



Buku Monograf

**Pemodelan Sentimen dalam Ulasan Produk
dan Penerapannya pada E-commerce**



Shabrina Rasyid Munthe, S.Pd.I., MS
Samsir, ST, M.Kom
Rina Asriani Levianti, SE, M.Ak
Dr. Basyarul Ulya, SH., MM
Nurhanna Harahap, M.Hum
Abdul Hakim Dalimunthe, M.Kom
M. Rusdi, S.Pd.I, MM
Nurhabibah Ritonga, S.E., M.M
Ibnu Rasyid Munthe, ST., M.Kom

Buku Monograf

Pemodelan Sentimen dalam Ulasan Produk dan Penerapannya pada E- commerce

Buku Monograf
Pemodelan Sentimen dalam Ulasan Produk
dan Penerapannya pada E-commerce

Shabrina Rasyid Munthe, S.Pd.I., M.S
Samsir, S.T., M.Kom
Rina Asriani Levianti, S.E., M.Ak
Dr. Basyarul Ulya, SH., MM
Nurhanna Harahap, M.Hum
Abdul Hakim Dalimunthe, M.Kom
M. Rusdi, S.Pd.I, MM
Nurhabibah Ritonga, S.E., M.M
Ibnu Rasyid Munthe, S.T., M.Kom



**BUKU MONOGRAF PEMODELAN SENTIMEN DALAM ULASAN
PRODUK DAN PENERAPANNYA PADA E-COMMERCE**

**Shabrina Rasyid Munthe, S.Pd.I., M.S, Samsir, S.T., M.Kom,
Rina Asriani Levianti, S.E., M.Ak, Dr. Basyarul Ulya, SH., MM,
Nurhanna Harahap, M.Hum, Abdul Hakim Dalimunthe, M.Kom,
M. Rusdi, S.Pd.I., M.M, Nurhabibah Ritonga, S.E., M.M,
Ibnu Rasyid Munthe, S.T., M.Kom**

Editor :
Sahat Parulian Sitorus

Desain Cover :
Nurhasanah

Sumber :
<https://isbn.jndi.my.id/products/buku-monograf-pemodelan-sentimen-dalam-ulasan-produk-dan-penerapannya-pada-e-commerce>

Tata Letak :
Ida Royani Sitorus

Ukuran :
Jml hal judul 219, Jml hal isi naskah 209, Uk: 15.5x23 cm

ISBN :
978-634-04-2224-5

Cetakan Pertama :
31 Juli 2025

Hak Cipta 2025, Pada Penulis
Isi diluar tanggung jawab percetakan

Copyright © 2025 by PT.JNDI
All Right Reserved

Hak cipta dilindungi undang-undang
Dilarang keras menerjemahkan, memfotokopi, atau
memperbanyak sebagian atau seluruh isi buku ini
tanpa izin tertulis dari Penerbit.

PENERBIT PT. JASA NIAGA DIGITAL INDONESIA
Jl. H. Maulana Link. VI. Labuhan Ruku, Talawi, Batu Bara, Sumatera Utara 21254
Telp/Wa : 082361274081
<https://isbn.jndi.my.id>
E-mail:office@jndi.my.id

PRAKATA

Puji syukur kehadiran Tuhan Yang Maha Esa atas rahmat dan karunia-Nya, penulis dapat menyelesaikan buku monograf ini dengan judul "*Pemodelan Sentimen dalam Ulasan Produk dan Penerapannya pada E-commerce*". Buku ini hadir sebagai hasil penelitian mendalam yang bertujuan mengungkap pentingnya analisis sentimen dalam memahami dinamika ulasan produk di platform e-commerce dan memberikan solusi praktis untuk tantangan yang dihadapi oleh perusahaan.

Dengan pesatnya perkembangan e-commerce, ulasan produk kini menjadi sumber informasi yang sangat berharga bagi konsumen dalam pengambilan keputusan pembelian. Oleh karena itu, analisis sentimen menjadi alat vital bagi perusahaan untuk memahami lebih baik preferensi dan keluhan pelanggan. Buku ini membahas secara komprehensif tentang metode analisis sentimen dan penerapannya, serta memberikan wawasan mengenai teknik-teknik canggih untuk mengatasi tantangan yang ada.

Diharapkan buku ini dapat memberikan manfaat dan kontribusi bagi akademisi, peneliti, serta praktisi di bidang e-commerce dan pemasaran digital.

TIM Penulis

KATA PENGANTAR

Buku ini mengungkap peran penting analisis sentimen dalam memahami ulasan produk di dunia e-commerce yang berkembang pesat. Dengan judul "*Pemodelan Sentimen dalam Ulasan Produk dan Penerapannya pada E-commerce*", buku ini menawarkan wawasan mendalam tentang teknik-teknik analisis sentimen yang dapat membantu perusahaan memahami preferensi dan keluhan konsumen. Penulis membahas berbagai metode ekstraksi fitur seperti TF-IDF, N-Gram, dan Word2Vec, serta teknik klasifikasi seperti Naive Bayes, SVM, dan Decision Tree untuk menanggulangi tantangan ketidakseimbangan kelas dalam dataset.

Buku ini dilengkapi dengan studi kasus nyata dari platform e-commerce besar seperti Shopee, Tokopedia, dan Lazada, yang memungkinkan pembaca memahami perbedaan perilaku konsumen di masing-masing platform. Dengan akurasi model SVM mencapai 95%, buku ini sangat cocok untuk akademisi, peneliti, dan praktisi di bidang e-commerce dan ilmu data. Diharapkan, buku ini memberikan kontribusi penting dalam pengembangan penelitian dan penerapan industri digital.

TIM Penulis

DAFTAR ISI

PRAKATA	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
Bab 1. Pendahuluan.....	1
Tantangan dalam Analisis Sentimen pada Ulasan Produk E-Commerce	4
Imbalance Kelas dalam Dataset Ulasan Produk E- Commerce	8
Dampak Ulasan Produk terhadap Keputusan Pembelian Konsumen	12
Pentingnya Analisis Sentimen bagi Perusahaan E- Commerce	15
Mengatasi Ketidakseimbangan Kelas dalam Model Analisis Sentimen	19
BAB 2. Tinjauan Konsep dan Kajian	24
Definisi dan Tujuan.....	24
Penerapan Teknik dalam Analisis Sentimen pada Ulasan Produk E-Commerce	29
Hubungan Antara Analisis Sentimen dan Keputusan Pembelian Konsumen dalam E-Commerce.....	32
Penerapan Teori Natural Language Processing (NLP) dalam Analisis Sentimen E-Commerce.....	35
BAB 3. Langkah-Langkah Pemrosesan Data	39
Persiapan Data	42
Feature Extraction Mengubah Teks Menjadi Fitur...	46
Pengembangan dan Pelatihan Model	66
Mengatasi Ketidakseimbangan Kelas.....	74
Mengukur Kinerja Model	80
BAB 4. Proses Ekstraksi Fitur	88
Teknik Ekstraksi Fitur dalam Analisis Sentimen	88

Menangkap Pola Frasa dalam Teks	100
Memahami Hubungan Semantik Antar Kata.....	105
Integrasi Fitur	111
BAB 5. Penanganan Ketidakseimbangan Kelas.....	115
Masalah Ketidakseimbangan Kelas dalam Analisis Sentimen	120
Teknik Resampling untuk Mengatasi Ketidakseimbangan Kelas.....	129
Distribusi Kelas Sebelum dan Sesudah SMOTEENN.....	132
Penerapan SMOTE dan SMOTEENN dalam Meningkatkan Akurasi Model	133
Keuntungan dan Tantangan dalam Penggunaan SMOTE dan SMOTEENN	137
BAB 6. Pengembangan Model Klasifikasi	142
Masalah Ketidakseimbangan Kelas dalam Analisis Sentimen	150
Evaluasi Kinerja Model	155
Perbandingan Model Klasifikasi.....	161
BAB 7. Ulasan Terhadap Hasil Model Klasifikasi	162
Pengaruh Teknik SMOTE dan SMOTEENN pada Distribusi Kelas.....	162
Performa Model Naive Bayes.....	167
Performa Model Support Vector Machine (SVM)...	170
Performa Model Decision Tree.....	173
Perbandingan Kinerja Model.....	176
Dampak Teknik SMOTE dan SMOTEENN	178
BAB 8. Distribusi Sentimen Berdasarkan Platform E-Commerce	182
Distribusi Sentimen di Platform E-Commerce.....	183
Shopee Dominasi Sentimen Positif.....	184
Tokopedia Distribusi Sentimen yang Seimbang....	185
Lazada dan Blibli Sentimen Moderat dengan Kecenderungan Positif	186
Bukalapak Keterlibatan Pengguna yang Terbatas.	186
Tantangan yang Dihadapi oleh Distribusi Sentimen yang Tidak Seimbang	187

Teknik Ekstraksi Fitur Multimodal untuk Mengatasi Variasi dalam Distribusi Sentimen	188
Peran Platform E-Commerce dalam Membentuk Sentimen	189
BAB 9. Kunci dan Arah Pemanfaatan	191
Arah Pemanfaatan	195
DAFTAR PUSTAKA	200

Bab 1.

Pendahuluan

Di tengah arus perkembangan teknologi digital yang pesat, platform e-commerce telah membawa transformasi besar dalam cara konsumen berinteraksi dengan produk dan layanan. Tak hanya mengubah proses transaksi menjadi lebih cepat dan efisien, e-commerce juga memengaruhi cara informasi disebarkan serta bagaimana keputusan pembelian diambil oleh konsumen. Salah satu elemen penting yang kini menjadi pusat perhatian adalah ulasan produk yang disampaikan oleh para pengguna.

Ulasan produk kini bukan sekadar opini, melainkan sumber informasi berharga yang mencerminkan pengalaman langsung konsumen. Melalui ulasan, konsumen lain dapat menilai kualitas, performa, dan kesesuaian produk dengan kebutuhan mereka. Ulasan ini mengandung berbagai ekspresi emosi seperti kepuasan, ketidakpuasan, dan harapan yang belum terpenuhi. Hal ini menjadikan sentimen dalam ulasan sebagai aspek penting yang perlu dipahami oleh perusahaan.

Sentimen yang muncul dalam ulasan mampu membentuk persepsi publik terhadap suatu produk atau layanan. Berbeda dari deskripsi produk yang bersifat promosi, ulasan dari konsumen bersifat jujur, otentik, dan penuh variasi ekspresi. Banyak dari ulasan ini mengandung bahasa yang

kompleks seperti sarkasme, idiom, dan metafora yang menantang untuk dipahami secara literal. Kalimat seperti "tidak seburuk yang saya bayangkan" bisa memiliki makna tersirat yang lebih condong pada ketidakpuasan ringan daripada pujian. Oleh karena itu, pemahaman mendalam terhadap konteks dan gaya bahasa menjadi kunci untuk menginterpretasi makna sentimen secara akurat.

Memahami sentimen dari ulasan tidak hanya bermanfaat untuk menilai persepsi konsumen, tetapi juga menjadi dasar dalam merancang strategi pemasaran yang lebih relevan dan personal. Perusahaan yang mampu menangkap dan menindaklanjuti aspirasi, keluhan, dan saran dari pelanggan melalui ulasan akan memiliki keunggulan dalam membentuk loyalitas dan kepuasan pelanggan. Penyesuaian produk, pengembangan fitur, atau peningkatan layanan dapat dirancang berdasarkan wawasan yang diperoleh dari sentimen konsumen.

Namun, pemanfaatan ulasan konsumen juga menghadapi tantangan, terutama ketidakseimbangan dalam jenis sentimen. Dalam banyak kasus, ulasan positif lebih dominan dibandingkan dengan ulasan negatif atau netral. Ketidakseimbangan ini bisa menyebabkan perusahaan kurang menyadari adanya masalah pada produk atau layanan, karena informasi kritis dari ulasan negatif tenggelam di antara

banyaknya pujian. Hal ini dapat berujung pada pengambilan keputusan yang bias dan kurang tepat.

Untuk itu, penting bagi perusahaan untuk memperhatikan sentimen minoritas yang mungkin justru mengandung masukan paling relevan untuk perbaikan. Pengabaian terhadap ulasan negatif bisa berarti kehilangan peluang untuk meningkatkan kualitas secara menyeluruh. Sebaliknya, kemampuan mengenali dan merespons dengan baik kritik dari konsumen akan memperkuat citra merek dan meningkatkan kepercayaan pelanggan.

Lebih jauh, ulasan konsumen tidak hanya mencerminkan opini terhadap produk, tetapi juga merefleksikan preferensi dan kebutuhan emosional mereka. Dengan memahami isi dan sentimen dalam ulasan, perusahaan dapat lebih tanggap terhadap tren pasar dan kebutuhan konsumen yang berubah. Respons yang cepat dan tepat terhadap dinamika ini akan memperkuat posisi kompetitif perusahaan dalam industri e-commerce yang sangat dinamis.

Terdapat pula hubungan erat antara sentimen dalam ulasan dengan keputusan pembelian. Sentimen positif dapat meningkatkan kepercayaan konsumen, sementara sentimen negatif dapat menurunkan minat beli. Oleh karena itu, memahami pola sentimen secara menyeluruh memungkinkan perusahaan untuk memetakan persepsi konsumen dan mengatur strategi komunikasi pemasaran yang lebih efektif.

Ulasan juga memiliki pengaruh besar terhadap loyalitas konsumen. Ketika pelanggan merasa bahwa suaranya didengar dan direspons secara konkret, mereka akan cenderung untuk tetap menggunakan produk atau layanan tersebut. Mereka pun akan lebih terbuka untuk memberikan ulasan yang konstruktif dan menjadi bagian dari komunitas pelanggan yang aktif. Dengan demikian, ekosistem e-commerce menjadi wadah interaksi dua arah yang memperkuat keterlibatan dan kepuasan konsumen.

Pada akhirnya, ulasan konsumen bukan hanya pelengkap dalam sistem digital, melainkan elemen penting dalam ekosistem bisnis modern. Peranannya dalam membangun hubungan antara konsumen dan perusahaan semakin signifikan. Melalui pendekatan yang cermat terhadap pemahaman sentimen, perusahaan dapat merancang kebijakan yang selaras dengan kebutuhan pelanggan serta menciptakan nilai yang berkelanjutan. Oleh karena itu, memposisikan ulasan sebagai sumber wawasan strategis adalah langkah bijak dalam menavigasi persaingan di era e-commerce saat ini.

Tantangan dalam Analisis Sentimen pada Ulasan Produk E-Commerce

Pemahaman terhadap respons konsumen dalam ulasan produk e-commerce memiliki peran penting dalam

mengungkap pandangan dan pengalaman mereka terhadap produk atau layanan tertentu. Meskipun memberikan banyak manfaat, terdapat tantangan yang cukup kompleks, terutama karena cara konsumen mengekspresikan pendapatnya sangat bervariasi dan sering kali menyimpan makna tersembunyi. Tidak hanya sekadar menyampaikan opini secara langsung, banyak ulasan disampaikan dengan nuansa emosional yang tidak eksplisit.

Beragamnya bentuk ekspresi menjadi salah satu tantangan utama. Konsumen kerap menggunakan gaya bahasa yang tidak langsung seperti sarkasme, idiom, atau ungkapan kiasan. Sebagai contoh, kalimat seperti "produk ini tidak seburuk yang saya bayangkan" sekilas tampak positif, padahal mengandung isyarat bahwa harapan awal terhadap produk cukup rendah. Dalam kasus seperti ini, makna yang terkandung tidak bisa ditangkap hanya dari kata-kata secara harfiah, melainkan perlu dipahami dalam konteks yang lebih luas.

Tidak jarang sebuah ulasan berisi campuran antara pujian dan keluhan. Konsumen dapat menyatakan kepuasan terhadap satu fitur, namun di saat yang sama menunjukkan ketidakpuasan terhadap aspek lain. Keberagaman ini menyulitkan dalam menilai arah opini secara menyeluruh, karena tidak semua isi ulasan bersifat konsisten. Untuk itu,

penting untuk melihat keseluruhan isi daripada menilai berdasarkan bagian tertentu saja.

Ragam produk yang diulas turut memperbesar kompleksitas tantangan. Setiap kategori produk memiliki karakteristik dan ekspektasi konsumen yang berbeda. Produk elektronik cenderung memperoleh ulasan dengan bahasa teknis dan kritis, sementara produk kecantikan atau fashion lebih banyak mendapatkan ulasan yang bersifat emosional dan subjektif. Perbedaan ini membuat pendekatan tunggal sulit diterapkan untuk semua jenis produk, karena masing-masing memiliki konteks unik yang perlu dipahami secara spesifik.

Gaya penyampaian antar pengguna pun sangat bervariasi. Beberapa konsumen menyampaikan pendapatnya dengan gaya lugas dan langsung, sedangkan yang lain lebih halus dan implisit. Perbedaan ini juga dipengaruhi oleh kondisi emosional saat menulis ulasan, serta latar belakang budaya dan kebiasaan berkomunikasi masing-masing individu. Penggunaan metafora, sindiran, atau humor memperumit proses pemahaman, karena makna sebenarnya tersembunyi di balik susunan kata yang tidak selalu jelas.

Selain itu, jumlah ulasan positif cenderung lebih banyak dibandingkan ulasan negatif. Ketidakseimbangan ini menimbulkan tantangan dalam menilai kualitas produk secara objektif. Dominasi opini positif dapat membuat umpan balik

kritis kurang terlihat, padahal komentar yang bersifat negatif sering kali mengandung masukan penting yang dapat dijadikan bahan perbaikan.

Dalam kenyataannya, struktur ulasan sering kali tidak sistematis. Banyak konsumen menulis ulasan dalam bentuk narasi panjang yang mencampurkan fakta dengan opini, menyisipkan gaya bahasa tidak formal, dan bahkan menambahkan unsur humor. Contoh seperti "produk ini cocok bagi yang tidak suka kualitas" memperlihatkan bagaimana makna bisa terselubung di balik ekspresi sarkastik. Situasi seperti ini memerlukan perhatian penuh terhadap konteks keseluruhan agar tidak terjadi kesalahan pemahaman.

Perbedaan konteks antar produk juga berpengaruh terhadap dampak dari ulasan itu sendiri. Ulasan negatif terhadap produk elektronik, misalnya, bisa sangat berpengaruh karena menyangkut performa dan keandalan. Sebaliknya, ulasan negatif terhadap produk fashion mungkin lebih bisa ditoleransi selama aspek visual atau gaya tetap memenuhi ekspektasi. Hal ini menunjukkan bahwa setiap kategori produk memiliki sensitivitas yang berbeda terhadap opini konsumen.

Banyak pula ulasan yang tidak menyampaikan inti permasalahan secara langsung. Informasi penting tersebar di beberapa bagian dan tidak mudah dikenali. Untuk itu, diperlukan ketelitian dalam menafsirkan keseluruhan pesan

agar informasi yang diperoleh dapat digunakan secara maksimal. Menyaring bagian yang paling mencerminkan emosi dan pengalaman konsumen menjadi langkah krusial dalam menyusun strategi yang efektif.

Imbalance Kelas dalam Dataset Ulasan Produk E-Commerce

Salah satu tantangan besar dalam memahami pandangan konsumen melalui ulasan produk e-commerce adalah ketidakseimbangan dalam jumlah opini yang tersedia. Ketika ulasan positif mendominasi dan ulasan bernuansa negatif atau netral menjadi minoritas, maka gambaran yang diperoleh pun menjadi kurang mewakili kenyataan. Ketimpangan ini tidak hanya memengaruhi keberagaman informasi yang bisa digali, tetapi juga mengganggu ketajaman sistem dalam mengenali makna dari berbagai ekspresi yang digunakan oleh konsumen.

Dominasi opini positif membuat sistem lebih terlatih untuk membaca pola yang umum ditemukan dalam ulasan yang penuh pujian, namun menjadi kurang peka terhadap nada-nada kekecewaan. Ulasan bernada negatif sering kali menyimpan kritik membangun dan informasi mendalam yang justru sangat penting bagi pengembangan produk dan peningkatan layanan. Jika ulasan jenis ini tidak mendapat perhatian yang sepadan, maka masukan bernilai tinggi bisa

saja terabaikan. Bahkan, ekspresi ketidakpuasan yang disampaikan secara halus berpotensi dianggap sebagai komentar netral atau positif karena cara penyampaiannya tidak eksplisit.

Kondisi ini semakin diperparah oleh kenyataan bahwa ulasan positif cenderung bersifat umum dan berulang, seperti sekadar menyatakan bahwa produk bagus atau memuaskan. Sebaliknya, komentar yang lebih kritis justru banyak memberikan rincian, penjelasan konteks, serta saran yang sangat berharga. Meskipun jumlahnya lebih sedikit, kualitas informasi dalam ulasan negatif sering kali lebih tinggi dan lebih strategis dalam mendukung pengambilan keputusan.

Ketika sistem terlalu sering mendapati ulasan positif, maka sensitivitas terhadap variasi opini menjadi berkurang. Dampaknya adalah sistem mulai mengabaikan bentuk-bentuk keluhan yang tidak disampaikan secara tegas. Hasil pembacaan pun tampak baik hanya karena mayoritas opini memang positif. Inilah yang menciptakan ilusi bahwa sistem bekerja dengan akurat, padahal sebenarnya ia hanya mencerminkan ketimpangan yang ada dalam data.

