikasi Sertifikat Halal dalam Manajemen Bisnis Industri Makanan dan Minuman*. 1(2), 116–129.
- Hasan, A. H. (2025). *Indonesia to issue 1.35 million free halal certificates in 2026*. ANTARA INDONESIAN NEWS AGENCY. https://en.antaranews.com/news/397507/indonesia-to-issue-135-million-free-halal-certificates-in-2026?utm_source=chatgpt.com
- Hasanah, E. N., & Agustiniingsih, D. D. (2020). *Analysis of “ Halal ” Word i n Social Media Using Text Mining and Word Networking*. 409(SoRes 2019), 125–129.
- Hidayat, B. shafira. (2024). *Persepsi Konsumen terhadap Produk Halal dari UMKM : Dampak terhadap Keputusan Pembelian*. 4(4), 619–625.
- Hidayat, H. H., Arsi, P., & Rahmawati, L. I. (2021). *Pemetaan Kata Kunci dan Polaritas Sentimen Pengguna Twitter Terhadap Kehalalan Produk Keyword Mapping and Sentiments Polarity of Twitter User Towards Product Halalness*. 21(1). <https://doi.org/10.30812/matrik.v21i1.1162>
- Hikmah, N., Fadillah, A., Intan, S. M., Yunus, M., Islam, U., & Utara, S. (2025). *Tantangan Sertifikasi Halal pada UMKM di Indonesia : Studi Kualitatif terhadap Aspek Sosialisasi , Biaya , dan Regulasi*. 5, 78–89.

- Humaida, A., Fasicha, I. D., Alghifari, M. R., & Lestari, P. S. (2024). *Potensi Industri Halal di Indonesia sebagai Negara Berpenduduk*. 6(1), 11–24.
- Irham, M. A. (2024). BPJPH Terbitkan 1,8 Juta Sertifikat Halal. *Hukum Online Com*.
- Izza, N. N., & Akbar, N. (2023). *Uncovering Halal Label Food Purchase Intention Behavior in Indonesia: A Twitter Analytics and SEM-Based Investigation*. 10(6), 551–570. <https://doi.org/10.20473/vol10iss20236pp551-570>
- Kumala, N., & Dana, R. D. (2023). *PENERAPAN ALGORITMA K-MEANS CLUSTERING BERDASARKAN HASIL PENJUALAN PRODUK INEZ COSMETIC DI TOKO ROSALIA*. 18, 50–56.
- Legowati, D. A., & Albab, F. N. U. (2019). *Pengaruh Attitude, Sertifikasi Halal, Promosi dan Brand terhadap Purchase Intention di Restoran Bersertifikasi Halal*. 2 (1)(0274), 39–53. <https://doi.org/10.12928/ijiefb.v2i1.787>
- Mahardi, Y. F., & Dadang Karya Bakti, P. N. (2025). *Pengaruh Bahan Baku Halal dan Persepsi Halal terhadap Minat Pembelian Menu KFC di Bandar Lampung dengan Country of Origin sebagai Variabel Mediasi*. 8(2), 452–464.
- Mardianto, D., Parakkasi, I., & Muthiadin, C. (2025). Peran Sertifikat Halal dalam Meningkatkan Kepercayaan Konsumen pada Produk Industri Pangan di Indonesia. *Jurnal Ilmiah Ekonomi, Akuntansi, Dan Pajak*, 2.
- Maslahah, W., & Destiarni, R. P. (2025). *KLASTERISASI UMKM TERHADAP IMPLEMENTASI SERTIFIKASI HALAL DI KABUPATEN PAMEKASAN*. 2(1).
- Mustajab, S. (2025). *Smart Halal Business : Inovasi Digital Dalam Sertifikasi , Branding , Dan Kepercayaan Konsumen*. 7(2), 88–103. <https://doi.org/10.3390/logistics8010021.1>
- Nasya Tanasya, M. S. A. (2022). *Exploring Public Sentiments on Halal Food : A Social Network Analysis*. 82–92.
- Nawangsih, I. (2023). *Analisa Penjualan Produk Kosmetik Dengan Metode Algoritma KMeans Di Toko Erremy*. 4(1), 140–145.
- Nur, A., Zulkifli, A., & Shafie, N. A. (2024). *Review of the Lazada application on Google Play Store : Sentiment Analysis*. 9(1). <https://doi.org/10.24191/jcrinn.v9i1.412>
- Prastiwi, R. (2025). *Jurnal Ekonomi Islam PENGARUH GAYA HIDUP HALAL TERHADAP PERILAKU KONSUMSI GENERASI MILENIAL*. 2.
- Puspitasari, R., Mahmudah, M., Indriani, D., Indawati, R., & Ardianah, E. (2025). *JURNAL BIOMETRIKA DAN KEPENDUDUKAN (Journal of Biometrics and Population) CLUSTERING OF DRUG , COSMETIC , TRADITIONAL MEDICINE , AND FOOD CRIME VULNERABILITY IN EAST JAVA USING THE K- MEDOIDS ALGORITHM*. 14(May), 34–44.
- Ramadhan, R. F., Wijoyo, S. H., & Saputra, M. C. (2023). *Penerapan Metode K-Means Clustering pada Ulasan Perumahan PT XYZ di Google Maps untuk Formulasi Strategi Bisnis dengan Analisis SWOT*. 7(6), 2879–2888.
- Rivaro, M., Gozali, F., & Brata, D. W. (2026). *Analisis Sentimen Pengguna Twitter / X Terhadap Fenomena # KaburAjaDulu Menggunakan Metode Support Vector Machine dan IndoBERT Embedding*. 10(2), 1–10.