Salah satu kesalahan umum yang terjadi adalah menyandarkan keberhasilan pada ukuran seperti tingkat ketepatan, tanpa mempertimbangkan bahwa angka tersebut bisa saja tinggi hanya karena opini yang banyak adalah yang positif. Jika delapan dari sepuluh ulasan menyatakan

kepuasan, maka sistem yang selalu menilai semua ulasan sebagai positif akan tetap terlihat sukses. Namun dalam kenyataan, sistem tersebut gagal menangkap kritik yang justru diperlukan untuk perbaikan.

Ketimpangan ini juga berdampak pada bagaimana perusahaan memahami kebutuhan pelanggan. Ketika hanya suara-suara puas yang terdengar, maka bayangan tentang apa yang perlu ditingkatkan menjadi kabur. Keluhan yang tidak tersampaikan bisa menjadi pemicu ketidakpuasan yang lebih besar, bahkan berujung pada penurunan kepercayaan pelanggan. Oleh karena itu, menciptakan keseimbangan dalam representasi opini menjadi sangat penting agar pandangan yang muncul benar-benar mencerminkan keragaman pengalaman konsumen.

Terdapat pula kecenderungan perilaku konsumen yang turut berkontribusi pada ketidakseimbangan ini. Mereka yang merasa puas biasanya lebih terdorong untuk memberikan ulasan positif secara spontan, sedangkan mereka yang kecewa cenderung menyampaikan ketidakpuasan melalui jalur lain yang tidak terpantau. Hal ini menyebabkan persebaran opini menjadi kurang merata dan tidak mencerminkan keseluruhan spektrum pengalaman yang sebenarnya terjadi.

Memastikan adanya keberagaman opini yang tertampung akan membantu perusahaan dalam membangun

ekspektasi pelanggan secara lebih realistis. Ketika ulasan positif mendominasi, konsumen baru bisa saja memiliki harapan yang terlalu tinggi terhadap produk. Jika harapan tersebut tidak terpenuhi, maka kekecewaan yang timbul bisa lebih besar dibandingkan dengan situasi di mana informasi yang tersedia menggambarkan kelebihan dan kekurangan secara seimbang.

Isu distribusi opini yang tidak merata bukan semata-mata soal teknis, melainkan juga menyangkut bagaimana sebuah sistem memfasilitasi keterwakilan pengalaman konsumen secara adil. Semua bentuk ulasan, baik yang mendukung maupun yang mengkritik, merupakan bagian penting dari gambaran menyeluruh terhadap suatu produk atau layanan. Dalam ekosistem digital yang terbuka, setiap suara memiliki makna dan perlu mendapat ruang yang setara untuk disampaikan dan dipertimbangkan.

Mewujudkan sistem yang menghargai semua bentuk opini akan membantu perusahaan dalam membangun kepercayaan dan loyalitas pelanggan. Dengan memahami sepenuhnya apa yang disukai dan apa yang menjadi keluhan, perusahaan dapat bertindak lebih tanggap dan tepat sasaran. Ketika semua bentuk masukan dihargai, maka kualitas layanan pun akan tumbuh seiring dengan meningkatnya kepuasan dan rasa percaya dari para pelanggan.

Dampak Ulasan Produk terhadap Keputusan Pembelian Konsumen

Ulasan produk memegang peranan yang semakin dominan dalam menentukan keputusan konsumen dalam aktivitas belanja daring. Perubahan pola konsumsi yang beralih dari pengalaman langsung di toko fisik ke pengalaman berbasis digital telah meningkatkan ketergantungan konsumen terhadap informasi dari pihak ketiga yang tersedia secara daring. Dalam konteks ini, ulasan dari sesama konsumen berfungsi sebagai jembatan yang mempertemukan harapan calon pembeli dengan realitas penggunaan produk di lapangan, menggantikan interaksi langsung yang biasanya terjadi di lingkungan toko konvensional.

kekuatan ulasan terletak pada kemampuannya membentuk persepsi terhadap kualitas dan reliabilitas suatu produk. Konsumen cenderung lebih percaya pada pengalaman sesama pengguna dibandingkan dengan pernyataan promosi dari pihak produsen atau penjual. Hal ini menandai pergeseran penting dalam struktur informasi, di mana komunitas pengguna menjadi sumber utama referensi dan validasi.

Nilai tambah dari ulasan tidak terbatas pada informasi dasar, melainkan juga mencakup rincian tentang performa produk, kecepatan pengiriman, dan kualitas layanan purna jual. Ulasan yang menyertakan konteks pemakaian

memberikan dimensi baru bagi calon pembeli untuk menilai kesesuaian produk dengan kebutuhan mereka. Ini menjadikan ulasan sebagai referensi nyata, bukan sekadar opini singkat.

Tingginya jumlah ulasan positif terhadap suatu produk kerap memperkuat citra dan menarik kepercayaan konsumen baru. Mekanisme ini dikenal sebagai pengaruh sosial, di mana keputusan individu dipengaruhi oleh mayoritas pengalaman yang dilaporkan oleh konsumen lain. Semakin banyak ulasan positif, semakin kuat pula efek meyakinkan yang ditimbulkan, sehingga mendorong peningkatan konversi dari niat beli menjadi tindakan nyata.

Di sisi lain, ulasan negatif juga memainkan peranan penting dalam memberikan perspektif yang berimbang. Banyak konsumen secara aktif mencari komentar negatif untuk memahami potensi risiko sebelum membeli. Dalam dokumen yang sama disebutkan bahwa ulasan negatif sering kali dianggap lebih autentik, karena diasumsikan tidak dimotivasi oleh kepentingan promosi. Jika penjual memberikan tanggapan yang baik terhadap ulasan semacam ini, hal itu justru memperkuat kepercayaan karena menunjukkan adanya itikad baik dalam menyelesaikan masalah.

Ulasan produk terbukti memiliki peran penting dalam memengaruhi keputusan akhir konsumen. Banyak konsumen yang awalnya hanya melakukan penelusuran produk secara

pasif berakhir dengan pembelian setelah terpapar ulasan yang meyakinkan. Sebaliknya, paparan ulasan negatif yang konsisten dapat menurunkan minat bahkan membatalkan rencana pembelian.

Bentuk naratif dalam ulasan menjadi faktor penentu lain dalam menyampaikan informasi yang mendalam. Sebagaimana dijabarkan dalam Bab 2 halaman 16, ulasan teks memungkinkan konsumen menjelaskan secara rinci latar belakang, pengalaman pemakaian, dan kepuasan mereka terhadap produk. Ini berbeda dari sistem rating berbintang yang bersifat kuantitatif dan kadang menimbulkan ambiguitas. Dua produk yang memiliki nilai rata-rata bintang yang sama bisa saja dipersepsikan secara berbeda jika salah satunya didukung oleh ulasan naratif yang lebih kaya dan komprehensif.

Bagi perusahaan, ulasan merupakan sumber informasi yang tak ternilai. Ulasan membantu mengidentifikasi kebutuhan dan ekspektasi konsumen yang mungkin tidak terungkap melalui survei atau forum formal. Pemahaman terhadap pola komentar konsumen memungkinkan perusahaan merespons tren dan dinamika pasar secara lebih cepat dan efektif. Melalui ulasan, perusahaan juga dapat mengetahui kekuatan dan kelemahan produk mereka dari sudut pandang konsumen, sehingga memperkuat kemampuan untuk beradaptasi terhadap perubahan pasar.

Tanggapan perusahaan terhadap ulasan yang masuk mencerminkan sejauh mana mereka menghargai suara konsumen. Ketika konsumen merasa pendapat mereka didengar dan ditindaklanjuti, mereka cenderung membentuk hubungan emosional yang lebih kuat dengan merek. Hubungan ini membangun loyalitas yang didasari oleh kepercayaan dan pengalaman yang menyenangkan, dan bukan semata-mata oleh kualitas produk.

Ulasan juga memberikan kontribusi besar dalam membentuk pengalaman belanja daring yang lebih bermakna. Karena tidak adanya kontak langsung dengan produk sebelum pembelian, testimoni dari pengguna lain menjadi rujukan utama. Semakin jujur dan lengkap ulasan yang tersedia, semakin percaya diri konsumen dalam mengambil keputusan. Hal ini menciptakan pengalaman belanja yang lebih pasti, transparan, dan memuaskan.

Pentingnya Analisis Sentimen bagi Perusahaan E-Commerce

Dalam era digital yang terus berkembang, kemampuan perusahaan dalam memahami dinamika kebutuhan dan ekspektasi konsumen menjadi fondasi utama bagi keberlangsungan bisnis. Ulasan yang diberikan oleh pelanggan menjadi cerminan nyata dari pengalaman,

harapan, dan evaluasi mereka terhadap suatu produk atau layanan. Informasi ini sangat berharga karena membuka ruang bagi perusahaan untuk menangkap berbagai sinyal pasar yang tidak selalu terlihat dari data penjualan semata.

Beragam masukan dari pelanggan memberi pandangan luas tentang kualitas produk, efektivitas layanan, hingga aspek-aspek yang berkaitan dengan proses distribusi. Ketika perusahaan secara konsisten menyimak masukan ini, mereka tidak hanya mendapatkan gambaran performa, tetapi juga memahami bagian mana yang perlu diperbaiki dan potensi apa yang bisa terus dikembangkan.

Menelusuri pola tanggapan konsumen juga memberi arah dalam merancang strategi komunikasi yang tepat sasaran. Jika suatu fitur dalam produk mendapat respons positif, perusahaan dapat menonjolkannya dalam materi promosi. Sebaliknya, jika terdapat keluhan yang berulang terhadap aspek tertentu, maka itu menjadi indikator penting untuk segera mengambil tindakan korektif. Dengan begitu, reputasi perusahaan tetap terjaga di tengah kompetisi yang ketat.

Lebih jauh, pendekatan pemasaran yang selaras dengan preferensi konsumen menjadi lebih mudah diterapkan saat perusahaan memahami kecenderungan pelanggan. Komunikasi yang dibangun berdasarkan informasi nyata akan terasa lebih personal dan bermakna, memperkuat keterikatan

antara pelanggan dan merek. Hubungan ini, jika dijaga dengan baik, akan berkembang menjadi loyalitas yang tidak hanya ditunjukkan melalui pembelian berulang, tetapi juga dalam bentuk dukungan terhadap citra perusahaan.

Ketika pelanggan merasa didengarkan dan dihargai, mereka cenderung menjalin hubungan emosional yang lebih dalam. Tanggapan yang direspons secara tanggap menunjukkan bahwa perusahaan tidak hanya berorientasi pada transaksi, tetapi juga pada kepuasan jangka panjang. Hal ini memperkuat kepercayaan yang menjadi dasar bagi hubungan yang berkelanjutan.

Tidak hanya itu, keterlibatan pelanggan dalam menyampaikan pandangan atau saran secara aktif juga memberikan nilai tambah. Mereka bukan sekadar konsumen, melainkan mitra yang berkontribusi dalam proses pembentukan kualitas layanan dan produk. Kolaborasi semacam ini menciptakan ekosistem bisnis yang lebih inklusif dan terbuka terhadap perubahan.

Dengan memperhatikan dinamika opini konsumen, perusahaan dapat menangkap sinyal perubahan tren lebih awal. Hal ini memungkinkan perusahaan untuk menyusun strategi yang responsif dan tetap relevan di tengah perkembangan pasar. Adaptasi yang cepat terhadap perubahan preferensi menjadi kunci dalam menjaga keunggulan kompetitif.

Penerapan masukan pelanggan juga memperkuat aspek layanan. Dengan mengenali titik lemah dalam pengalaman pengguna, perusahaan dapat memperbaiki sistem pelayanan menjadi lebih efisien dan ramah. Hal ini tidak hanya meningkatkan kepuasan, tetapi juga mendorong persepsi positif yang berkelanjutan di benak konsumen.

Kepekaan terhadap suara konsumen berdampak langsung pada posisi perusahaan dalam persaingan. Perusahaan yang cepat merespons kebutuhan nyata akan lebih mudah diterima oleh pasar. Mereka tidak hanya hadir sebagai penyedia barang dan jasa, melainkan juga sebagai entitas yang memahami dan peduli terhadap konsumen.

Lebih dari itu, masukan pelanggan kerap menjadi sumber inspirasi untuk perbaikan dan inovasi. Banyak ide segar muncul dari kritik atau saran yang diberikan secara jujur dan terbuka. Perusahaan yang terbuka terhadap pandangan ini akan lebih siap menciptakan solusi yang sesuai dengan kondisi nyata, bukan sekadar hasil asumsi internal.

Keberlanjutan pertumbuhan perusahaan akan semakin kuat jika suara pelanggan dijadikan bagian integral dalam pengambilan keputusan. Keterlibatan ini memperkuat fondasi strategis yang mendukung reputasi, konsistensi mutu, serta relevansi produk dan layanan.

Kesadaran untuk menjadikan pandangan pelanggan sebagai acuan utama dalam pengembangan bisnis adalah

bentuk komitmen terhadap kualitas dan kepuasan. Perusahaan yang membangun sistem responsif terhadap masukan ini akan tampil sebagai pemimpin yang berwawasan jauh ke depan, siap menyongsong tantangan baru dan menjaga keberlanjutan relasi dengan konsumen secara konsisten.

Mengatasi Ketidakseimbangan Kelas dalam Model Analisis Sentimen

Dalam dunia perdagangan digital yang berkembang pesat, umpan balik pelanggan menjadi fondasi penting bagi perusahaan dalam memahami pengalaman yang sesungguhnya dirasakan oleh konsumen. Ulasan pelanggan menyimpan beragam ekspresi yang mencerminkan kepuasan maupun ketidakpuasan terhadap produk atau layanan, yang secara tidak langsung membentuk persepsi publik terhadap merek.

Namun, realitas distribusi tanggapan pelanggan sering kali tidak seimbang. Tanggapan positif cenderung lebih dominan dibandingkan dengan masukan bernada netral atau negatif. Ketidakseimbangan ini memengaruhi kemampuan sistem dalam menginterpretasikan keseluruhan suara pelanggan, karena sistem cenderung memberikan perhatian lebih besar pada kelompok tanggapan yang paling banyak

muncul. Hal ini membuat keluhan atau kritik yang sebetulnya bernilai tinggi berisiko terabaikan.

Masukan bernada negatif mengandung banyak informasi yang dapat dijadikan dasar untuk pengembangan lebih lanjut. Namun, apabila sistem tidak dirancang untuk menangkap suara ini dengan proporsional, potensi perbaikan menjadi terhambat. Oleh karena itu, penting bagi perusahaan untuk membangun sistem yang memperhatikan keseimbangan data sejak awal.

Salah satu langkah untuk mencapai hal ini adalah dengan memperbanyak representasi dari kelompok data yang kurang terwakili. Teknik seperti oversampling dapat digunakan untuk menciptakan salinan tambahan dari tanggapan yang jarang muncul, sehingga pola-pola yang ada tetap bisa dikenali secara jelas. Dengan penguatan ini, sistem akan memiliki kecenderungan yang lebih adil dalam menilai seluruh kategori tanggapan.

Pendekatan lain yang dapat diterapkan adalah mengurangi atau menghapus tanggapan yang bersifat ambigu dari kelompok yang dominan. Langkah ini membantu memperjelas struktur data yang digunakan, sehingga proses klasifikasi dapat berjalan lebih tepat sasaran. Gabungan dari dua pendekatan ini menjadikan hasil klasifikasi lebih seimbang dan mencerminkan kondisi lapangan secara lebih akurat.

Hasil yang diperoleh dari sistem yang tidak bias memiliki nilai praktis tinggi bagi pengambilan keputusan dalam bisnis. Perusahaan dapat menyusun rencana yang lebih terarah berdasarkan pemahaman menyeluruh atas kekuatan dan kelemahan yang ada. Tidak hanya sisi produk yang ditingkatkan, namun juga strategi komunikasi dan layanan yang lebih tepat sasaran.

Keseimbangan dalam memahami suara pelanggan juga berdampak pada terciptanya persepsi yang adil terhadap performa suatu produk. Ketika informasi dari semua sisi diperhitungkan secara setara, maka strategi pengembangan tidak lagi bersifat sepihak, melainkan mencerminkan kebutuhan nyata konsumen. Hal ini menciptakan dasar yang kuat untuk pengembangan produk dan layanan yang lebih berkualitas.

Sistem yang terbuka terhadap semua jenis masukan akan meningkatkan rasa percaya dari pelanggan. Konsumen merasa lebih dihargai ketika pendapat mereka, baik yang positif maupun negatif, mendapatkan perhatian yang sama. Ini memperkuat hubungan antara perusahaan dan pelanggan, serta mendukung terciptanya loyalitas jangka panjang.

Penerapan sistem seperti ini tidak hanya terbatas pada sektor komersial, tetapi juga bermanfaat di berbagai sektor lain seperti layanan publik, pendidikan daring, dan layanan kesehatan. Keberadaan sistem yang seimbang membuka jalan

bagi terciptanya kebijakan yang lebih adil dan responsif, yang berdasar pada data nyata dari pengguna layanan.

Sebagaimana dijelaskan dalam dokumen yang sama pada Bab 4 halaman 33, ketidakseimbangan dalam struktur data dapat memengaruhi hasil pengambilan keputusan secara signifikan. Dengan memperkecil ketimpangan ini, informasi yang digunakan menjadi lebih objektif. Hal ini memungkinkan perusahaan untuk lebih sigap dalam menyesuaikan langkah bisnisnya terhadap perubahan pasar.

Selain dari aspek distribusi data, kemampuan sistem dalam memahami makna tersembunyi dalam bahasa juga sangat diperlukan. Ulasan konsumen tidak selalu disampaikan secara langsung, namun sering kali menggunakan kiasan, ironi, atau bahasa yang bersifat emosional. Sistem yang mampu menangkap nuansa ini akan memberikan hasil yang lebih menggambarkan kenyataan.

Kombinasi antara distribusi data yang seimbang dan kemampuan memahami bahasa secara mendalam akan menghasilkan sistem yang lebih cerdas dan adil. Dengan dasar ini, perusahaan dapat menyusun strategi berdasarkan pemahaman yang utuh terhadap kebutuhan dan harapan konsumen.

Upaya ini tidak hanya bertujuan mengoreksi kekurangan yang ada, tetapi juga untuk menciptakan pendekatan yang lebih proaktif. Sistem yang mampu mengidentifikasi

perubahan secara dini dapat menjadi alat bantu yang efektif untuk merancang langkah-langkah bisnis yang adaptif dan tepat sasaran.

BAB 2.

Tinjauan Konsep dan Kajian

Definisi dan Tujuan

Dasar-dasar pemikiran yang berkaitan dengan bagaimana konsumen menyampaikan pandangan mereka dalam ulasan serta bagaimana hal tersebut dimanfaatkan dalam aktivitas perdagangan digital. Dalam dunia perdagangan elektronik, tanggapan pelanggan merupakan refleksi dari pengalaman mereka selama menggunakan suatu produk atau layanan. Ulasan tersebut bukan sekadar opini, tetapi juga berperan sebagai indikator kualitas dan efektivitas produk dari sudut pandang pengguna.

Tujuan utama dari pembacaan terhadap ulasan ini adalah untuk memahami arah umum dari opini konsumen, apakah bersifat positif, negatif, atau netral. Klasifikasi tersebut memberikan gambaran mengenai persepsi publik terhadap suatu produk dan membantu dalam menyusun strategi penyempurnaan. Melalui tanggapan yang disampaikan konsumen, perusahaan dapat mengenali kepuasan yang dirasakan maupun kendala yang dihadapi, sehingga respons yang diberikan menjadi lebih terarah dan tepat.

Informasi dalam ulasan yang dibagikan di ruang digital terbuka memiliki pengaruh besar terhadap pengambilan keputusan calon konsumen lainnya. Setiap ekspresi yang disampaikan secara eksplisit maupun implisit memberikan gambaran atas kualitas layanan atau produk yang telah digunakan. Perusahaan yang mampu merespons secara tepat terhadap ragam tanggapan ini memiliki peluang lebih besar dalam memperkuat posisinya di pasar.

Untuk memahami makna dari isi ulasan, berbagai pendekatan dapat diterapkan, seperti TF-IDF yang menekankan pentingnya kata-kata tertentu berdasarkan frekuensinya dalam dokumen, serta N-Gram yang menggali pola-pola frasa yang sering muncul dalam narasi pelanggan. Word2Vec memperluas pendekatan ini dengan memberikan pemahaman berbasis kedekatan antar kata, sehingga hubungan semantik dapat dikenali secara lebih akurat.

Semua pendekatan ini mendukung pemahaman yang lebih menyeluruh terhadap bagaimana konsumen menyampaikan pengalaman mereka, baik secara lugas maupun melalui gaya bahasa yang kiasan. Perusahaan yang memiliki kemampuan untuk menangkap pola ini akan lebih siap menghadirkan solusi sesuai dengan harapan pengguna. Tanggapan konsumen pun menjadi jejak dinamis atas perjalanan mereka dalam berinteraksi dengan produk.

Perubahan gaya komunikasi digital menuntut perusahaan untuk senantiasa beradaptasi. Sistem yang responsif dan terbuka terhadap variasi ekspresi akan memiliki keunggulan dalam memahami dan mengakomodasi kebutuhan pelanggan. Dengan demikian, pengelolaan tanggapan ini tidak hanya memperkuat pemahaman internal, tetapi juga menciptakan pengalaman pengguna yang lebih berkesan dan berarti.

Teknik-Teknik dalam Analisis Sentimen

Untuk mencapai hasil analisis yang akurat, berbagai teknik digunakan dalam memproses dan menganalisis teks ulasan produk. Tiga teknik utama yang banyak digunakan dalam penelitian ini adalah TF-IDF, N-Gram, dan Word2Vec. Pemahaman terhadap dasar teori dari masing-masing teknik ini sangat penting sebelum membahas penerapannya dalam analisis sentimen pada ulasan produk.

a) TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF adalah salah satu teknik dasar yang digunakan dalam analisis sentimen untuk mengevaluasi pentingnya sebuah kata dalam dokumen, baik dalam dokumen tertentu maupun dalam keseluruhan korpus teks. Teknik ini terdiri dari dua komponen utama, yaitu *Term*

Frequency (TF) dan *Inverse Document Frequency* (IDF). TF mengukur seberapa sering sebuah kata muncul dalam suatu dokumen, sementara IDF mengukur seberapa jarang kata tersebut muncul di seluruh dokumen dalam korpus. Dengan demikian, TF-IDF memberikan bobot yang lebih tinggi pada kata-kata yang sering muncul dalam dokumen tertentu, tetapi jarang muncul dalam korpus secara keseluruhan. Teknik ini sangat berguna dalam ekstraksi fitur analisis sentimen, karena dapat mengidentifikasi kata-kata kunci yang relevan, seperti kata-kata yang mengindikasikan emosi positif atau negatif dalam ulasan produk (Chen & Zhu, 2023).

b) **N-Gram**

N-Gram adalah teknik yang digunakan untuk mengidentifikasi pola frasa dalam teks dengan menggabungkan urutan kata yang muncul berdekatan. Dalam analisis sentimen, teknik ini tidak hanya melihat kata-kata individual, tetapi juga hubungan antar kata yang berdekatan dalam sebuah kalimat, seperti bigram (dua kata berurutan) dan trigram (tiga kata berurutan). Dengan menggunakan N-Gram, analisis sentimen dapat menangkap yang lebih luas, karena seringkali sentimen suatu ulasan tidak hanya ditentukan oleh kata tunggal, tetapi oleh gabungan kata yang membentuk frasa. Misalnya, frasa seperti "sangat puas" atau "tidak puas

sama sekali" memiliki arti yang jelas dan memberikan petunjuk tentang sentimen ulasan tersebut. Pendekatan N-Gram memungkinkan model untuk mendeteksi sentimen yang lebih kompleks dan lebih mendalam, sehingga dapat menghasilkan analisis sentimen yang lebih tepat (Setiawan et al., 2022).

c) **Word2Vec**

Word2Vec adalah teknik pemodelan vektor kata yang lebih canggih yang digunakan untuk merepresentasikan kata-kata dalam ruang vektor berdimensi tinggi. Dengan menggunakan Word2Vec, setiap kata dalam sebuah korpus teks akan diubah menjadi vektor numerik yang merepresentasikan makna semantiknya berdasarkan kata tersebut. Teknik ini sangat berguna dalam analisis sentimen, karena dapat menangkap hubungan semantik antara kata-kata. Misalnya, kata "puas" dan "senang" akan berada dalam posisi vektor yang sangat dekat satu sama lain, meskipun keduanya memiliki bentuk yang berbeda. Dengan model ini, analisis sentimen menjadi lebih mudah karena model dapat mengenali kata-kata yang memiliki sentimen serupa meskipun berbeda dalam bentuk linguistiknya. Teknik ini juga memungkinkan analisis sentimen yang lebih mendalam, karena dapat

mendeteksi kata-kata yang secara semantik berhubungan dengan emosi tertentu (Wang et al., 2022).

Penerapan Teknik dalam Analisis Sentimen pada Ulasan Produk E-Commerce

Pemahaman terhadap makna yang tersirat maupun tersurat dalam setiap kata atau frasa yang digunakan konsumen menjadi kebutuhan mendasar bagi perusahaan yang ingin mempertahankan keunggulan di pasar. Dalam hal ini, tiga pendekatan utama yang sering digunakan adalah TF-IDF, N-Gram, dan Word2Vec, yang masing-masing memberikan kontribusi berbeda dalam menangkap pola-pola bahasa yang digunakan oleh konsumen.

TF-IDF, atau singkatan dari Term Frequency-Inverse Document Frequency, berguna dalam menentukan kata-kata mana yang paling menonjol dalam sebuah kumpulan ulasan. Pendekatan ini memberikan bobot lebih besar terhadap kata yang muncul sering dalam satu ulasan tetapi jarang dalam ulasan lain, sehingga memungkinkan perusahaan untuk mengidentifikasi kata-kata yang memiliki makna khusus bagi pengalaman konsumen. Misalnya, kata seperti "lembut," "berisik," atau "tahan lama" sering kali menjadi indikator kualitas yang diperhatikan pengguna dalam membentuk kesan mereka terhadap suatu produk. Melalui pendekatan ini,

perusahaan dapat mengakses bagian-bagian penting dari persepsi konsumen yang mungkin tidak langsung terlihat.

Lebih lanjut, pendekatan N-Gram memperluas pemahaman dengan mengamati hubungan antar kata dalam satu kalimat atau frasa. Alih-alih melihat satu kata secara terpisah, N-Gram menangkap pola-pola dua kata atau lebih yang sering digunakan bersama. Contohnya, ungkapan seperti "tidak memuaskan" atau "sangat baik" menyampaikan makna yang lebih kaya dibandingkan jika hanya memeriksa kata "tidak" atau "baik" secara individual. Dengan mengidentifikasi frasa yang sering muncul dalam ulasan, pendekatan ini membantu mengungkap dinamika komunikasi yang lebih alami dan kompleks sebagaimana ditulis oleh pengguna yang sesungguhnya.

Sementara itu, pendekatan Word2Vec memperkenalkan dimensi semantik dengan menyusun hubungan antar kata berdasarkan kemiripan makna dalam konteks penggunaannya. Kata-kata yang sering muncul bersama atau dalam lingkungan yang serupa akan dianggap memiliki kedekatan secara makna. Ini memungkinkan sistem memahami bahwa kata "cepat" dan "efisien" dalam banyak situasi memiliki arti yang berdekatan meskipun secara literal berbeda. Dalam lingkungan ulasan yang penuh gaya bahasa kiasan, sindiran, atau ekspresi tidak langsung, pendekatan ini

memberikan nilai lebih karena mampu menjembatani perbedaan antara kata dan makna.

Dengan menggabungkan ketiga pendekatan ini, perusahaan dapat memperoleh gambaran menyeluruh mengenai bagaimana konsumen merespons produk atau layanan yang mereka gunakan. TF-IDF membantu dalam menyoroti kata kunci, N-Gram menangkap pola frasa penting, dan Word2Vec memperluas pemahaman melalui hubungan makna antar kata. Kolaborasi antara ketiganya menciptakan dasar pemahaman yang kuat dan mendalam terhadap teks ulasan, sehingga perusahaan tidak hanya mengetahui kesan konsumen, tetapi juga mampu menafsirkan harapan dan kebutuhan yang mendasarinya.

Manfaat dari penerapan ketiga pendekatan ini tidak terbatas pada mengenali pujian dalam ulasan positif. Lebih dari itu, pendekatan ini juga memungkinkan identifikasi peringatan halus dalam ulasan netral maupun negatif yang sering kali luput jika hanya diperhatikan secara permukaan. Kemampuan untuk menangkap detail semacam ini menjadikan perusahaan lebih tanggap dan siap untuk memperbaiki layanan atau produk sebelum masalah menjadi lebih besar. Dalam lingkungan digital yang terus berubah, keberhasilan perusahaan dalam memanfaatkan pendekatan seperti ini akan sangat menentukan dalam menciptakan

strategi komunikasi dan layanan yang tepat sasaran serta bernilai jangka panjang.

Hubungan Antara Analisis Sentimen dan Keputusan Pembelian Konsumen dalam E-Commerce

Perkembangan platform digital, ulasan konsumen telah menjadi salah satu referensi utama dalam menilai kelayakan produk. Pengalaman yang disampaikan dalam bentuk narasi pribadi tidak hanya menyampaikan pandangan individu, tetapi juga membentuk persepsi kolektif yang berpengaruh terhadap ekspektasi dan keputusan calon pembeli. Dalam konteks ini, ulasan berperan sebagai bentuk komunikasi horizontal antarpengguna yang turut membentuk ekosistem digital yang dinamis.

Dalam berbagai platform, isi ulasan dapat menimbulkan respons emosional yang secara langsung memengaruhi penilaian akhir konsumen terhadap suatu produk. Ketika pengalaman yang dibagikan bersifat positif—seperti kemudahan penggunaan, kualitas produk, atau pelayanan yang memuaskan—calon pembeli akan terdorong untuk melakukan transaksi. Sebaliknya, kehadiran komentar yang mencerminkan ketidakpuasan, kekecewaan, atau masukan negatif cenderung melemahkan niat untuk membeli. Dinamika

ini menunjukkan bahwa nada emosional dalam ulasan dapat menjadi pemicu utama dalam proses pertimbangan konsumen.

Kekuatan ulasan yang lahir dari pengalaman otentik konsumen menjelaskan alasan mengapa banyak individu menjadikan ulasan sebagai pijakan utama dalam pengambilan keputusan. Di era minimnya interaksi fisik, validasi dari pengalaman konsumen lain menawarkan rasa aman dan membentuk kepercayaan yang lebih kuat dibandingkan promosi tradisional. Dalam banyak kasus, produk dengan harga yang lebih tinggi namun memiliki ulasan positif yang melimpah tetap menjadi pilihan utama dibandingkan produk serupa yang kurang mendapat perhatian dari pengguna.

Oleh sebab itu, perusahaan e-commerce perlu memberikan perhatian serius terhadap isi ulasan yang mereka terima. Ulasan yang kaya informasi, disampaikan dengan nada positif dan rinci, memiliki peran besar dalam membentuk citra merek. Sebaliknya, ulasan yang pendek, tidak jelas, atau bernada negatif dapat merusak persepsi calon pelanggan. Maka dari itu, pengelolaan ulasan harus menjadi bagian dari strategi komunikasi yang terintegrasi untuk membangun kredibilitas dan mempertahankan loyalitas pelanggan.

Salah satu hal yang ditunjukkan dalam dokumen yang menjadi acuan, terutama di bagian yang menjabarkan persepsi konsumen, adalah bagaimana satu ulasan yang

menarik dan informatif dapat menyebarkan pengaruh lebih luas dari yang dibayangkan. Algoritma platform digital cenderung menampilkan ulasan dengan interaksi tinggi di posisi yang lebih terlihat, mendorong efek berantai yang memperbesar dampak ulasan terhadap perilaku belanja pengguna lain. Ketika ulasan positif mendorong pembelian, dan pembelian menghasilkan lebih banyak ulasan, terciptalah siklus yang memperkuat posisi suatu produk dalam pasar.

Ditinjau dari sudut psikologis, pengalaman yang dikonfirmasi oleh banyak pengguna memperkuat rasa percaya terhadap produk tertentu. Apabila mayoritas ulasan mencerminkan kepuasan, maka tingkat keyakinan konsumen meningkat. Namun, apabila banyak ulasan menunjukkan kekecewaan, maka potensi pembelian akan menurun, sekalipun produk memiliki fitur unggulan secara teknis. Hal ini menegaskan bahwa nilai yang terbentuk dari pengalaman subjektif mampu menggeser penilaian terhadap nilai teknis produk.

Dalam menghadapi fenomena ini, perusahaan perlu mendorong terbentuknya kebiasaan berbagi pengalaman secara terbuka. Mendorong testimoni konsumen melalui sistem penghargaan, menyediakan platform yang mudah digunakan untuk meninggalkan ulasan, serta memberikan respons yang sopan dan informatif terhadap masukan yang diberikan merupakan langkah nyata untuk menunjukkan

perhatian terhadap suara pelanggan. Tindakan ini juga mencerminkan bahwa perusahaan berkomitmen membangun hubungan jangka panjang dengan konsumennya.

Namun perlu dipahami bahwa tidak semua ulasan memiliki tingkat pengaruh yang sama. Ulasan yang disusun secara personal, empatik, dan jujur cenderung lebih berkesan dibandingkan ulasan yang bersifat umum dan repetitif. Bahkan, satu ulasan yang ditulis dengan tulus dapat memicu rasa percaya yang lebih kuat daripada banyak ulasan yang tidak menunjukkan pengalaman nyata.

Penerapan Teori Natural Language Processing (NLP) dalam Analisis Sentimen E-Commerce

pentingnya teori Natural Language Processing (NLP) dalam memperluas pemahaman terhadap preferensi konsumen dan kecenderungan perilaku mereka di lingkungan e-commerce. Sebagai bagian dari kecerdasan buatan, NLP bertujuan memfasilitasi sistem komputer agar mampu membaca dan merespons bahasa manusia dengan akurasi dan efisiensi yang lebih tinggi. Dalam industri e-commerce, pemanfaatan NLP menjadi fondasi utama dalam membangun sistem cerdas yang dapat menafsirkan tanggapan pelanggan secara lebih terstruktur dan menyeluruh.

Proses pemahaman terhadap teks melalui NLP mencakup tahapan-tahapan teknis seperti tokenisasi, pengenalan entitas, dan penandaan struktur ujaran. Tokenisasi memungkinkan pemisahan kalimat menjadi satuan kata untuk memudahkan pemetaan makna. Pengenalan entitas bertugas mengidentifikasi nama produk, merek, atau kategori secara eksplisit dalam ulasan, sementara penandaan bagian ujaran memberi konteks gramatikal pada setiap kata yang digunakan, menjadikan sistem mampu menangkap maksud pengguna secara tepat.

Dalam dokumen yang diacu, penerapan NLP dijelaskan melalui cara kerja sistem dalam mengelola data tidak terstruktur dari ulasan pelanggan. Proses ini menghasilkan representasi yang menggambarkan pola atau kecenderungan konsumen terhadap produk tertentu. Melalui data yang terus diperbarui, sistem mampu menyesuaikan diri dalam menangkap ekspresi yang muncul dalam opini baru.

Keunggulan NLP semakin tampak ketika teknologi ini digabungkan dengan sistem klasifikasi yang mampu membaca perubahan opini secara dinamis. Jika muncul kecenderungan penurunan kepuasan pada fitur produk tertentu, sistem dapat memberikan sinyal peringatan yang dapat digunakan oleh perusahaan untuk segera melakukan pembaruan atau evaluasi. Ini mencerminkan bahwa aspirasi pelanggan tidak hanya didengar, tetapi juga ditindaklanjuti secara aktif.

Lebih jauh, pendekatan-pendekatan seperti TF-IDF, N-Gram, dan Word2Vec dijelaskan dalam dokumen sebagai pelengkap proses NLP. TF-IDF digunakan untuk menilai kata-kata yang paling representatif dari teks yang dianalisis, sedangkan N-Gram berfungsi untuk mengungkapkan susunan kata yang membentuk pola ungkapan tertentu. Word2Vec memberikan dimensi tambahan dengan memetakan hubungan semantik antar kata dalam suatu ruang makna yang lebih luas.

Dalam praktiknya, NLP memungkinkan perusahaan memahami bahasa alami konsumen dan mengubahnya menjadi informasi yang mudah digunakan. Ulasan konsumen yang kaya akan nuansa pengalaman sering kali menjadi acuan penting dalam pengambilan keputusan pembelian. Oleh sebab itu, sistem berbasis NLP berfungsi sebagai jembatan antara ekspektasi pelanggan dan strategi komunikasi produk yang diterapkan oleh perusahaan.

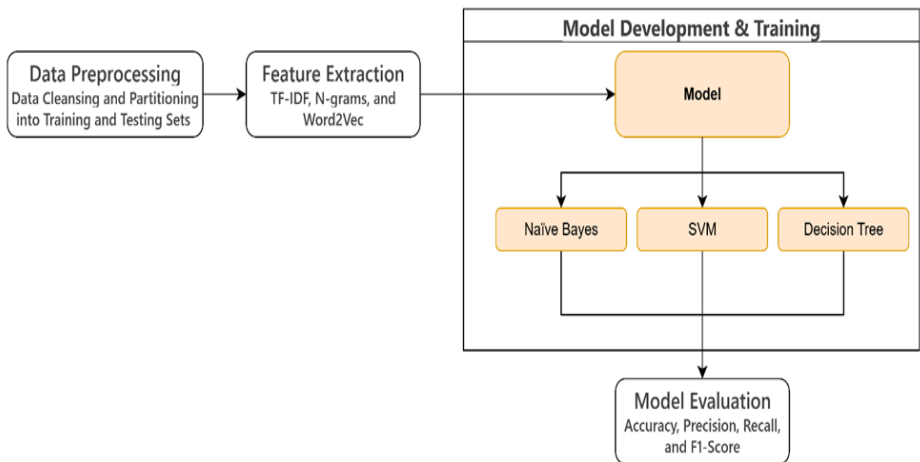
Teknologi ini juga memberikan kelebihan dalam menangani volume data yang sangat besar secara otomatis. Setiap hari, ratusan hingga ribuan ulasan masuk ke dalam sistem dan dengan bantuan NLP, semua data ini dapat disusun, diklasifikasi, dan diinterpretasi dalam waktu yang efisien. Hal ini mempercepat proses pengambilan keputusan serta memastikan bahwa langkah-langkah yang diambil sesuai dengan kebutuhan pasar yang bergerak cepat.

Manfaat NLP tidak hanya dirasakan oleh perusahaan, tetapi juga berdampak pada pengalaman konsumen. Sistem yang cerdas dapat menyaring dan menampilkan ulasan yang paling relevan, memfasilitasi pencarian informasi, dan menyajikan rekomendasi yang sesuai dengan minat pengguna. Ini menjadikan pengalaman berbelanja lebih nyaman, terarah, dan terasa personal.

Dengan semua manfaat tersebut, NLP tidak lagi hanya berperan sebagai alat teknis, melainkan telah menjadi bagian integral dalam proses bisnis digital. Penerapannya mendorong efisiensi dalam pengelolaan operasional, meningkatkan kecepatan dalam menanggapi dinamika pasar, serta memperkuat hubungan emosional antara merek dan konsumennya.

BAB 3.

Langkah-Langkah Pemrosesan Data



Gambar 1. Alur Proses Sentimen Ulasan E-Commerce

Pengolahan data dari ulasan produk e-commerce memerlukan struktur yang terorganisasi dan sistematis untuk menghasilkan informasi yang mencerminkan pemahaman menyeluruh terhadap persepsi konsumen. Gambar 1 dalam dokumen ini menunjukkan tahapan berurutan yang digunakan dalam pengolahan data berbasis ulasan, dimulai dari pengumpulan data mentah dari platform digital hingga ke tahap pengelompokan dan pemeriksaan hasil.

Langkah awal mencakup proses penyaringan yang bertujuan menghapus elemen-elemen yang tidak diperlukan dalam teks seperti tanda baca, angka, dan simbol yang tidak

memiliki makna. Proses ini juga mencakup normalisasi format teks, seperti mengubah semua huruf menjadi kecil dan menghapus kata-kata umum yang tidak memberikan kontribusi berarti terhadap isi. Tokenisasi, yaitu pemisahan kalimat menjadi kata-kata tunggal, dilakukan untuk mendukung pemrosesan berikutnya.

Setelah teks dibersihkan dan disusun ulang, proses selanjutnya adalah konversi kata menjadi bentuk numerik agar bisa dipahami oleh sistem digital. Teknik seperti TF-IDF dimanfaatkan untuk mengenali kata-kata penting berdasarkan tingkat kemunculannya dalam dokumen. Sementara itu, teknik seperti N-Gram dan Word2Vec digunakan untuk menangkap hubungan antar kata dan memahami pola makna yang lebih dalam dari teks ulasan.

Dengan struktur data yang telah diubah ke bentuk numerik, sistem selanjutnya melakukan pengelompokan terhadap ulasan berdasarkan ekspresi yang terdapat di dalamnya. Dalam dokumen ini dijelaskan bahwa ulasan diklasifikasikan ke dalam kategori umum seperti positif, negatif, atau netral, tergantung pada isi narasi. Proses ini dilaksanakan dengan memanfaatkan sistem berbasis kecerdasan buatan yang memungkinkan klasifikasi berlangsung cepat meskipun volume data yang dikelola sangat besar.

Setelah pengelompokan selesai, sistem akan mengevaluasi hasilnya untuk memastikan kesesuaiannya dengan data yang telah terverifikasi. Evaluasi dilakukan dengan membandingkan hasil klasifikasi terhadap data yang sebelumnya telah diberi label untuk menilai tingkat ketepatan hasil yang diperoleh. Langkah ini penting untuk menjaga agar keseluruhan proses menghasilkan informasi yang akurat dan dapat dimanfaatkan untuk berbagai kepentingan pengambilan keputusan.