- Rizka, S. K., Purnamadewi, Y. L., & Hasanah, N. (2020). *Produk Roti dalam Pola Konsumsi Pangan dan Keberadaan Label Halal dalam Keputusan Konsumsi Masyarakat (Kasus: Kota Bogor)*. 6(1), 15–27. <https://doi.org/10.29244/jam.6.1.15-27>
- Salam, R. R., Jamil, M. F., & Ibrahim, Y. (2023). *Sentiment Analysis of Cash Direct Assistance Distribution for Fuel Oil Using Support Vector Machine Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggunakan Support Vector Machine*. 3(April), 27–35.
- Salima, D. J. N., & Ma'arif, M. R. (2024). *Indonesian Journal on Data Science Analisis Persepsi Publik di Media Sosial Terhadap Produk Halal Di Indonesia Menggunakan Text Mining*. 2(2), 67–76.
- Sucipto, S., Al-Mubarak, M. S., & Setiyawan, D. T. (2021). *HIERARCHICAL CLUSTERING BAHAN MENU DI KANTIN UNIVERSITAS UNTUK MENUNJANG IMPLEMENTASI SISTEM JAMINAN HALAL*. 31(1), 60–69.
- Susilawati, W., Aulia, R. N., Alamanda, D. T., Ramdani, A., & Kurniasih, A. (2021). *Analisis Persepsi Warganet Mengenai Halal Menggunakan Twitter*. 49~56.
- Syaifudin, Y. W., & Irawan, R. A. (2018). *IMPLEMENTASI ANALISIS CLUSTERING DAN SENTIMEN DATA TWITTER PADA OPINI WISATA PANTAI MENGGUNAKAN METODE K-MEANS*. 4, 189–194.
- Timur, Y. P., Ratnasari, R. T., Pitchay, A. A., & Jamilu, A. (2023). *Investigating Netizen Sentiment Toward Halal Certification in Indonesia Using Machine Learning*. 10(6), 525–540. <https://doi.org/10.20473/vol10iss20236pp525-540>
- Wakhidah, N. (2014). *CLUSTERING MENGGUNAKAN K-MEANS ALGORITHM (K-MEANS ALGORITHM CLUSTERING)*. 21 (1), 70–80.
- Wakhidah, N. (2024). *CLUSTERING MENGGUNAKAN K-MEANS ALGORITHM (K-MEANS ALGORITHM CLUSTERING)*. 5.
- Wulan Ariby, Naila Deswita, R. R. A. H. (2025). *UPAYA MENINGKATKAN KESADARAN KONSUMEN MUSLIM DALAM MEMILIH PRODUK HALAL DAN SYARIAH. Economic Insight: Journal of Economic and Management*. <https://journal.aspublisher.co.id/index.php/econinsight/index>
- Yuniar, N., & Heikal, J. (2024). *Segmentasi sertifikat halal pada wilayah Indonesia menggunakan k-means clustering di badan penyelenggara jaminan produk halal*. 1(1), 23–42.
- Zaskya, M., Boham, A., & Lotulung, L. J. H. (2020). *Twitter Sebagai Media Mengungkapkan Diri Pada Kalangan Milenial*. 1–8.
- Zuraidah, Nuzula, S. F., & Latifa, A. (2023). *Halal Certification as An Effort to Improve the Branding of Micro Small and Medium Enterprises Products in Keling Village, Kepung District, Kediri Regency*. 2(2), 92–98.

[illegible][illegible][illegible]

[illegible][illegible][illegible]

[illegible][illegible]

2. Hasil *Preprocessing*

[illegible][illegible][illegible]

[illegible]

Word Cloud - Sentimen Negatif



b. *Cluster 1* (Netral)

Word Cloud - Sentimen Netral



c. *Cluster 2* (Positif)