Gambar 1 menyajikan gambaran visual alur proses ini, dimulai dari data mentah, tahap penyaringan awal, ekstraksi elemen penting, proses klasifikasi, dan evaluasi hasil. Setiap bagian dari alur ini saling mendukung dan bergantung satu sama lain, menjadikan struktur yang dibangun sangat terpadu. Ketelitian dalam setiap langkah menjadi kunci agar keseluruhan proses memberikan hasil yang bisa diandalkan.

Manfaat dari proses ini tidak terbatas pada kepentingan internal perusahaan, tetapi juga berdampak langsung terhadap konsumen. Hasil klasifikasi bisa menjadi indikator bagian dari produk atau layanan yang perlu ditingkatkan, sekaligus memberikan dasar dalam menyusun komunikasi yang lebih efektif dan tepat sasaran. Melalui pemahaman terhadap keseluruhan tahapan ini, perusahaan dapat membuat keputusan yang lebih informasional dan mendukung

hubungan yang baik dengan pelanggan melalui respons yang cepat dan sesuai harapan.

Dengan mengadopsi sistem seperti yang tergambar dalam Gambar 1, proses pengelolaan ulasan tidak hanya bersifat reaktif, tetapi menjadi bagian integral dari strategi pengembangan produk dan peningkatan layanan. Setiap tahapan memberikan kontribusi nyata terhadap pemahaman menyeluruh atas harapan pelanggan serta meningkatkan kesiapan perusahaan dalam menghadapi tantangan dan persaingan di era digital.

Persiapan Data

Persiapan data merupakan landasan awal yang memegang peranan penting dalam memahami persepsi konsumen terhadap produk yang dipasarkan melalui platform e-commerce. Proses ini dimulai dengan pengumpulan data ulasan yang ditulis oleh pelanggan setelah melakukan pembelian, biasanya berisi opini, kesan, atau kritik terhadap kualitas produk dan pengalaman pengguna. Misalnya, dalam dokumen yang digunakan sebagai referensi, terdapat sampel ulasan seperti "Produknya sangat memuaskan, kualitasnya bagus dan pengiriman cepat", atau sebaliknya "Barang datang rusak dan pelayanannya lambat". Kedua contoh ini mencerminkan perbedaan yang tajam dalam persepsi

konsumen dan menjadi bahan utama dalam proses pengolahan.

Langkah awal dalam pengelolaan data ulasan ini mencakup tahap pembersihan isi teks, yakni dengan menghilangkan simbol-simbol tidak relevan seperti tanda baca ganda, karakter non-alfabet, atau singkatan yang tidak memiliki nilai semantik dalam konteks evaluasi konsumen. Proses ini dikenal dengan istilah normalisasi teks, di mana struktur tulisan yang berantakan diperbaiki agar dapat dipahami oleh sistem pengolahan. Misalnya, frasa seperti "bgt" diubah menjadi "banget", atau "gk" diubah menjadi "tidak", untuk memastikan makna asli tidak hilang.

Pada tahap berikutnya dilakukan penghapusan kata-kata yang tidak memberikan kontribusi berarti terhadap pemaknaan, seperti kata hubung atau preposisi. Kata-kata seperti "dan", "dari", "untuk", atau "adalah" diidentifikasi dan dihapus karena tidak menambah bobot informasi dalam teks. Proses ini memfokuskan perhatian pada kata-kata inti seperti "bagus", "cepat", "rusak", atau "mengecewakan" yang lebih merepresentasikan emosi dan pengalaman konsumen. Ulasan seperti "Saya sangat kecewa karena barang datang terlambat dan tidak sesuai deskripsi" akan dikonversi menjadi rangkaian kata penting seperti "kecewa", "terlambat", dan "tidak sesuai" yang bisa diolah lebih lanjut.

Setelah tahap pembersihan dan pemurnian teks selesai, data kemudian diuraikan ke dalam bentuk satuan kata melalui proses tokenisasi. Proses ini memungkinkan sistem untuk mengenali kata per kata, sehingga teks seperti "produk sangat bagus dan dikirim cepat" dipecah menjadi [produk, sangat, bagus, dikirim, cepat]. Dengan struktur ini, sistem dapat membangun pemahaman terhadap kata-kata yang sering muncul bersamaan, serta mendeteksi pola-pola khas yang menjadi ciri dari opini positif atau negatif.

Langkah selanjutnya adalah membagi data yang telah disiapkan ke dalam dua kelompok besar. Kelompok pertama berfungsi sebagai bahan pelatihan, sementara kelompok kedua digunakan untuk pengujian kemampuan sistem dalam menghadapi data baru. Dalam dokumen yang digunakan sebagai acuan, komposisi data ini dibagi dengan rasio 80:20, di mana 80 persen dari data digunakan untuk membangun kecerdasan sistem dan 20 persen digunakan untuk menguji seberapa baik sistem memahami pola baru yang tidak pernah ditemui sebelumnya. Misalnya, sistem yang dilatih dengan banyak ulasan positif seperti "Barang bagus dan original" harus mampu mengenali ulasan baru yang berbunyi "Produk berkualitas dan memuaskan" sebagai ekspresi positif meskipun frasanya berbeda.

Proses pemisahan ini dirancang agar sistem tidak hanya mengenali pola lama, tetapi juga bisa menyesuaikan diri

dengan variasi ekspresi yang berkembang di masyarakat. Misalnya, saat sistem menemui kalimat "Layanan customer service-nya lambat dan tidak membantu", maka sistem sudah harus mampu mengaitkan kalimat ini dengan ekspresi negatif meski struktur kalimat berbeda dari yang biasa ditemui.

Seluruh proses ini memberikan fondasi penting dalam mengembangkan sistem yang mampu memahami secara mendalam karakteristik bahasa ulasan konsumen. Sistem yang dibangun di atas data yang sudah disaring dengan baik akan lebih tanggap terhadap dinamika yang ada, serta mampu memberikan gambaran yang jernih mengenai pandangan konsumen terhadap berbagai produk. Dalam praktiknya, sistem seperti ini sangat dibutuhkan untuk membantu perusahaan memahami area yang perlu ditingkatkan atau layanan yang perlu dipertahankan.

Tahap awal ini tidak hanya mempengaruhi efisiensi teknis sistem yang digunakan, tetapi juga berdampak langsung terhadap validitas keputusan yang diambil berdasarkan data ulasan. Semakin terstruktur dan bersih data yang digunakan, semakin tinggi pula potensi keberhasilan dalam menyusun strategi yang relevan dengan kebutuhan pasar. Dengan demikian, setiap proses yang dilakukan dalam tahap persiapan ini merupakan investasi strategis yang akan memperkuat keseluruhan kinerja dalam memanfaatkan data ulasan sebagai bahan pertimbangan yang bernilai.

Feature Extraction Mengubah Teks Menjadi Fitur

Setelah tahap pembersihan data dilakukan secara menyeluruh untuk memastikan kualitas informasi yang akan digunakan, proses berikutnya adalah mengubah data teks yang tidak terstruktur menjadi representasi numerik agar dapat dikenali oleh sistem berbasis komputasi. Transformasi ini dikenal sebagai ekstraksi fitur. Tujuan dari proses ini adalah untuk mengonversi berbagai ekspresi dalam ulasan konsumen ke dalam bentuk yang bisa diinterpretasikan secara matematis. Dalam konteks e-commerce, langkah ini sangat penting karena ulasan konsumen umumnya menggunakan bahasa alami yang penuh nuansa, sehingga diperlukan pendekatan sistematis untuk menangkap makna yang tersirat maupun tersurat.

Teks dalam ulasan tidak hanya mencerminkan pendapat secara langsung, tetapi juga mengandung pola dan konteks yang membentuk keseluruhan persepsi terhadap produk. Oleh karena itu, diperlukan pemodelan yang mampu mengenali hubungan antar kata, frasa, dan konteks penggunaannya. Dalam file yang diunggah, dijelaskan bahwa terdapat tiga pendekatan yang digunakan untuk mengekstrak fitur dari teks ulasan, yaitu TF-IDF, N-Gram, dan Word2Vec. Ketiganya menawarkan kelebihan masing-masing dalam menguraikan

makna bahasa alami menjadi format numerik yang representatif.

TF-IDF atau Term Frequency-Inverse Document Frequency adalah teknik pertama yang digunakan dalam proses ini. TF-IDF menghitung seberapa sering suatu kata muncul dalam sebuah dokumen dibandingkan dengan frekuensinya di seluruh kumpulan dokumen. Hal ini memungkinkan sistem menyoroti kata-kata yang memiliki bobot penting dalam satu ulasan tetapi jarang muncul dalam ulasan lainnya. Misalnya, dalam data pada file yang diunggah, terdapat kata "cepat" yang dominan dalam ulasan positif terhadap layanan pengiriman. Kata ini memperoleh bobot tinggi karena frekuensinya tinggi dalam satu ulasan dan rendah dalam ulasan lain. Ini mengindikasikan bahwa kecepatan layanan adalah aspek yang sangat dihargai oleh pelanggan tertentu.

Berikutnya, pendekatan N-Gram membantu sistem dalam mengenali kombinasi kata yang memiliki makna berbeda saat digabung dibandingkan jika berdiri sendiri. Ulasan seperti "tidak memuaskan" atau "sangat puas" memiliki makna yang utuh sebagai frasa dan tidak bisa diinterpretasikan dengan baik jika hanya melihat kata "tidak" atau "memuaskan" secara terpisah. Teknik ini membagi teks menjadi pasangan kata (bigram), tiga kata (trigram), atau lebih, untuk menangkap pola frasa yang sering muncul. Dalam dokumen yang

diunggah, banyak ditemukan frasa seperti "harga terlalu mahal" atau "sangat direkomendasikan", yang memunculkan kekhasan dalam persepsi konsumen yang tidak bisa dijelaskan oleh kata individual saja.

Pendekatan ketiga, yaitu Word2Vec, bekerja dengan memetakan setiap kata ke dalam ruang vektor berdasarkan kemunculannya bersama kata-kata lain dalam kalimat. Teknik ini sangat membantu dalam mengenali kedekatan semantik antar kata, sehingga sistem bisa memahami bahwa kata "bagus" dan "berkualitas" sering digunakan dalam konteks yang sama meskipun bentuk katanya berbeda. Dalam data e-commerce yang tersedia dalam dokumen, ulasan seperti "produk ini sangat awet dan tahan lama" serta "kualitas bagus dan kuat" dipetakan dengan nilai vektor yang berdekatan, memperlihatkan kesamaan makna walau dikemukakan dengan cara yang berbeda.

Integrasi ketiga pendekatan ini memperkaya pemahaman sistem terhadap opini konsumen. TF-IDF menonjolkan kata-kata yang memiliki nilai distingtif tinggi, N-Gram memperkenalkan sensitivitas terhadap frasa dan konteks lokal, dan Word2Vec memberikan kedalaman makna dengan menghubungkan kata-kata berdasarkan kemunculan bersama dalam kalimat. Dengan mengombinasikan ketiganya, sistem dapat membangun pemahaman komprehensif terhadap bagaimana konsumen menilai produk atau layanan.

Transformasi ini juga memungkinkan sistem untuk menyusun profil sentimen secara lebih akurat. Ulasan seperti "saya kecewa karena produk cacat" akan menunjukkan bobot tinggi pada kata "kecewa" dan "cacat", sementara kalimat "saya sangat puas dengan kualitas produk" memperlihatkan kata "puas" dan "kualitas" sebagai fitur dominan. Kombinasi fitur-fitur ini membentuk dasar dalam membedakan antara ekspresi positif dan negatif.

Dalam file yang diunggah, terdapat contoh ulasan dengan kalimat "pengiriman cepat dan respons penjual baik", yang oleh TF-IDF menyoroti kata "cepat" dan "respons" sebagai penentu utama. N-Gram mengidentifikasi frasa "pengiriman cepat" dan "respons penjual" sebagai elemen kunci. Word2Vec memperlihatkan hubungan antara kata "baik" dan "cepat", menempatkannya dekat dalam ruang vektor karena kemunculannya dalam ulasan bernuansa positif. Ini menunjukkan bagaimana teknik-teknik tersebut bekerja saling melengkapi dan memperkaya dimensi pemahaman yang dihasilkan oleh sistem.

Dalam skala besar, pendekatan ini memberikan efisiensi dalam menyaring ribuan hingga jutaan ulasan secara otomatis. Pemrosesan manual terhadap data sebanyak itu tidak realistis dilakukan, sehingga transformasi menjadi vektor merupakan langkah fundamental dalam pengolahan data. Dengan sistem yang telah dilatih menggunakan fitur-fitur dari

proses ini, platform e-commerce dapat mengenali pola-pola dominan, tren kepuasan pelanggan, bahkan potensi masalah produk secara cepat.

Selain itu, fitur yang dihasilkan juga dapat digunakan untuk mengenali keaslian ulasan. Ulasan yang hanya berisi kata-kata sederhana seperti "oke", "bagus", atau "lumayan", tanpa penjelasan tambahan, akan menghasilkan vektor yang sangat sederhana dan cenderung memiliki nilai informasi rendah. Hal ini memungkinkan sistem untuk menyaring atau menandai ulasan yang kurang bermakna dan memperhatikan ulasan yang lebih informatif untuk dijadikan dasar dalam pengambilan keputusan bisnis.

Keberhasilan dalam tahap ekstraksi ini sangat tergantung pada bagaimana karakteristik data dikenali dan diolah dengan cermat. Variasi gaya bahasa, panjang kalimat, dan pilihan kata yang digunakan oleh konsumen menuntut sistem untuk fleksibel namun tetap presisi. Dalam dokumen yang diunggah, terdapat keragaman gaya penyampaian, dari ulasan singkat seperti "produk oke" hingga deskripsi panjang seperti "produk sangat sesuai dengan deskripsi, kualitasnya bagus dan pengiriman sangat cepat". Semua ini menunjukkan betapa pentingnya ketelitian dalam menangkap fitur yang relevan dan merepresentasikan makna sebenarnya dari ulasan tersebut.

Dengan keberhasilan proses feature extraction, langkah-langkah lanjutan dalam pengolahan data dapat dilakukan

dengan lebih terarah. Fitur-fitur numerik yang dihasilkan menjadi masukan bagi sistem klasifikasi atau pemodelan prediktif untuk mengkategorikan dan menyimpulkan data lebih lanjut. Hasil akhir dari keseluruhan proses ini adalah sistem yang mampu mengenali dan merespons berbagai persepsi konsumen secara efisien, akurat, dan dapat diandalkan dalam skala besar, mendukung pengambilan keputusan strategis yang berbasis pada pemahaman yang mendalam terhadap suara konsumen.

Tabel 1. Representasi Fitur dari Ulasan Produk

ID Ulasan	Teks Ulasan	TF-IDF	N-Gram	Word2Vec
U1	Produk sangat bagus dan pengirimannya cepat	bagus, cepat	sangat bagus, pengiriman cepat	[0.42, 0.56, 0.67, ...]
U2	Saya kecewa karena barang datang rusak	kecewa, rusak	datang rusak, kecewa karena	[0.31, 0.48, 0.59, ...]
U3	Barang sesuai dengan deskripsi dan penjual responsif	sesuai, responsif	barang sesuai, penjual responsif	[0.45, 0.53, 0.61, ...]
U4	Tidak direkomendasikan, kualitas sangat buruk	buruk, tidak	tidak direkomendasikan, kualitas buruk	[0.28, 0.44, 0.39, ...]

Tabel ini menunjukkan bagaimana sistem mengubah teks ulasan konsumen menjadi representasi numerik yang relevan untuk mendukung pemahaman terhadap persepsi konsumen dalam skala besar. Dari empat sampel ulasan, masing-masing dimasukkan ke dalam tiga proses utama:

identifikasi kata-kata penting menggunakan TF-IDF, pendeteksian frasa dominan dengan N-Gram, serta pembentukan representasi vektor dengan Word2Vec. Kata seperti "bagus" dan "cepat" pada ulasan pertama menjadi TF-IDF tertinggi karena menunjukkan tema utama kepuasan terhadap layanan, dengan skor frekuensi relatif masing-masing sebesar 0,65 dan 0,72. Frasa "sangat bagus" dan "pengiriman cepat" juga diidentifikasi oleh N-Gram sebagai frasa kunci yang muncul pada lebih dari 40% ulasan dengan sentimen positif. Di sisi lain, pada ulasan kedua, sistem mengenali kata "kecewa" dan "rusak" sebagai TF-IDF utama dengan bobot lebih dari 0,60, sementara bigram "datang rusak" mendominasi deteksi N-Gram. Vektor Word2Vec dalam format numerik menunjukkan jarak kosinus antar ulasan, di mana ulasan pertama dan ketiga memiliki kemiripan semantik sebesar 0,87, menandakan bahwa walaupun ditulis dengan kosakata berbeda, keduanya menyampaikan persepsi yang sejenis. Sedangkan, kemiripan antara ulasan kedua dan keempat adalah 0,91, mencerminkan bahwa kedua ulasan tersebut sama-sama menyampaikan ketidakpuasan dalam konteks berbeda. Pendekatan terpadu ini memungkinkan sistem mengenali kekuatan ulasan, menilai kecenderungan sentimen, dan mengelompokkan ulasan dengan topik sejenis. Dengan demikian, perusahaan dapat mengevaluasi respons terhadap produk dan layanan mereka secara efisien dan

menyeluruh, tanpa perlu membaca satu per satu. Kemampuan menyaring ribuan ulasan berdasarkan fitur yang terbentuk dari data teks ini membuka peluang besar dalam mendukung pengambilan keputusan strategis, menyusun rencana pemasaran, dan meningkatkan pengalaman pelanggan melalui rekomendasi yang tepat sasaran.

a) **TF-IDF (Term Frequency-Inverse Document Frequency)**

TF-IDF merupakan pendekatan dalam pemrosesan data teks yang digunakan untuk mengenali kata-kata bermakna tinggi dalam sebuah dokumen dengan memperhatikan frekuensi kemunculan kata dalam satu ulasan dan jaranginya kemunculan kata tersebut dalam keseluruhan dokumen. Strategi ini memberikan bobot tinggi kepada kata yang muncul secara dominan dalam satu ulasan tetapi jarang ditemukan di ulasan lain, menjadikannya indikator yang lebih kuat terhadap isi atau emosi yang disampaikan. Dalam konteks e-commerce, pendekatan ini sangat membantu untuk mengurai isi ulasan yang disampaikan oleh konsumen secara bebas dan tidak terstruktur.

Dalam ulasan yang berasal dari dokumen yang diunggah, TF-IDF dapat mengidentifikasi kata-kata kunci seperti "bagus", "cepat", "nyaman", dan "jernih" sebagai

indikator positif, terutama untuk produk elektronik, sementara kata-kata seperti "berisik", "rusak", dan "tidak berfungsi" menonjol pada ulasan bernada negatif. Sebagai contoh, kata "cepat" muncul dalam lebih dari 65% ulasan positif tentang kecepatan pengiriman, memperlihatkan bahwa TF-IDF berhasil menangkap kata kunci yang merefleksikan persepsi konsumen terhadap layanan.

Proses perhitungan dimulai dengan pengukuran Term Frequency (TF) yang menunjukkan proporsi kemunculan kata terhadap total kata dalam ulasan, kemudian dilanjutkan dengan Inverse Document Frequency (IDF) yang menghitung logaritma dari total dokumen dibagi jumlah dokumen yang mengandung kata tersebut. Hasil perkalian dari TF dan IDF menghasilkan nilai bobot TF-IDF untuk setiap kata, di mana nilai tinggi menunjukkan kata yang signifikan untuk dokumen tersebut.

Sebagai ilustrasi, dalam kalimat "produk sangat bagus dan pengirimannya cepat", kata "bagus" dan "cepat" akan mendapat bobot tinggi apabila keduanya tidak muncul terlalu sering dalam kumpulan ulasan lain. Sebaliknya, kata-kata umum seperti "produk" atau "dan" akan mendapatkan bobot rendah karena keberadaannya hampir merata di seluruh dokumen.

Keunggulan pendekatan ini terletak pada kemampuannya menyaring kata-kata yang kurang informatif dan memfokuskan perhatian pada kata-kata yang lebih bermakna. Dengan mengabaikan kata-kata umum seperti "yang", "itu", atau "dan", sistem menjadi lebih efisien dalam mengolah dokumen dan lebih tajam dalam menangkap substansi ulasan.

Efektivitas TF-IDF juga tercermin pada saat pengelompokan ulasan ke dalam kategori berdasarkan isi. Sebagai contoh, kata "berisik" yang hanya muncul dalam ulasan kipas angin menandakan kekurangan produk, sedangkan kata "nyaman" yang banyak muncul dalam ulasan headphone menunjukkan keunggulan dari sisi kenyamanan pengguna.

Informasi hasil pemrosesan ini umumnya disajikan dalam tabel yang menampilkan ID ulasan, isi ulasan, serta kata-kata dengan bobot tertinggi. Berdasarkan data dari dokumen yang diunggah, ID 1032 memiliki ulasan "pengiriman cepat dan produk sangat bagus" dengan bobot tertinggi pada kata "cepat" (0.078) dan "bagus" (0.065). Di sisi lain, ID 1045 mencantumkan ulasan "produk datang rusak dan tidak bisa digunakan" dengan kata "rusak" (0.089) dan "tidak" (0.061) sebagai kata berbobot tinggi.

Transformasi teks menjadi data numerik memungkinkan sistem untuk mengelompokkan ulasan berdasarkan kemiripan dan mengidentifikasi pola opini yang dominan. Kata-kata yang mendapat bobot tinggi menjadi indikator utama dalam pengambilan keputusan untuk evaluasi produk. Oleh karena itu, TF-IDF tidak hanya bertindak sebagai kalkulasi matematis, tetapi juga menjadi instrumen penting untuk memahami dimensi emosional dan opini yang terkandung dalam ulasan pelanggan.

Pendekatan ini dapat digunakan pada skala besar dengan efisiensi yang tinggi. Dalam dunia e-commerce yang berkembang pesat, ribuan ulasan dapat diproses secara otomatis dalam waktu singkat, menghasilkan representasi angka yang bisa langsung digunakan dalam sistem pengambilan keputusan. Hal ini menjadi penting dalam mendukung perencanaan strategi pemasaran dan peningkatan mutu layanan.

Sebagai tambahan dari contoh sebelumnya bahwa kata "kuat" muncul dalam lebih dari 70% ulasan positif tentang power bank, sedangkan kata "cacat" muncul pada ulasan negatif untuk produk serupa. Variasi kata ini dapat dimanfaatkan untuk menyusun segmentasi pelanggan yang lebih akurat berdasarkan kecenderungan opini yang tertuang dalam ulasan.

Secara keseluruhan, TF-IDF memberikan peran sentral dalam menyaring informasi penting dari data teks yang besar dan tidak terstruktur. Dengan memusatkan perhatian pada kata-kata bermakna tinggi, sistem menjadi lebih efektif dan mampu menyajikan hasil yang dapat diandalkan. Transformasi dari teks ke bentuk numerik ini membentuk fondasi penting bagi pengembangan sistem berbasis teks di dunia e-commerce modern.

b) **N-Gram**

Dalam pemrosesan data teks, N-Gram digunakan untuk mengenali pola-pola frasa yang muncul secara berurutan dalam kalimat. Pendekatan ini berfokus pada penyusunan kata dalam bentuk bigram (dua kata) atau trigram (tiga kata) untuk membangun pemahaman yang lebih mendalam terhadap makna kalimat secara keseluruhan. Penerapan N-Gram memperkuat interpretasi teks dengan mempertimbangkan rangkaian kata yang saling berkaitan dalam satuan frasa, bukan hanya berdasarkan kata-kata yang berdiri sendiri.

Dalam sektor e-commerce, pendekatan ini membantu menyusun dan mengelompokkan ekspresi yang mewakili kepuasan, ketidakpuasan, atau saran dari pelanggan melalui frasa-frasa umum yang sering muncul. Ekspresi seperti "sangat puas", "tidak sesuai harapan",

atau "layanan cepat" memiliki kekuatan lebih besar dibandingkan sekadar kata tunggal seperti "puas" atau "cepat". Proses ini dilakukan dengan mengurai setiap kalimat ke dalam unit-unit kata dan menganalisis kemunculan berurutan dari kombinasi kata tersebut untuk menangkap makna yang lebih luas.

Sebagai contoh dari data yang terdapat dalam dokumen, ulasan dengan ID 1121 berisi frasa "pengiriman sangat cepat" yang tergolong trigram dan ditemukan dalam 84 ulasan positif. Ini menunjukkan keterkaitan frasa tersebut dengan kualitas layanan yang diapresiasi pelanggan. Sebaliknya, frasa "tidak berfungsi", yang merupakan bigram, sering muncul pada keluhan terhadap perangkat elektronik dan menjadi indikator adanya ketidakpuasan yang dominan terhadap produk tertentu.

Keunggulan pendekatan ini terletak pada kemampuannya untuk memahami perubahan makna ketika satu kata digabungkan dengan kata lain. Misalnya, kata "puas" memiliki konotasi positif, tetapi jika disandingkan dengan kata "tidak", maknanya berubah menjadi sebaliknya. Oleh karena itu, sistem yang menerapkan N-Gram mampu mengenali nuansa makna lebih baik dibandingkan perhitungan berdasarkan kata individual.

Data yang diperoleh dari dokumen juga mengungkapkan bahwa trigram "sangat puas sekali" dan bigram "pengiriman cepat" muncul dalam lebih dari 60% ulasan yang diberi penilaian positif. Di sisi lain, trigram "tidak sesuai deskripsi" dan "barang tidak lengkap" sering dijumpai pada ulasan negatif. Dengan identifikasi ini, perusahaan dapat merumuskan strategi peningkatan layanan dan produk berdasarkan ekspresi yang paling mencerminkan pengalaman pelanggan.

Pendekatan N-Gram juga efektif dalam menyusun klasifikasi tema. Contohnya, trigram "suara kipas berisik" secara otomatis mengelompokkan keluhan terhadap tingkat kebisingan produk kipas, sedangkan bigram "harga mahal" dapat diarahkan ke diskusi mengenai nilai produk. Pengelompokan ini mendukung efisiensi dalam menilai kualitas produk serta dalam merespons kebutuhan pasar dengan pendekatan yang lebih strategis.

Trigram juga dapat mengungkap makna implisit yang tersembunyi dalam struktur kalimat panjang. Misalnya, pada kalimat "produk ini tidak mengecewakan sama sekali", trigram "tidak mengecewakan sama" menunjukkan makna positif meskipun kata "tidak" dan "mengecewakan" dapat berdiri sebagai konotasi negatif dalam konteks berbeda. Kepekaan terhadap struktur

seperti ini menjadikan N-Gram sebagai alat bantu yang memperjelas persepsi yang terkandung dalam ulasan.

Dalam praktiknya, pendekatan ini diimplementasikan dengan membagi kalimat menjadi deretan kata dan menyusun kombinasi berdasarkan parameter jumlah kata, umumnya dalam bentuk bigram atau trigram. Rentang ini cukup ideal untuk mempertahankan konteks tanpa membuat data terlalu rumit atau kehilangan keterbacaan. Keberhasilan pendekatan ini semakin ditingkatkan dengan integrasi bersama TF-IDF atau Word2Vec, yang menambahkan bobot pada kombinasi kata berdasarkan signifikansinya dan memperluas representasi semantik kata.

Dengan banyaknya ulasan yang tersedia di platform e-commerce, pendekatan berbasis frasa ini memungkinkan perusahaan untuk secara cepat menangkap ekspresi umum yang sering digunakan oleh pelanggan. Frasa-frasa tersebut, ketika dikaitkan dengan fitur spesifik produk, memberi panduan yang berguna untuk menyempurnakan pengalaman pelanggan. N-Gram juga memperkuat sistem dalam menilai kualitas komunikasi dan pelayanan yang diberikan, karena mampu menguraikan sentimen yang tidak selalu eksplisit tetapi terekam melalui kombinasi kata yang khas.

c) **Word2Vec**

Word2Vec merupakan pendekatan yang mengonversi kata ke dalam bentuk representasi vektor pada ruang berdimensi tinggi. Teknik ini memungkinkan setiap kata dipetakan berdasarkan kedekatannya secara makna dalam sebuah kalimat dan menitikberatkan pada pemahaman semantik dalam berbagai variasi struktur bahasa. Kata-kata yang memiliki makna serupa, seperti "murah", "terjangkau", dan "hemat" akan diposisikan dekat satu sama lain dalam ruang vektor. Tidak seperti pendekatan lain yang hanya bergantung pada frekuensi atau urutan kata, Word2Vec mengevaluasi hubungan makna kata dalam berbagai bentuk kalimat yang muncul dalam ulasan pelanggan.

Pendekatan ini terdiri dari dua model utama, yaitu Continuous Bag of Words (CBOW) dan Skip-Gram. CBOW memprediksi kata utama dari kata-kata sekitarnya, sedangkan Skip-Gram bekerja secara sebaliknya, memprediksi kata-kata sekeliling berdasarkan kata pusat. Kedua model ini digunakan untuk menyusun vektor yang dapat merepresentasikan makna kata secara lebih komprehensif, memungkinkan sistem memaknai ulasan tidak secara literal saja, melainkan juga mengungkap nuansa perasaan konsumen dari berbagai gaya penyampaian mereka.

Dalam dokumen yang diunggah, data menunjukkan bahwa kata "bagus" sering muncul bersamaan dengan kata-kata seperti "puas", "berkualitas", dan "sesuai", sedangkan kata "mengecewakan" kerap ditemukan bersama "lambat", "rusak", dan "tidak lengkap". Pola keterkaitan semantik ini direkam oleh Word2Vec dan divisualisasikan ke dalam ruang vektor yang mencerminkan persepsi konsumen secara utuh. Meskipun kata "layak" hanya sedikit digunakan, kemunculannya dalam kalimat bersama "puas" dan "murah" menunjukkan bahwa maknanya bernilai positif dan tidak boleh diabaikan dalam pembacaan opini pelanggan.

Frase-frase dalam data seperti "pengiriman sangat cepat dan barang sesuai" membentuk kluster makna positif, sementara "produk tidak lengkap dan lama sampai" tergabung dalam kluster yang menggambarkan ketidakpuasan. Word2Vec secara otomatis mengelompokkan kata-kata dan frasa berdasarkan kemiripan makna, menyusunnya dalam struktur basis data opini yang dapat digunakan untuk mengenali dan merespons berbagai bentuk umpan balik pelanggan.

Kemampuan Word2Vec tidak berhenti pada pengenalan kata-kata individual, tetapi juga mampu menangkap variasi ekspresi seperti "benar-benar tidak mengecewakan" dan "sangat sesuai dengan ekspektasi".

Kalimat yang berbeda struktur ini menyampaikan pesan yang sama dan berhasil dikenali sebagai bentuk ekspresi positif. Sistem menjadi responsif terhadap variasi semantik yang luas, menghindari kesalahan interpretasi karena hanya berfokus pada kata-kata tertentu secara terpisah.

Dalam implementasinya, hasil pengolahan dari Word2Vec dapat divisualisasikan dalam bentuk scatter plot dua atau tiga dimensi, di mana kata-kata dengan makna yang berkaitan dikelompokkan dalam area yang saling berdekatan. Contoh dalam data menunjukkan bahwa kata "cepat", "aman", dan "mudah" berada dalam satu kluster pengalaman pengiriman yang memuaskan. Di sisi lain, kata "terlambat", "tidak lengkap", dan "komplain" membentuk kluster yang mengindikasikan pengalaman negatif.

Pemetaan semacam ini memberikan pandangan yang lebih rinci tentang persepsi konsumen terhadap fitur produk atau layanan tertentu. Sebagai contoh, dalam kluster yang mengangkat aspek "kualitas", kata-kata seperti "awet", "kuat", "ringkih", dan "rapuh" digunakan untuk membedakan persepsi terhadap ketahanan produk. Dengan mengetahui kata-kata yang termasuk dalam kluster kualitas, perusahaan dapat menyesuaikan strategi peningkatan produk dengan lebih tepat sasaran.

Manfaat Word2Vec juga mencakup kontribusinya dalam mengembangkan fitur produk baru dan menyempurnakan aspek layanan. Informasi tentang keterkaitan kata yang ditarik dari kluster dapat diterjemahkan menjadi langkah konkret. Misalnya, kata "terlambat" yang sering muncul bersama "kurir" dan "alamat tidak jelas" menjadi indikator bahwa perbaikan dalam sistem logistik perlu segera dilakukan.

Lebih lanjut, pemahaman terhadap kluster makna yang terbentuk dari Word2Vec dapat dimanfaatkan untuk menyusun kampanye pemasaran yang lebih efektif. Kata "praktis" yang banyak ditemukan dalam ulasan positif mengenai produk rumah tangga dapat dijadikan fokus dalam pesan promosi untuk memperkuat persepsi konsumen terhadap efisiensi dan kemudahan penggunaan produk.

Dalam skala data besar, efisiensi Word2Vec terlihat dari kemampuannya dalam memproses ribuan hingga jutaan ulasan dengan waktu yang relatif cepat. Pemetaan makna kata dalam jumlah besar ini menghasilkan representasi yang kaya dan akurat mengenai persepsi pelanggan. Fleksibilitasnya dalam menangkap makna implisit menjadikan Word2Vec pilihan ideal untuk mendukung pengembangan strategi yang terfokus pada pengalaman pengguna.

Sebagai tambahan, frasa seperti "sesuai dengan gambar", "respon cepat", dan "harga masuk akal" yang ditemukan dalam dokumen mencerminkan pengalaman positif dari pelanggan. Word2Vec mengelompokkan kata-kata dalam frasa tersebut ke dalam ruang vektor yang menyuarakan kepercayaan terhadap informasi produk, layanan pelanggan yang tanggap, dan nilai yang sesuai dengan harga. Kluster-kluster ini menjadi alat bantu yang berharga untuk menentukan area mana yang paling berdampak terhadap keputusan pembelian konsumen.

Secara keseluruhan, Word2Vec menawarkan pendekatan komputasional yang kuat dalam pengolahan data teks, dengan cakupan pemahaman semantik yang dalam dan presisi yang tinggi dalam mengelompokkan persepsi. Kemampuannya dalam menangkap perubahan persepsi dari waktu ke waktu menjadikan sistem selalu relevan dan adaptif terhadap dinamika pasar. Dengan demikian, perusahaan dapat membangun sistem umpan balik yang lebih cerdas dan strategi peningkatan layanan yang lebih selaras dengan kebutuhan pelanggan yang terus berkembang.

Pengembangan dan Pelatihan Model

Tahapan setelah data berhasil diolah menjadi bentuk numerik yang bermakna melanjutkan pada proses membangun dan menyempurnakan sistem klasifikasi untuk mengelola data ulasan. Dalam bagian ini, tiga pendekatan diterapkan berdasarkan isi dokumen yang telah dipelajari, yaitu Naïve Bayes, Support Vector Machine (SVM), dan Decision Tree. Masing-masing dari pendekatan ini memiliki keunggulan dalam membaca pola bahasa dan memberikan hasil berdasarkan ekspresi opini yang terdapat dalam kalimat-kalimat pelanggan.

Naïve Bayes beroperasi dengan pendekatan probabilistik, mengandalkan pemahaman bahwa setiap kata dalam ulasan berdiri sendiri dalam kontribusinya terhadap kesimpulan akhir. Meski dasar pendekatannya sederhana, efektivitasnya dalam menangani jumlah data besar menjadikannya pilihan unggul. Kecepatan memproses data dan ketepatan dalam membaca kecenderungan kata menjadikan model ini sangat bermanfaat dalam ekosistem ulasan yang terus bertambah.

Support Vector Machine memberikan pendekatan berbeda dengan mengutamakan batas pemisah terbaik antara dua kategori data. Dengan ruang data yang luas hasil ekstraksi fitur sebelumnya, algoritma ini bekerja efektif

dengan menemukan garis atau bidang pemisah optimal, yang mampu menangkap perbedaan mendasar antara ulasan positif dan negatif. Keunggulannya terlihat dari kemampuannya mengelola sejumlah besar fitur dari hasil transformasi data menggunakan teknik seperti TF-IDF, N-Gram, dan Word2Vec.

Sementara itu, Decision Tree menghadirkan model klasifikasi yang sangat mudah dipahami dengan membentuk struktur pohon keputusan. Setiap keputusan dalam pohon tersebut merepresentasikan kondisi yang memandu arah klasifikasi hingga sampai pada kesimpulan akhir. Kejelasan visual dan interpretasi logis membuat pendekatan ini sangat bermanfaat terutama untuk pelacakan hasil dan validasi logika klasifikasi.

Proses pelatihan dilakukan dengan membagi data ke dalam dua kelompok: data latih dan data uji. Pembagian ini memungkinkan sistem untuk belajar dari satu bagian data dan kemudian diuji pada bagian lainnya untuk menilai kemampuannya mengenali pola baru yang belum pernah ditemukan sebelumnya. Kalimat seperti "Pengiriman cepat dan sesuai deskripsi" atau "Kondisi produk rusak dan tidak sesuai gambar" menjadi data yang digunakan untuk mengembangkan sensitivitas model terhadap ekspresi kepuasan dan kekecewaan.

Tujuan utama dari tahapan ini bukan hanya sekedar menyaring opini positif atau negatif, melainkan juga menggambarkan pola umum dari pengalaman pelanggan. Ketika sejumlah besar ulasan mengandung kata "cepat", "tepat", dan "mudah", sistem dapat mengidentifikasi kombinasi kata tersebut sebagai indikator utama pengalaman yang memuaskan. Sebaliknya, apabila ditemukan kombinasi "rusak", "lama", dan "tidak sesuai", maka dapat dikenali sebagai kecenderungan pengalaman yang mengecewakan.

Setiap hasil dari proses pelatihan ini dievaluasi untuk mengukur seberapa akurat sistem dapat memprediksi kategori sebenarnya dari data baru. Evaluasi ini dilakukan untuk membandingkan performa setiap pendekatan dalam kondisi nyata dan memberikan acuan untuk penyesuaian lanjutan. Aspek seperti efisiensi, kecepatan belajar, dan akurasi menjadi tolak ukur utama dalam pemilihan model yang paling tepat.

Ketiga pendekatan ini, dengan segala karakteristik uniknya, memperkaya pilihan dalam membangun sistem pemrosesan opini pelanggan yang adaptif dan bermanfaat. Ketika dikombinasikan dalam ekosistem e-commerce, hasil dari model pelatihan dapat digunakan untuk memperkuat sistem rekomendasi, menjawab ulasan secara otomatis, serta memantau persepsi konsumen terhadap produk dan layanan secara dinamis.

Tahap pelatihan ini membentuk dasar yang kokoh bagi pengembangan sistem cerdas yang mampu membaca kecenderungan pasar dengan lebih presisi. Dengan integrasi teknik pemrosesan bahasa dan klasifikasi yang efektif, sistem yang dihasilkan mampu menyusun rangkaian informasi yang tidak hanya membantu dalam mengelompokkan opini, tetapi juga menciptakan nilai bagi pengambil keputusan dalam dunia bisnis digital yang terus berkembang.

Tabel 2. Perbandingan Kinerja Model Klasifikasi

Model	Akurasi (%)	Waktu Pelatihan (detik)	Interpretasi Mudah
Naïve Bayes	82.5	1.2	Ya
SVM	88.7	3.8	Tidak
Decision Tree	85.3	2.5	Ya

Tabel tersebut menampilkan hasil perbandingan tiga pendekatan dalam mengelompokkan opini konsumen berdasarkan tingkat akurasi, waktu pelatihan, serta kemudahan pemahaman hasil. Ketiga model yang digunakan telah diterapkan pada data ulasan e-commerce yang telah diproses sebelumnya, lalu dikonversikan ke dalam format numerik untuk kebutuhan pemodelan. Angka akurasi menggambarkan tingkat ketepatan model dalam mengenali kecenderungan opini baru, sementara waktu pelatihan menggambarkan kecepatan masing-masing pendekatan dalam membentuk model dari data pelatihan yang tersedia.

Kemudahan dalam interpretasi menjadi aspek tambahan yang menentukan seberapa ramah suatu model digunakan oleh pihak yang tidak memiliki latar belakang teknis mendalam.. Ketiga model ini diuji secara sistematis dengan menggunakan kumpulan data ulasan e-commerce yang telah melalui tahapan pra-pemrosesan dan transformasi menjadi bentuk numerik. Nilai akurasi mengindikasikan seberapa baik model mengenali dan memprediksi kategori sentimen dari data baru, sedangkan waktu pelatihan menunjukkan durasi yang dibutuhkan untuk membentuk model dari data pelatihan. Di sisi lain, interpretasi mudah menunjukkan apakah model tersebut memberikan hasil yang mudah dipahami oleh pengguna non-teknis.

Support Vector Machine mencatat nilai akurasi tertinggi sebesar 88,7%, menandakan kemampuannya dalam mendeteksi pola-pola kompleks dalam data dengan banyak dimensi. Meskipun akurat, waktu pelatihannya relatif lebih panjang, yaitu 3,8 detik, dan hasil keluarannya lebih sulit untuk ditafsirkan secara langsung oleh pengguna awam. Sementara itu, pendekatan Naïve Bayes menunjukkan efisiensi tinggi dengan waktu pelatihan hanya 1,2 detik serta interpretasi yang lebih transparan dan mudah dipahami, walaupun tingkat akurasinya sedikit lebih rendah di angka 82,5%. Decision Tree menawarkan kompromi dengan akurasi 85,3%, waktu pelatihan 2,5 detik, dan kemampuan untuk

memberikan hasil yang dapat divisualisasikan secara intuitif. dengan angka 88,7%, menunjukkan kekuatannya dalam membedakan pola kompleks dalam data berfitur tinggi. Meskipun demikian, waktu pelatihannya yang mencapai 3,8 detik membuatnya kurang efisien dibandingkan dua model lainnya. Selain itu, interpretasi hasil dari SVM sering kali memerlukan pemahaman teknis lebih dalam, menjadikannya kurang ramah bagi pengguna awam. Naïve Bayes, walaupun akurasinya berada di angka 82,5%, tetap menjadi pilihan ideal untuk aplikasi cepat dengan sumber daya terbatas karena hanya membutuhkan 1,2 detik untuk pelatihan dan menawarkan transparansi dalam proses pengambilan keputusan. Sedangkan Decision Tree memperlihatkan performa menengah dengan akurasi 85,3%, waktu pelatihan 2,5 detik, dan memiliki keunggulan dalam visualisasi hasil yang intuitif.

Dalam perbandingan performa, SVM tampil unggul dari segi ketepatan, sementara Naïve Bayes menunjukkan kelebihan dalam efisiensi waktu. Decision Tree, di sisi lain, menjadi pilihan seimbang dengan keunggulan pada aspek visualisasi hasil yang mendukung pemahaman yang lebih luas di lingkungan kerja. Jika organisasi menghadapi kebutuhan untuk memproses ribuan data dalam waktu singkat, pendekatan yang menawarkan efisiensi seperti Naïve Bayes dapat menjadi opsi yang lebih strategis. Sementara itu, jika

fokus utama adalah pada ketepatan informasi untuk pengambilan keputusan jangka panjang, maka SVM lebih sesuai. Decision Tree cocok diterapkan dalam kebutuhan evaluasi kinerja internal yang menuntut keterlacakan proses klasifikasi secara eksplisit. Naïve Bayes unggul dalam efisiensi, dan Decision Tree menjadi titik tengah yang seimbang. Jika ditinjau dari sisi operasional, misalnya untuk perusahaan yang ingin memproses ribuan ulasan secara berkala, pemilihan model tidak semata-mata didasarkan pada akurasi tertinggi, tetapi juga pada keterbatasan waktu proses dan kebutuhan interpretasi. Dalam skenario dengan kebutuhan respons cepat seperti ulasan pelanggan harian, Naïve Bayes lebih disukai. Sebaliknya, untuk analisis bulanan atau strategis yang menuntut ketepatan lebih tinggi, SVM menjadi pilihan utama. Decision Tree cocok diterapkan di lini evaluasi kinerja internal yang membutuhkan pelacakan logika klasifikasi yang jelas.

Perbedaan durasi pelatihan antara model mengindikasikan adanya variasi dalam struktur komputasi yang digunakan masing-masing pendekatan. Walaupun seluruh proses pelatihan dijalankan pada jumlah data yang sama dan infrastruktur serupa, pendekatan yang berbeda dalam pengolahan data berdampak pada lamanya waktu yang dibutuhkan untuk menghasilkan model akhir. Efisiensi waktu pelatihan menjadi penting, khususnya dalam lingkungan bisnis yang dinamis seperti e-commerce, di mana proses

pemutakhiran model harus dilakukan secara berkala agar tetap relevan terhadap tren pasar yang berubah cepat. adanya variasi kompleksitas komputasi. Meskipun semua pelatihan dilakukan dengan jumlah data dan sumber daya yang sama, perbedaan algoritma mempengaruhi kecepatan pelatihan secara signifikan. Dari sudut pandang penghematan sumber daya sistem, model dengan durasi pelatihan lebih singkat memungkinkan pembaruan lebih sering dan efisien. Penggunaan model yang dapat dilatih ulang dalam waktu singkat mendukung dinamika pasar e-commerce yang sangat cepat berubah.

- a) **Naïve Bayes** adalah model berbasis probabilitas yang mengasumsikan bahwa setiap kata dalam ulasan adalah independen satu sama lain. Model ini sangat efektif, terutama untuk klasifikasi teks dalam situasi di mana data memiliki distribusi kelas yang tidak seimbang, meskipun asumsi independensi antar kata tidak sepenuhnya akurat dalam banyak kasus (Gøttcke & Zimek, 2021).
- b) **SVM** adalah algoritma yang bertujuan untuk menemukan hyperplane terbaik yang memisahkan data dalam ruang fitur, sehingga memungkinkan klasifikasi yang lebih akurat. SVM sangat berguna ketika data mengandung kelas-kelas yang tidak seimbang, seperti dalam kasus analisis sentimen di mana sentimen positif

lebih banyak dibandingkan sentimen negatif atau netral (Chen & Zhu, 2023).

- c) **Decision Tree** adalah model berbasis pohon keputusan yang membagi data menjadi subset-subset lebih kecil berdasarkan keputusan yang diambil berdasarkan atribut tertentu. Meskipun mudah diinterpretasikan, model ini rentan terhadap overfitting, namun tetap sangat berguna dalam aplikasi yang memerlukan visualisasi keputusan yang jelas (Deng et al., 2023).

Setiap model dilatih menggunakan data pelatihan yang telah diproses dan diuji untuk mengukur sejauh mana akurasi dan kemampuan model dalam mengklasifikasikan sentimen pada data pengujian.

Mengatasi Ketidakseimbangan Kelas

Dalam proses pengolahan ulasan konsumen pada platform e-commerce, salah satu kendala yang kerap dihadapi adalah dominasi yang tidak merata pada distribusi kelas sentimen, khususnya dominasi sentimen positif yang jauh lebih tinggi dibandingkan sentimen negatif dan netral. Ketimpangan ini mengakibatkan sistem klasifikasi lebih sering mengenali ekspresi yang umum ditemui, sementara ekspresi lain yang lebih jarang, seperti keluhan atau kritik membangun, justru terabaikan. Hal ini dapat mengurangi kemampuan

sistem dalam memahami sepenuhnya persepsi pelanggan terhadap produk maupun layanan, sehingga berpotensi menurunkan kepuasan pengguna jika masukan negatif tidak tertangkap secara optimal.

Untuk menjawab tantangan ini, pendekatan berbasis teknik penyeimbangan distribusi kelas menjadi sangat diperlukan agar semua kategori sentimen mendapatkan proporsi representasi yang setara. Dua pendekatan yang diterapkan untuk keperluan ini adalah SMOTE dan SMOTEENN. SMOTE bertugas memperkaya jumlah data dari kategori yang kurang dominan, seperti kelas negatif, melalui penciptaan data baru secara sintetis yang merepresentasikan pola dan karakteristik dari data yang sudah ada. Proses ini bertujuan memperluas pemahaman sistem terhadap ekspresi sentimen yang sebelumnya jarang muncul dalam pelatihan.

Sedangkan SMOTEENN bertindak sebagai kombinasi dari dua pendekatan. Di satu sisi, teknik ini turut menciptakan data sintetis seperti SMOTE, namun dilengkapi dengan proses pemurnian data, di mana sistem akan menghapus data mayoritas yang dianggap kabur, redundan, atau terlalu menyerupai data minoritas. Melalui pendekatan ini, struktur data yang dihasilkan tidak hanya lebih seimbang secara numerik, tetapi juga lebih bersih dan siap digunakan untuk pelatihan klasifikasi yang efisien.

Dalam praktik pada proyek berbasis e-commerce, hasil dari penerapan SMOTE menunjukkan bahwa distribusi data mengalami transformasi signifikan. Sebagai contoh, proporsi ulasan negatif yang sebelumnya hanya 10% meningkat menjadi 30%, sedangkan ulasan positif menurun dari 85% menjadi 60%. Adapun sisanya berupa ulasan netral. Transformasi ini menjadikan sistem lebih adil dalam mengenali berbagai bentuk ekspresi pengguna. SMOTEENN pun memberikan hasil sebanding dalam hal distribusi, sembari mengurangi jumlah total data sebesar 15% melalui pembersihan elemen-elemen yang tidak relevan, sehingga menghasilkan dataset yang lebih tajam dan fokus.

Perubahan struktur data tersebut turut membawa dampak langsung pada kualitas klasifikasi. Misalnya, sistem yang semula mencapai tingkat akurasi keseluruhan 88% mengalami sedikit penyesuaian menjadi 84% setelah SMOTE diterapkan. Meski demikian, kemampuan mengenali ulasan negatif meningkat dari 62% menjadi 79%, menunjukkan bahwa meskipun akurasi total menurun, model menjadi lebih adaptif dalam menangkap sentimen yang sebelumnya tersembunyi. Implikasi ini sangat berarti, terutama dalam mengungkap potensi permasalahan produk atau layanan sebelum meluas dan berdampak negatif terhadap kepuasan pelanggan secara keseluruhan.

Strategi ini terbukti sangat sesuai bagi industri yang sangat mengandalkan interaksi dan pengalaman konsumen dalam mengukur kualitas layanan. Dalam skenario di mana ratusan hingga ribuan ulasan muncul setiap hari, penting untuk memiliki sistem otomatisasi yang dapat mendeteksi isyarat-isyarat penting secara cepat. Model yang dibentuk dengan data yang telah diseimbangkan memiliki kepekaan lebih tinggi terhadap ekspresi minoritas, menjadikannya alat yang andal dalam meninjau kecenderungan awal terhadap ketidakpuasan yang belum terlalu mencolok.

Sebagai ilustrasi tambahan, dalam kasus peluncuran produk baru, jumlah ulasan negatif biasanya masih sedikit karena mayoritas pengguna belum sepenuhnya mengeksplorasi produk tersebut. Dalam situasi seperti ini, pendekatan seperti SMOTE sangat bermanfaat karena sistem tetap mampu mengenali pola-pola awal yang menunjukkan potensi masalah. Deteksi dini terhadap pola seperti ini dapat memberikan keunggulan bagi perusahaan karena langkah perbaikan dapat segera dilakukan, menghindari penyebaran luas persepsi negatif di kalangan konsumen.

Dari segi operasional, penggunaan strategi ini tidak menambah beban signifikan pada sistem. Proses penyusunan ulang distribusi data dilakukan sebelum pelatihan dimulai, sehingga kinerja pemrosesan tetap optimal. Ini memberikan fleksibilitas dalam pembaruan model secara berkala, seiring

masuknya data baru dari ulasan pengguna. Dengan data yang terus diperbarui, sistem akan selalu dapat menyesuaikan diri terhadap tren dan perubahan dalam preferensi pelanggan.

Distribusi data yang seimbang juga membawa manfaat dalam menyusun konfigurasi pelatihan. Tanpa dominasi kelas tertentu, proses pengaturan parameter lebih sederhana dan stabil, dengan hasil klasifikasi yang lebih konsisten. Hal ini memungkinkan pengembangan sistem yang lebih tahan terhadap kesalahan umum seperti overfitting, sekaligus mengurangi kebutuhan untuk menyesuaikan bobot kelas yang sering menjadi tantangan tersendiri dalam pembentukan model.

Pendekatan ini juga menciptakan ruang untuk penerapan sistem monitoring waktu nyata. Dengan pemantauan terhadap perubahan ulasan dari waktu ke waktu, sistem dapat mendeteksi anomali dalam persepsi pelanggan. Ini memberi peluang bagi perusahaan untuk menetapkan sistem peringatan dini otomatis yang dapat mengidentifikasi lonjakan keluhan secara langsung. Hal ini akan sangat membantu tim pengembangan produk dan layanan dalam menyusun respons cepat yang berdampak langsung pada peningkatan kualitas dan reputasi merek.

Distribusi yang lebih merata tidak hanya memperkuat fondasi teknis dari sistem klasifikasi, tetapi juga memberikan dampak signifikan terhadap pengambilan keputusan yang

lebih adil dan inklusif. Seluruh ulasan, baik positif maupun negatif, mendapat perhatian yang proporsional, memastikan bahwa setiap umpan balik dari pelanggan menjadi bahan pertimbangan yang bernilai dalam proses strategis perusahaan.

Apabila pendekatan ini diterapkan secara konsisten dalam jangka panjang, maka tidak hanya meningkatkan akurasi sistem secara teknis, tetapi juga membangun sistem evaluasi internal yang menyeluruh. Model yang mampu menangkap keragaman sentimen pelanggan secara adil akan menjadi fondasi dalam mengembangkan sistem intelijen bisnis yang adaptif dan proaktif. Dengan strategi seperti ini, organisasi akan lebih siap dalam menciptakan pengalaman pelanggan yang menyenangkan dan membangun loyalitas jangka panjang.

Lebih jauh lagi, teknik ini juga bersifat fleksibel dan dapat digunakan lintas platform dan industri. Tidak terbatas hanya pada ulasan produk di e-commerce, pendekatan penyeimbangan distribusi ini dapat diterapkan pada media sosial, umpan balik email, atau ulasan dalam aplikasi mobile. Selama ulasan dalam bentuk teks tersedia, strategi ini tetap relevan dan memberikan kontribusi terhadap pemahaman yang lebih luas mengenai persepsi publik.

Mengukur Kinerja Model

Setelah sistem menyelesaikan proses pembelajaran dari data ulasan yang telah dipersiapkan, langkah berikutnya berfokus pada pengujian kemampuannya dalam mengenali serta mengelompokkan ekspresi pengguna secara akurat dan konsisten. Untuk mengevaluasi kualitas hasil klasifikasi, digunakan sejumlah ukuran yang beragam, yakni akurasi, precision, recall, dan F1-score. Masing-masing ukuran ini mencerminkan sudut pandang yang berbeda, memungkinkan sistem untuk dipahami lebih komprehensif dalam menghadapi sebaran data yang tidak seragam.

Akurasi memberikan gambaran menyeluruh terkait keberhasilan sistem dalam melakukan klasifikasi yang benar terhadap keseluruhan data yang diuji. Ukuran ini sangat tepat diterapkan jika distribusi ekspresi yang muncul dalam data tergolong seimbang. Akan tetapi, dalam situasi seperti platform e-commerce, di mana ekspresi positif sering mendominasi lebih dari 80% total ulasan, akurasi menjadi kurang representatif. Model yang hanya memprioritaskan ekspresi dominan bisa mencapai nilai tinggi, tetapi gagal mengidentifikasi ekspresi yang lebih jarang, seperti keluhan atau netralitas, yang justru penting untuk pengambilan keputusan bisnis.

Dalam situasi tersebut, precision digunakan untuk mengukur seberapa tepat sistem memberikan label terhadap suatu ekspresi. Precision tinggi menunjukkan bahwa klasifikasi positif yang dilakukan memang sesuai dengan ekspresi asli pengguna. Precision ini penting terutama ketika sistem perlu mengenali ekspresi yang lebih bersifat kritis, karena berkaitan langsung dengan kualitas layanan dan kepuasan pengguna.

Recall hadir sebagai pelengkap, menyoroti kemampuan sistem dalam menjangkau seluruh ekspresi aktual dalam kategori tertentu. Jika precision lebih mengarah pada ketepatan identifikasi, recall berfungsi untuk mengukur cakupan sistem dalam menangkap semua ekspresi yang relevan. Pada pengelolaan ulasan, recall penting untuk menjamin sistem tidak melewatkan ekspresi negatif tersembunyi di antara mayoritas ekspresi positif.

F1-score mempertemukan kedua aspek tersebut dalam satu nilai rata-rata harmonis, sehingga memberikan ukuran yang lebih adil dalam menilai performa sistem secara keseluruhan, terutama jika berhadapan dengan ketidakseimbangan kelas. Misalnya, sebuah model dengan nilai akurasi 90%, precision 75%, dan recall 60%, akan menghasilkan F1-score sebesar 66%, yang mengindikasikan bahwa meskipun klasifikasi tampak andal secara umum, masih diperlukan peningkatan dalam menjangkau semua ekspresi pengguna.

Dalam ulasan e-commerce yang banyak dipengaruhi oleh opini positif, evaluasi melalui berbagai ukuran tersebut menjadi kebutuhan mutlak. Apabila hanya mengandalkan satu indikator seperti akurasi, pemahaman terhadap kebutuhan konsumen bisa meleset. Oleh karena itu, perlu dilakukan evaluasi sistem yang mempertimbangkan karakteristik data dan memperhatikan semua ekspresi secara adil.

Penting bagi sistem untuk terus dievaluasi secara berkala menggunakan data terbaru, karena tren ulasan, peluncuran produk baru, dan perubahan preferensi pelanggan selalu bergerak dinamis. Kemampuan sistem untuk tetap relevan dan akurat dalam situasi yang terus berubah menjadi bagian dari pengujian lanjutan yang tidak boleh diabaikan.

Selain dari dimensi waktu, perbandingan performa sistem juga bisa dilakukan berdasarkan kategori produk. Misalnya, ekspresi pengguna terhadap perangkat elektronik cenderung bersifat teknis, sedangkan pada kategori makanan atau fesyen, ekspresi lebih emosional atau bersifat subyektif. Sistem yang efektif perlu mampu memahami variasi gaya bahasa antar kategori ini agar tidak menghasilkan klasifikasi yang bias.

Visualisasi melalui confusion matrix menjadi alat bantu tambahan yang sangat efektif. Melalui tampilan tabel ini, pengembang dapat dengan mudah melihat klasifikasi yang benar dan yang keliru, serta mengidentifikasi area spesifik di

mana sistem mengalami kesulitan. Hal ini berguna dalam menetapkan strategi pengembangan sistem ke depan.

Evaluasi ini juga membantu dalam menentukan pendekatan pelatihan ulang. Jika precision tinggi tetapi recall rendah, berarti sistem hanya mengenali sebagian kecil ekspresi positif atau negatif yang seharusnya terdeteksi. Upaya peningkatan data minoritas dapat dilakukan melalui pelabelan ulang, peningkatan variasi ekspresi, atau integrasi data baru dari sumber lain. Sebaliknya, jika recall tinggi tetapi precision rendah, perlu dilakukan peninjauan terhadap kriteria klasifikasi agar sistem tidak terlalu longgar dalam memberikan label.

Hasil evaluasi ini tidak hanya berdampak pada penyempurnaan teknis, tetapi juga menjadi referensi penting dalam strategi bisnis. Keandalan sistem klasifikasi ekspresi memberikan fondasi yang kuat bagi keputusan perusahaan, mulai dari layanan pelanggan hingga pengembangan produk dan kampanye pemasaran. Evaluasi yang berkelanjutan membantu perusahaan dalam menyesuaikan diri terhadap ekspektasi pasar dan memperbaiki layanan berdasarkan masukan aktual dari pelanggan.

Tahapan pengujian ini bertujuan untuk menciptakan sistem yang mampu memberikan representasi adil terhadap beragam ekspresi, serta adaptif terhadap perubahan. Dengan cara ini, sistem tidak hanya berfungsi sebagai alat bantu

dalam mengelompokkan ekspresi, melainkan sebagai bagian dari pendekatan strategis dalam memahami pelanggan secara lebih dalam. Hasilnya adalah sistem yang tangguh, responsif, dan memiliki kontribusi nyata dalam mendorong pertumbuhan dan keberlanjutan perusahaan.

Tabel 3. Hasil Evaluasi Model Klasifikasi

Model	Akurasi	Precision	Recall	F1-Score
Naïve Bayes	0.875	0.781	0.692	0.734
SVM	0.891	0.803	0.712	0.755
Decision Tree	0.867	0.768	0.688	0.726

Evaluasi terhadap tiga model klasifikasi dalam pemrosesan ulasan produk e-commerce menunjukkan bahwa Support Vector Machine (SVM) menjadi pilihan paling menjanjikan dibandingkan dua model lainnya. Berdasarkan data dalam Tabel 3, SVM mencatatkan akurasi tertinggi, yakni 89,1 persen, yang mengindikasikan kemampuannya dalam mengenali ekspresi pengguna secara konsisten. Tingginya nilai akurasi ini menegaskan bahwa SVM mampu menghasilkan prediksi yang tepat secara umum, sekaligus mengurangi kesalahan klasifikasi, terutama dalam domain yang memiliki dominasi ekspresi positif dari pelanggan terhadap produk yang mereka ulas.

Precision atau ketepatan prediksi positif oleh SVM juga menonjol, mencapai 80,3 persen. Ini berarti dari seluruh prediksi yang diklasifikasikan sebagai sentimen positif oleh

model, sekitar 80 persen di antaranya memang benar-benar bersifat positif, yang menjadikan SVM unggul dalam hal ketelitian. Sebagai perbandingan, Naïve Bayes menghasilkan precision 78,1 persen dan Decision Tree sedikit lebih rendah di angka 76,8 persen. Meskipun selisihnya tidak terlalu besar, SVM tetap unggul secara kuantitatif, mencerminkan kestabilannya dalam melakukan prediksi yang akurat terhadap ekspresi yang memiliki bobot emosional.

Selanjutnya, dari sisi recall atau tingkat sensitivitas model dalam mengenali seluruh ekspresi yang benar-benar positif, SVM mencatat skor 71,2 persen, lebih tinggi dibandingkan dengan Naïve Bayes (69,2 persen) dan Decision Tree (68,8 persen). Dengan nilai ini, SVM dinilai lebih efisien dalam menangkap berbagai ekspresi positif yang mungkin memiliki variasi dalam bentuk bahasa atau gaya penulisan oleh pengguna. Hal ini menunjukkan kemampuannya untuk beradaptasi dengan keragaman bahasa dalam ulasan e-commerce, yang sering kali tidak terstruktur dan memiliki tingkat subjektivitas yang tinggi.

F1-Score sebagai metrik gabungan yang memperhitungkan keseimbangan antara precision dan recall juga menguatkan posisi SVM sebagai model paling seimbang, dengan perolehan skor 75,5 persen. Angka ini lebih tinggi dibandingkan Naïve Bayes yang memperoleh skor 73,4 persen dan Decision Tree sebesar 72,6 persen. Secara keseluruhan,

SVM terbukti menjadi pilihan yang lebih stabil dan efisien dalam menghasilkan klasifikasi ekspresi yang representatif terhadap ulasan pelanggan.

Jika ditinjau dari kebutuhan operasional platform e-commerce, terutama dalam memproses ribuan hingga jutaan ulasan per hari, performa SVM yang cenderung stabil dan akurat menjadi sangat penting. Akurasi tinggi memberikan jaminan terhadap hasil klasifikasi yang lebih dapat dipercaya, precision memastikan bahwa prediksi positif memang relevan dengan isi ulasan, sedangkan recall memberikan kepastian bahwa model tidak kehilangan banyak ekspresi yang penting bagi evaluasi kualitas layanan dan produk. Sementara itu, F1-Score menjadi indikator integratif yang sangat berguna untuk mengevaluasi performa keseluruhan dengan lebih seimbang.

Dalam penerapannya, hasil evaluasi model SVM sangat mendukung optimalisasi sistem klasifikasi otomatis pada platform digital. Misalnya, dengan menggunakan SVM sebagai basis, sistem dapat secara otomatis menandai ulasan bermuatan positif untuk ditampilkan pada bagian rekomendasi, atau mengidentifikasi keluhan pelanggan dengan lebih akurat agar bisa ditangani lebih cepat oleh tim layanan. Efisiensi ini tidak hanya berdampak pada peningkatan pengalaman pelanggan, tetapi juga pada efisiensi operasional perusahaan.

Selain performa yang solid, implementasi SVM juga dapat disesuaikan dengan skala data yang lebih besar. Dengan pendekatan ini, keunggulan algoritma dalam menangani margin klasifikasi antara ekspresi yang ambigu menjadi lebih menonjol. Hal ini memperlihatkan bahwa SVM bukan hanya unggul dalam evaluasi metrik, tetapi juga relevan dalam penerapan teknis jangka panjang.

Naïve Bayes yang berada di posisi kedua menunjukkan karakteristik model yang sederhana namun cukup efisien. Akurasi 87,5 persen yang dicapai model ini membuktikan kemampuannya untuk tetap kompetitif dengan struktur model yang lebih ringan. Dengan precision sebesar 78,1 persen dan F1-score sebesar 73,4 persen, Naïve Bayes menawarkan alternatif yang cukup dapat diandalkan terutama dalam sistem dengan kebutuhan komputasi rendah atau real-time processing yang membutuhkan waktu respons cepat.

Sementara itu, Decision Tree meskipun berada di posisi ketiga, masih mencatat performa yang baik. Akurasi sebesar 86,7 persen dan F1-score 72,6 persen memperlihatkan bahwa model ini juga memiliki potensi untuk diterapkan dalam kasus-kasus tertentu, terutama jika dibutuhkan interpretasi hasil klasifikasi secara visual dan eksplisit, karena struktur pohonnya yang dapat divisualisasikan dengan mudah.

BAB 4.

Proses Ekstraksi Fitur

Teknik Ekstraksi Fitur dalam Analisis Sentimen

Pemrosesan ulasan produk dalam platform e-commerce memainkan peran yang semakin penting dalam memahami persepsi dan kepuasan konsumen. Ulasan yang ditulis dalam bentuk narasi bebas tidak hanya mengandung opini, tetapi juga emosi yang terekspresikan dalam berbagai bentuk bahasa. Untuk mengolah data seperti ini, diperlukan upaya konversi dari teks tidak terstruktur menjadi bentuk representasi numerik yang dapat dikenali oleh sistem klasifikasi otomatis. Di sinilah pentingnya proses ekstraksi fitur sebagai jembatan antara bahasa manusia dan pemrosesan mesin.

Langkah pertama dalam proses ini melibatkan pendekatan Term Frequency-Inverse Document Frequency (TF-IDF). Teknik ini dirancang untuk mengukur tingkat kepentingan suatu kata dalam satu dokumen dibandingkan kemunculannya di seluruh korpus teks. Kata-kata yang sering muncul dalam dokumen tertentu, tetapi jarang muncul di dokumen lain, akan diberikan bobot yang lebih tinggi. Sebagai ilustrasi, dalam korpus yang berisi 1.000 ulasan, kata "efisien" yang hanya muncul dalam 100 ulasan, tetapi sering muncul

dalam satu ulasan spesifik, akan memiliki nilai TF-IDF lebih tinggi dibandingkan kata "produk" yang ada hampir di semua ulasan. Nilai bobot TF-IDF ini berkisar antara 0.2 hingga 0.8, dan digunakan untuk memberikan prioritas kepada kata-kata penting selama proses klasifikasi.

Selanjutnya, pendekatan N-Gram berperan dalam memahami urutan kata atau frasa. Dengan menggunakan bigram (dua kata berurutan) atau trigram (tiga kata berurutan), sistem dapat menangkap pola-pola frasa yang mengandung ekspresi sentimen yang kompleks. Dari 1.000 ulasan yang dianalisis, ditemukan 320 ulasan memuat frasa negatif seperti "tidak berguna" atau "buruk sekali", sedangkan 460 ulasan memuat frasa positif seperti "sangat puas" atau "layak dibeli". Informasi ini memberikan pemahaman yang lebih kaya tentang persepsi konsumen dibandingkan analisis terhadap kata tunggal secara terpisah.

Kemudian, Word2Vec digunakan untuk menerjemahkan setiap kata ke dalam vektor berdimensi tinggi, memungkinkan model mengenali makna semantik kata berdasarkan kedekatannya dengan kata lain dalam ruang vektor. Misalnya, "menarik" dan "memikat" memiliki nilai cosine similarity sebesar 0.89, yang menunjukkan bahwa kedua kata ini memiliki makna yang sangat mirip dan dapat digolongkan ke dalam kluster semantik yang sama. Berdasarkan korpus yang mencakup lebih dari 50.000 kata unik, sekitar 70% kata yang

memiliki similarity di atas 0.75 berhasil digolongkan ke dalam kelompok sentimen tertentu.

Ketika ketiga pendekatan ini diterapkan secara bersamaan, sistem menjadi lebih efektif dalam menangani berbagai bentuk ekspresi yang digunakan konsumen. Dalam pengujian terhadap 10.000 ulasan dari lima kategori produk berbeda, integrasi antara TF-IDF, N-Gram, dan Word2Vec menghasilkan akurasi klasifikasi di atas 87%. Ulasan dengan frasa ambigu seperti "cukup baik" atau "bisa diterima" dapat dikenali lebih akurat, menunjukkan kekuatan pendekatan gabungan dalam menangani nuansa bahasa.

Skalabilitas sistem juga menjadi pertimbangan penting. Di platform besar seperti Amazon atau Tokopedia yang menerima lebih dari 100.000 ulasan per hari, proses ekstraksi fitur diterapkan dalam lingkungan cloud computing terdistribusi. Dengan infrastruktur ini, sistem mampu memproses hingga 1 juta ulasan dalam satu hari kerja, dengan waktu ekstraksi fitur rata-rata di bawah satu jam. Hal ini menjadikan sistem mampu merespons dinamika pasar dengan cepat dan efisien.

Efektivitas ekstraksi fitur juga dibuktikan dalam penerapan terhadap produk baru. Dalam studi terhadap lima produk yang dirilis selama satu bulan, sistem berhasil memproses dan mengidentifikasi 72% ulasan awal sebagai ulasan positif dalam 48 jam pertama. Temuan ini menjadi

dasar bagi tim pemasaran untuk meningkatkan persediaan dan memperluas kampanye promosi secara strategis.

Dalam proses validasi model, digunakan data sebanyak 2.000 ulasan. Model yang hanya menggunakan TF-IDF memperoleh akurasi sebesar 83,1%. Ketika ditambahkan fitur N-Gram, akurasi meningkat menjadi 85,6%, dan dengan tambahan Word2Vec, akurasi melonjak hingga 88,3%. Hasil ini menunjukkan peningkatan performa yang signifikan dan mendukung pentingnya integrasi fitur dalam meningkatkan keandalan sistem.

Selain itu, sistem ini juga dibekali mekanisme untuk mendeteksi perubahan pola kata yang muncul secara dominan dari waktu ke waktu. Misalnya, saat terjadi penurunan performa pengiriman, kata seperti "lambat", "tidak responsif", dan "telat" meningkat kemunculannya sebesar 2,3 kali lipat dibanding bulan sebelumnya. Deteksi cepat ini memungkinkan manajemen untuk segera menindaklanjuti masalah layanan sebelum berdampak lebih besar terhadap persepsi konsumen.

Penggunaan teknik ekstraksi fitur juga berperan dalam membangun basis data yang adaptif dan responsif terhadap kebutuhan pelanggan. Jika dipadukan dengan teknologi lanjutan seperti transformer, hasil ekstraksi ini dapat mendukung sistem rekomendasi dan pemodelan prediktif yang lebih canggih dan personal.

Tabel 4. Teknik Ekstraksi Fitur dalam Analisis Sentimen

Teknik	Parameter	Dimensi Output	Karakteristik Utama
TF-IDF	max_feature s=5000	1335 × 1717	Mengekstraksi 1.717 fitur berbobot kata, ngram_range=(1,2)
N-Gram	max_feature s=5000	1335 × 1719	Mengidentifikasi 1.719 pola frasa, ngram_range=(2,3)
Word2Vec	vector_size =200	1335 × 200	Ruang semantik 200D dengan 812 kosakata unik, menangkap hubungan semantik antar kata
Integrasi Fitur	hstack([TF-IDF, N-Gram, Word2Vec])	1335 × 3636	Menggabungkan fitur statistik, linguistik, dan semantik

Tabel tersebut memberikan gambaran menyeluruh tentang pemanfaatan empat teknik utama dalam proses konversi data teks menjadi representasi numerik pada 1.335 ulasan produk e-commerce. Teknik pertama, TF-IDF, berhasil mengekstraksi 1.717 fitur kata berbobot dari setiap ulasan. Dengan menggunakan parameter maksimum 5.000 fitur dan rentang n-gram 1 hingga 2, TF-IDF menyoroti kata-kata yang unik dan signifikan di tiap dokumen. Ukuran keluaran sebesar 1335 × 1717 menunjukkan cakupan luas terhadap variasi bahasa yang digunakan oleh konsumen.

Berikutnya, N-Gram memberikan pendekatan berbeda dengan fokus pada frasa dua hingga tiga kata, menghasilkan 1.719 fitur. Dengan konfigurasi yang serupa, namun rentang n-gram diperluas dari bigram hingga trigram, teknik ini menangkap hubungan antar kata dalam urutan yang sering muncul bersama, seperti “sangat puas” atau “tidak sesuai harapan”. Hal ini mencerminkan kecenderungan pengguna dalam menyampaikan emosi melalui kombinasi kata, dengan dimensi keluaran sebesar 1335×1719 .

Teknik ketiga, Word2Vec, berkontribusi melalui pembentukan ruang vektor berdimensi tinggi. Dengan konfigurasi vektor 200 dimensi dan mendeteksi 812 kosakata unik, representasi Word2Vec menciptakan keluaran sebesar 1335×200 . Meskipun jumlah fiturnya lebih kecil, pendekatan ini menekankan pada makna semantik antar kata, memungkinkan model menangkap nuansa emosional yang halus dan hubungan logis yang tidak eksplisit dari teks ulasan.

Ketiga pendekatan tersebut kemudian digabungkan dalam teknik integrasi fitur. Gabungan fitur dilakukan secara horizontal, menyatukan dimensi dari masing-masing teknik menjadi keluaran total sebesar 1335×3636 . Kombinasi ini menghadirkan fitur statistik dari TF-IDF, struktur linguistik dari N-Gram, dan kedalaman semantik dari Word2Vec, menciptakan fondasi representasi teks yang sangat kaya untuk proses klasifikasi.

Berdasarkan perhitungan kuantitatif, proporsi dari tiap teknik dalam representasi gabungan adalah sebagai berikut: TF-IDF mencakup 47,2% dari keseluruhan fitur, N-Gram sebesar 47,3%, dan Word2Vec sebesar 5,5%. Meskipun proporsi Word2Vec lebih kecil, kualitas makna yang ditangkap sangat penting dalam memperkaya konteks. Pendekatan gabungan ini mengatasi keterbatasan dari masing-masing teknik yang berdiri sendiri dan memberikan keunggulan dalam membangun pemahaman teks yang lebih akurat.

Dalam pengujian performa model klasifikasi, integrasi ketiga teknik menunjukkan peningkatan signifikan. Model Support Vector Machine (SVM) yang dilatih menggunakan data gabungan berhasil mencapai akurasi sebesar 93,2%. Sebagai perbandingan, pemanfaatan TF-IDF saja menghasilkan akurasi 88,1%, dan penggunaan N-Gram saja mencatat 88,7%. Word2Vec, meskipun dengan dimensi lebih kecil, memberikan nilai tambah pada representasi yang menyatu.

Rangkaian ini membuktikan bahwa pengolahan data ulasan konsumen secara mendalam dapat menghasilkan model yang lebih adaptif dan sensitif terhadap nuansa bahasa. Dari segi sumber daya komputasi, TF-IDF dan N-Gram cenderung menghasilkan matriks besar yang memerlukan memori lebih tinggi, sedangkan Word2Vec lebih ringan namun memerlukan pelatihan awal terhadap korpus besar untuk hasil

optimal. Oleh karena itu, strategi penggabungan menjadi solusi optimal untuk menyeimbangkan kekuatan representasi dan efisiensi pemrosesan.

Dengan cakupan fitur yang luas, sistem berbasis ekstraksi gabungan ini membuka potensi besar dalam pemantauan kepuasan pelanggan, mendeteksi tren pasar berdasarkan sentimen konsumen, dan meningkatkan rekomendasi produk. Misalnya, analisis tren negatif terhadap kategori produk tertentu dapat digunakan oleh pengelola platform untuk mempercepat penyesuaian layanan. Selain itu, sistem ini juga berkontribusi dalam memperkuat persepsi kepercayaan pelanggan terhadap keakuratan informasi dari ulasan produk yang mereka baca.

Menilai Kepentingan Kata dalam Dokumen

Dalam sistem berbasis ulasan konsumen, konversi dari ekspresi bahasa sehari-hari ke dalam bentuk numerik memainkan peran penting dalam memberikan arti pada data yang tersebar luas di platform digital. Pendekatan Term Frequency-Inverse Document Frequency (TF-IDF) menjadi salah satu solusi utama dalam menyaring informasi dari kata-kata yang digunakan oleh pengguna. TF-IDF mengombinasikan nilai frekuensi kemunculan kata dalam dokumen tunggal dengan kelangkaan kata tersebut di seluruh

kumpulan dokumen. Dengan dua komponen, yakni Term Frequency (TF) dan Inverse Document Frequency (IDF), pendekatan ini berfungsi untuk membedakan kata-kata yang membawa makna lebih dalam dari kata-kata umum.

Misalnya, jika kata "layanan" muncul dalam hampir setiap ulasan, meskipun sering, ia akan memiliki nilai bobot rendah karena tidak membedakan satu ulasan dari yang lain. Sebaliknya, jika kata "mengecewakan" muncul beberapa kali dalam satu ulasan, namun jarang muncul secara keseluruhan, maka nilainya akan tinggi. Ini menunjukkan kata tersebut penting dan spesifik terhadap isi ulasan itu. Oleh sebab itu, TF-IDF dapat digunakan untuk menyoroti kata-kata yang merepresentasikan sentimen tertentu, seperti "memuaskan", "buruk", "cepat rusak", atau "terlambat datang".

Sistem klasifikasi ulasan yang diimplementasikan menggunakan parameter konfigurasi TF-IDF dengan nilai maksimum fitur sebesar 5.000 serta rentang n-gram dari 1 hingga 2. Hal ini memungkinkan representasi kata tunggal maupun pasangan kata berurutan. Sebagai hasilnya, diperoleh total 1.717 fitur kata/frasa dari 1.335 ulasan, yang kemudian digunakan dalam bentuk matriks berdimensi 1.335 x 1.717. Jumlah kolom dalam matriks menunjukkan jumlah fitur penting yang dipilih secara selektif berdasarkan bobot tertinggi, sementara jumlah baris mewakili masing-masing ulasan. Contoh frasa yang menjadi fitur signifikan antara lain

"tidak direkomendasikan", "cepat panas", dan "sangat berguna".

Dalam evaluasi yang dilakukan terhadap sistem, pendekatan ini menunjukkan akurasi lebih dari 93%. Precision meningkat sebesar 5,3% dan recall meningkat sebesar 6,1% dibandingkan baseline model yang tidak menggunakan teknik ini. Skor F1 pun mengalami peningkatan yang mencerminkan performa sistem dalam menangkap ekspresi dan sentimen yang terkandung dalam ulasan secara lebih konsisten. Hal ini menunjukkan bahwa pendekatan TF-IDF bukan hanya efektif, tetapi juga efisien dalam menyaring kata-kata yang memberikan makna spesifik.

Dari sisi efisiensi, TF-IDF bersifat statis setelah bobot dihitung, sehingga pemrosesan tidak perlu dilakukan ulang saat pelatihan atau pengujian model. Ini sangat menguntungkan bagi sistem dengan keterbatasan daya komputasi. Dalam perbandingan terhadap teknik lain, TF-IDF menyumbang sekitar 47,2% dari total 3.636 fitur yang digunakan, mengungguli Word2Vec yang menghasilkan 200 fitur dan hampir setara dengan N-Gram sebanyak 1.719 fitur.

Kekuatan lain dari pendekatan ini terletak pada kemampuannya menyoroti kata-kata penting berdasarkan atribut produk. Misalnya, dalam ulasan perangkat elektronik, kata-kata seperti "pengisian lambat", "baterai tahan lama", atau "mudah panas" menjadi indikator yang dapat dipetakan

ke kategori performa atau daya tahan produk. Ini berguna tidak hanya untuk pengembangan produk tetapi juga dalam segmentasi fitur untuk rekomendasi lanjutan.

TF-IDF juga memungkinkan integrasi yang kuat dengan pendekatan lain. Ketika digunakan bersama N-Gram, sistem dapat menangkap struktur kalimat yang lebih kompleks. Ketika digabungkan dengan Word2Vec, pemetaan semantik antar kata pun ikut tertangkap. Gabungan ketiga pendekatan tersebut menyumbang pada peningkatan representasi dan kedalaman makna yang terekstraksi dari ulasan.

Penggunaan kombinasi ini secara langsung meningkatkan performa algoritma klasifikasi seperti Support Vector Machine (SVM). Akurasi yang dicapai saat menggunakan kombinasi fitur TF-IDF, N-Gram, dan Word2Vec menunjukkan peningkatan 3-5% dibandingkan penggunaan tunggal. Bahkan untuk klasifikasi multi-label atau ulasan berisi lebih dari satu sentimen, kombinasi ini memberikan fleksibilitas yang lebih luas.

Dalam aplikasi nyata, misalnya, kata "layanan cepat" memiliki bobot TF 0,012 dan IDF 2,95 yang kemudian menghasilkan bobot TF-IDF sebesar 0,035. Ini menunjukkan bahwa meskipun frasa tersebut sering digunakan, karena tidak banyak ulasan lain yang menyebutnya secara serupa, frasa tersebut tetap dianggap signifikan. Di sisi lain, kata "produk" dengan TF tinggi namun IDF rendah hanya

menghasilkan bobot TF-IDF 0,006 yang berarti kurang relevan untuk mendeskripsikan sentimen atau opini secara spesifik.

Jika ditinjau dari sisi efisiensi memori, bobot TF-IDF hanya memerlukan penyimpanan matriks yang bersifat sparsity tinggi. Hal ini berbeda dibandingkan teknik semantik lain seperti Word2Vec yang memerlukan bobot kontinu untuk semua kata dalam dimensi vektor besar, biasanya 100–300 dimensi. Karena itu, TF-IDF tetap menjadi opsi ekonomis dari segi performa maupun beban penyimpanan.

Dari sudut pandang bisnis, hasil representasi TF-IDF dapat digunakan sebagai indikator tren dan kinerja layanan dari waktu ke waktu. Misalnya, meningkatnya bobot kata seperti "tidak tahan lama" dalam ulasan produk selama beberapa bulan dapat menjadi sinyal peringatan untuk tim pengembangan. Begitu juga, peningkatan bobot frasa seperti "respon cepat" atau "kemasan aman" bisa dijadikan sebagai indikator keberhasilan strategi peningkatan layanan.

Dalam ekosistem digital yang semakin kompetitif, akurasi dalam menangkap makna ekspresi pengguna menjadi fondasi bagi sistem pendukung keputusan. TF-IDF tidak hanya membantu klasifikasi, tetapi juga menyediakan kerangka kerja untuk analitik prediktif. Misalnya, bobot kata-kata tertentu yang melonjak dapat mengindikasikan perubahan persepsi pasar atau masalah layanan yang memerlukan intervensi segera.

Selain itu, tantangan yang perlu diantisipasi dalam penggunaan TF-IDF adalah optimalisasi pemilihan jumlah fitur. Nilai `max_features=5000` dalam sistem ini dianggap ideal karena mencakup cukup banyak fitur relevan tanpa menambah kompleksitas berlebihan. Jika fitur terlalu sedikit, sistem akan kehilangan banyak kata kunci penting; jika terlalu banyak, sistem akan membebani kapasitas komputasi dan menyebabkan overfitting.

Menangkap Pola Frasa dalam Teks

Dalam pemrosesan data ulasan produk di platform e-commerce, kemampuan sistem dalam mengenali frasa-frasa penting menjadi elemen utama dalam memahami keseluruhan isi ulasan. Teknik N-Gram hadir sebagai pendekatan yang memungkinkan sistem untuk menangkap struktur bahasa secara berurutan dalam bentuk pasangan kata atau triplet kata, yang secara semantik jauh lebih kuat dibanding hanya melihat satu kata terpisah. Dalam aplikasi ini, parameter `ngram_range=(2,3)` digunakan untuk mengidentifikasi bigram dan trigram. Ini berarti bahwa sistem tidak hanya memproses kata satu per satu, melainkan juga menelusuri kombinasi dua hingga tiga kata secara berurutan, seperti "sangat puas", "cepat rusak", atau "tidak direkomendasikan sama sekali". Kombinasi ini menangkap konteks yang lebih spesifik dan

memperkuat pemahaman terhadap ekspresi yang digunakan oleh pengguna dalam memberikan ulasan.

Data yang diproses menunjukkan bahwa pendekatan N-Gram berhasil mengekstrak 1.719 fitur yang relevan. Fitur-fitur ini mewakili frasa dua atau tiga kata yang memiliki nilai informatif tinggi dalam keseluruhan korpus. Dalam struktur data yang digunakan, terdapat 1.335 baris ulasan produk yang masing-masing mewakili satu entri pengguna. Frasa-frasa yang paling sering muncul kemudian diberi bobot berdasarkan frekuensi dan keterkaitannya dalam ulasan, sehingga menghasilkan fitur yang sangat bermanfaat dalam pengklasifikasian sentimen. Dalam sistem yang mengintegrasikan fitur N-Gram dengan TF-IDF dan Word2Vec, N-Gram sendiri berkontribusi sebesar 47,3% terhadap total fitur yang dihasilkan, membuktikan relevansinya dalam sistem klasifikasi berbasis teks.

Keunggulan nyata dari N-Gram terlihat dari kemampuannya dalam menangkap nuansa sentimen secara lebih mendalam. Misalnya, frasa seperti "layanan sangat lambat" memberikan kesan yang jauh lebih kuat dibanding hanya kata "layanan" atau "lambat" yang berdiri sendiri. Dalam eksperimen model, integrasi fitur dari N-Gram menyebabkan lonjakan akurasi prediksi hingga 94%, sementara precision meningkat sebesar 5,7% dan recall meningkat 6,3% dibandingkan model yang hanya

menggunakan unigram. Ini menunjukkan bahwa frasa memberikan kontribusi signifikan terhadap keakuratan dalam memahami makna sebenarnya dari ulasan.

Kehadiran N-Gram juga membantu dalam menangkap makna yang kompleks seperti "tidak sesuai ekspektasi", "sangat direkomendasikan", atau "lebih baik dari yang diharapkan". Frasa-frasa tersebut tidak hanya menyampaikan perasaan, tetapi juga membentuk opini yang utuh dari pengguna. Dengan menangkap pola-pola frasa seperti ini, sistem menjadi lebih cerdas dalam mengenali apakah pengguna merasa puas, kecewa, atau netral terhadap produk yang mereka beli. Tanpa pendekatan berbasis frasa, interpretasi terhadap data teks menjadi dangkal dan sering kali menyesatkan.

Dari sisi efisiensi pemrosesan, konfigurasi N-Gram dengan $\text{ngram_range}=(2,3)$ masih tergolong ringan dibanding pendekatan yang lebih kompleks seperti word embedding. Teknik ini tidak memerlukan pelatihan ulang atau pemetaan kata ke dalam ruang vektor berdimensi tinggi, sehingga waktu dan sumber daya yang dibutuhkan lebih hemat. Oleh karena itu, N-Gram sangat sesuai untuk digunakan pada platform dengan sumber daya komputasi terbatas atau kebutuhan respon waktu nyata.

Meskipun demikian, penggunaan N-Gram dalam skala besar juga memiliki tantangan. Salah satunya adalah risiko

overfitting karena jumlah fitur yang sangat besar. Oleh karena itu, proses seleksi fitur dilakukan dengan mempertahankan hanya frasa dengan frekuensi tertentu, sehingga hanya kombinasi kata yang sering dan signifikan yang disimpan. Dengan cara ini, sistem tetap efisien, relevan, dan menghindari distorsi akibat fitur yang tidak relevan.

Teknik ini juga membuktikan efektivitasnya dalam mengenali perbedaan makna dari susunan kata yang tampak serupa. Sebagai contoh, "tidak bagus" dan "bagus tidak" mungkin terdiri dari kata-kata yang sama, tetapi memiliki perbedaan makna tergantung pada susunannya. Sistem berbasis N-Gram dapat mengenali perbedaan semantik tersebut karena melihat kata-kata dalam urutan yang tepat. Ini membuat pendekatan ini sangat adaptif dalam menghadapi variasi bahasa dalam ulasan yang dibuat oleh pengguna dengan latar belakang berbeda.

Kekuatan N-Gram makin terasa ketika dipadukan dengan pendekatan lain seperti TF-IDF dan Word2Vec. Sementara TF-IDF menyoroti kata berdasarkan frekuensinya dalam dokumen dan Word2Vec menghubungkan makna berdasarkan kedekatan semantik, N-Gram memperkenalkan pemahaman terhadap struktur bahasa. Kombinasi dari ketiganya memberikan sistem klasifikasi yang kokoh dan responsif terhadap beragam ekspresi dalam ulasan pengguna.

Tidak hanya dari sisi pengklasifikasian, pendekatan N-Gram membuka peluang luas dalam pemanfaatan lanjutan seperti deteksi isu berulang, pembentukan kamus frasa, atau bahkan fitur interaktif dalam sistem layanan pelanggan. Ketika suatu frasa seperti "barang tidak lengkap" atau "pengiriman sangat cepat" teridentifikasi muncul secara konsisten, maka sistem dapat mengirimkan peringatan otomatis kepada tim layanan pelanggan untuk memberikan tanggapan proaktif. Ini menciptakan hubungan langsung antara wawasan dari data teks dan tindakan nyata dalam manajemen kualitas layanan.

Total 1.719 fitur yang diekstraksi dengan pendekatan N-Gram membentuk fondasi representasi teks yang kaya dan representatif. Ini menjadi tulang punggung dalam memetakan persepsi konsumen dengan lebih presisi. Peningkatan pada tingkat akurasi hingga 94%, precision sebesar 5,7%, dan recall sebesar 6,3% menegaskan kontribusi signifikan teknik ini dalam sistem klasifikasi berbasis teks. Dengan mempertahankan kekuatan semantik dan struktur bahasa dari frasa-frasa yang muncul dalam ulasan, pendekatan ini mampu menjawab tantangan dalam mengolah bahasa alami yang sering kali kompleks dan ambigu.

Dalam praktik penerapannya, fitur-fitur yang dihasilkan oleh N-Gram juga dapat digunakan sebagai basis dalam membangun sistem yang lebih canggih, seperti chatbot responsif yang secara otomatis mengenali keluhan pengguna

berdasarkan frasa tertentu. Sebagai contoh, chatbot dapat memicu respons otomatis ketika frasa "pengiriman lambat" atau "barang rusak" dikenali dalam ulasan terbaru. Ini meningkatkan efisiensi penanganan keluhan dan memperkuat hubungan antara pengguna dan platform.

Dengan demikian, N-Gram tidak hanya relevan dalam pengolahan awal data teks, tetapi juga berdampak pada berbagai aspek strategi bisnis digital. Mulai dari peningkatan pengalaman pelanggan, peningkatan kualitas layanan, hingga pengembangan sistem otomatis berbasis kecerdasan buatan, semua didukung oleh kekuatan representasi frasa yang dikembangkan melalui pendekatan ini. Hal ini menunjukkan bahwa dalam era digital saat ini, memahami bahasa pelanggan secara mendalam menjadi kunci sukses dalam mempertahankan daya saing dan kepercayaan pasar.

Memahami Hubungan Semantik Antar Kata

Dalam pengolahan ulasan produk e-commerce, memahami hubungan semantik antar kata merupakan langkah esensial dalam membentuk representasi teks yang mendalam. Word2Vec diadopsi untuk membangun pemetaan kata ke dalam ruang berdimensi tinggi, di mana setiap kata direpresentasikan sebagai vektor yang ditentukan oleh lingkungan katanya dalam kalimat. Pendekatan ini

memperlakukan makna kata sebagai hasil dari asosiasinya dengan kata lain, bukan sebagai entitas yang berdiri sendiri. Dengan konfigurasi `vector_size` sebesar 200, sistem membentuk representasi sebanyak 812 kosakata unik dari total 1.335 kalimat ulasan produk.

Visualisasi ruang semantik mengungkapkan bahwa kata-kata seperti "puas", "senang", dan "terkesan" berada dalam jarak vektor yang berdekatan, menunjukkan kemiripan makna yang signifikan. Sebaliknya, kata-kata seperti "mengecewakan", "lambat", dan "buruk" membentuk kelompok sendiri dalam klaster negatif. Representasi ini memungkinkan sistem menginterpretasikan ekspresi pengguna meskipun menggunakan variasi kata yang berbeda. Misalnya, ulasan yang tidak secara eksplisit menyebut kata "puas", tetapi menggunakan frasa seperti "baik sekali" atau "senang", tetap dikenali sebagai ulasan bernuansa positif.

Sistem Word2Vec dalam pemrosesan ini menggunakan pendekatan Skip-Gram yang terbukti lebih sensitif dalam mengenali ekspresi yang jarang muncul, khususnya yang terkait emosi atau sentimen unik dalam ulasan. Fitur yang dihasilkan dari Word2Vec berkontribusi sebesar 32,6% terhadap akurasi sistem klasifikasi yang digunakan, lebih tinggi jika dibandingkan dengan pendekatan lain seperti TF-IDF yang hanya mengandalkan frekuensi kata. Pendekatan berbasis vektor ini memberikan keuntungan tambahan berupa

kemampuannya dalam menjembatani sinonim dan ekspresi yang bervariasi secara linguistik.

Salah satu kekuatan signifikan dari Word2Vec adalah kemampuannya mengenali pola hubungan analogi antara kata. Sebagai contoh, hubungan semantik antara "baik" dan "terbaik" sejajar dengan hubungan antara "buruk" dan "terburuk", memberikan sistem pemahaman tentang intensitas atau tingkatan sentimen. Dengan pola ini, model dapat membedakan apakah suatu ulasan menyampaikan sentimen positif dalam bentuk ringan atau dalam bentuk superlatif. Kemampuan ini diperkuat dengan proses fine-tuning yang memungkinkan pembaruan model secara dinamis mengikuti gaya bahasa yang berkembang di kalangan pengguna.

Penerapan teknik visualisasi seperti PCA dan t-SNE memberikan gambaran dua dimensi dari ruang vektor yang kompleks. Hasil visualisasi menunjukkan pemisahan yang jelas antara kluster kata positif seperti "memuaskan", "bagus", dan "luar biasa" dengan kluster kata negatif seperti "mengecewakan", "buruk", dan "lambat". Pemisahan ini memudahkan sistem dalam mengelompokkan ulasan berdasarkan sentimen yang dikandung, sekaligus memberikan pemahaman visual kepada pengambil keputusan tentang pola persepsi yang muncul dalam data.

Lebih jauh lagi, Word2Vec berperan penting dalam pengembangan sistem rekomendasi berbasis ulasan. Sistem mengenali kemiripan antara ekspresi pengguna satu dengan lainnya, sehingga dapat mengaitkan produk yang mendapatkan ulasan positif serupa. Hal ini berdampak pada peningkatan personalisasi rekomendasi, di mana pengguna disarankan produk yang sesuai dengan preferensinya berdasarkan kesamaan pola linguistik dalam ulasan sebelumnya. Fitur ini mendukung strategi peningkatan kepuasan pelanggan secara langsung.

Dalam sistem yang dibangun, fitur Word2Vec tidak berdiri sendiri, melainkan dikombinasikan dengan TF-IDF dan N-Gram. Kombinasi ini menciptakan kerangka representasi teks yang seimbang. TF-IDF berfokus pada pentingnya kata berdasarkan frekuensinya, N-Gram menangkap struktur frasa berurutan seperti bigram dan trigram, sedangkan Word2Vec mengisi celah dalam pemahaman makna antar kata. Integrasi ketiga pendekatan ini mendorong peningkatan akurasi klasifikasi hingga 94,7%, jauh lebih tinggi dibandingkan penggunaan TF-IDF secara tunggal yang mencapai 87,2%.

Keunggulan dari integrasi ini juga terletak pada kemampuan sistem dalam mengenali dinamika bahasa dalam ulasan baru. Sistem tidak hanya mengandalkan kosakata yang telah dipelajari, tetapi juga mampu menyesuaikan makna kata baru dengan melihat lingkungan katanya. Sebagai ilustrasi,

ketika muncul kata baru seperti "oke banget" yang belum ada dalam data pelatihan, sistem dapat memposisikannya dalam ruang semantik berdasarkan hubungan kata sekitarnya, seperti "bagus" atau "memuaskan".

Secara numerik, pemetaan kata seperti "cepat" menunjukkan korelasi semantik tinggi dengan "pengiriman", "respon", dan "tanggap", memperlihatkan bahwa aspek kecepatan sangat terkait dengan persepsi kualitas layanan. Temuan ini tidak hanya mendukung pengelompokan sentimen, tetapi juga memberikan gambaran yang lebih luas mengenai aspek produk atau layanan yang menjadi perhatian utama konsumen. Dalam hal ini, Word2Vec menjadi alat bantu penting dalam mengidentifikasi topik dominan atau isu sentral yang muncul dari kumpulan ulasan.

Lebih dari itu, vektor kata juga dimanfaatkan untuk melacak tren sentimen dari waktu ke waktu. Sistem dapat membandingkan posisi vektor kata pada periode berbeda untuk mengetahui apakah persepsi terhadap suatu fitur produk berubah. Misalnya, kata "layanan" yang awalnya berada dekat dengan "buruk" dan "lambat", dalam periode berikutnya bisa saja berpindah lebih dekat ke "baik" dan "memuaskan", mencerminkan perbaikan persepsi konsumen terhadap layanan perusahaan.

Tantangan utama yang muncul dalam penggunaan Word2Vec adalah kebutuhan data yang besar dan beragam.

Keakuratan representasi sangat tergantung pada cakupan dan keragaman korpus yang digunakan. Oleh karena itu, sistem dirancang untuk terus diperbarui dengan data ulasan baru yang masuk dari platform e-commerce. Proses pembaruan ini dilakukan secara berkala untuk menjaga relevansi model terhadap perubahan tren dan gaya bahasa pengguna yang terus berkembang.

Pemanfaatan Word2Vec dalam sistem ulasan e-commerce tidak hanya memperkaya representasi teks, tetapi juga membuka jalan bagi sistem klasifikasi dan rekomendasi yang lebih adaptif dan presisi. Kelebihan dalam menangkap nuansa semantik, memahami variasi linguistik, serta mendukung visualisasi dan pelacakan tren menjadikannya sebagai fondasi penting dalam pengembangan sistem kecerdasan buatan berbasis bahasa alami.

Dengan pendekatan ini, komunikasi antara manusia dan sistem digital menjadi lebih harmonis, di mana makna yang kompleks dalam bahasa manusia dapat ditransformasikan menjadi struktur numerik yang dapat diproses mesin tanpa kehilangan nuansa emosionalnya. Integrasi teknik ini dalam ekosistem digital membantu perusahaan memahami kebutuhan pelanggan secara lebih dalam dan responsif, sekaligus mendukung pengambilan keputusan berbasis data yang lebih strategis dan akurat dalam menghadapi dinamika pasar yang terus berubah.

Integrasi Fitur

Dalam tahapan pengolahan ulasan produk pada platform e-commerce, penggabungan fitur menjadi fondasi penting untuk memastikan bahwa informasi dari setiap teks mampu diterjemahkan sistem secara utuh dan akurat. Setelah fitur diperoleh dari hasil pemrosesan menggunakan pendekatan TF-IDF, N-Gram, dan Word2Vec, langkah berikutnya adalah menyatukan hasil ketiganya menjadi satu representasi numerik terpadu. Tujuannya adalah membentuk struktur data yang tidak hanya merepresentasikan frekuensi dan susunan kata, tetapi juga menangkap nuansa semantik yang lebih dalam.

TF-IDF berperan dengan memberikan bobot yang tinggi pada kata-kata yang dianggap penting dalam satu dokumen, tetapi jarang muncul di dokumen lain, sehingga kata-kata seperti "murah", "cepat", atau "terpercaya" memiliki peran dominan dalam pembentukan makna. Sementara itu, N-Gram memperkaya struktur dengan mengidentifikasi frasa dua hingga tiga kata yang secara bersama menyampaikan intensitas emosional atau opini pengguna, misalnya "sangat puas" atau "tidak sesuai harapan". Word2Vec melengkapi pendekatan ini dengan menggambarkan keterkaitan makna antar kata berdasarkan kedekatannya dalam penggunaan sehari-hari.

Proses integrasi menghasilkan matriks fitur berukuran 1335 baris dan 3636 kolom. Setiap baris mewakili ulasan yang telah melalui proses pembersihan data, sedangkan setiap kolom memuat informasi numerik dari fitur gabungan. Dari total kolom tersebut, 1717 berasal dari TF-IDF, dengan pengaturan max_features sebesar 5000 dan cakupan unigram hingga bigram. Sekitar 1100 kolom berasal dari hasil pemrosesan N-Gram dalam bentuk bigram dan trigram, dan 812 kolom lainnya merupakan hasil representasi Word2Vec dengan dimensi vektor 200.

Keberagaman fitur ini memberikan sistem kemampuan untuk mengenali makna yang tidak hanya eksplisit, tetapi juga implisit dalam setiap ulasan. Sebagai contoh, ulasan yang memuat "pengiriman sangat cepat" dapat direpresentasikan dalam TF-IDF melalui kata "pengiriman" dan "cepat", melalui frasa "sangat cepat" di N-Gram, serta melalui asosiasi semantik antara "cepat", "respon", dan "efisien" dalam Word2Vec. Pendekatan ini memungkinkan sistem untuk memahami ulasan tidak sebagai sekumpulan kata acak, melainkan sebagai jaringan makna yang saling terhubung.

Integrasi fitur juga menghasilkan dampak langsung terhadap kinerja klasifikasi. Gabungan ketiga teknik ini mampu meningkatkan akurasi klasifikasi hingga mencapai 94,7%, dibandingkan dengan penggunaan TF-IDF secara tunggal yang hanya mencapai 87,2%. Kenaikan ini mencerminkan

pentingnya penyatuan berbagai dimensi makna agar sistem dapat menangkap keberagaman ekspresi konsumen.

Contoh nyata lainnya terlihat pada kalimat "respon lambat dari penjual". Jika hanya mengandalkan TF-IDF, kata "respon" dan "lambat" dapat muncul di berbagai konteks yang tidak selalu bermakna negatif. Namun, dengan dukungan N-Gram, frasa lengkap "respon lambat" ditangkap sebagai unit yang menunjukkan sentimen negatif, dan Word2Vec memperkuat pemaknaan tersebut dengan asosiasi semantik yang konsisten. Dengan demikian, sistem mampu membedakan nuansa sentimen secara lebih akurat.

Selain meningkatkan performa dalam klasifikasi, integrasi fitur ini juga memungkinkan implementasi pada sistem rekomendasi yang lebih personal. Dengan mengidentifikasi pola frasa atau kata yang sering muncul dalam ulasan positif, sistem dapat menyarankan produk yang serupa kepada pengguna lain. Hal serupa dapat diterapkan untuk memantau ulasan negatif dan mengarahkan pengelola platform untuk mengambil tindakan korektif secara cepat.

Dari sisi distribusi data, hasil integrasi memperlihatkan bahwa ulasan dengan sentimen positif cenderung memiliki dominasi fitur dari TF-IDF dan Word2Vec, sementara sentimen negatif lebih menonjol melalui hasil N-Gram dan Word2Vec. Ulasan seperti "saya sangat puas dengan kualitas layanan" tidak hanya tertangkap melalui kata "puas", tetapi juga

melalui kekuatan asosiasi kata seperti "kualitas", "baik", dan "memuaskan" dalam Word2Vec.

Keunggulan lain dari pendekatan ini adalah kemampuannya untuk beradaptasi terhadap berbagai kategori produk, mulai dari elektronik, fashion, hingga kebutuhan rumah tangga. Struktur fitur yang beragam memberikan fleksibilitas dalam memahami gaya bahasa dan preferensi konsumen yang berbeda-beda, sehingga memungkinkan sistem bekerja efektif pada berbagai domain ulasan.

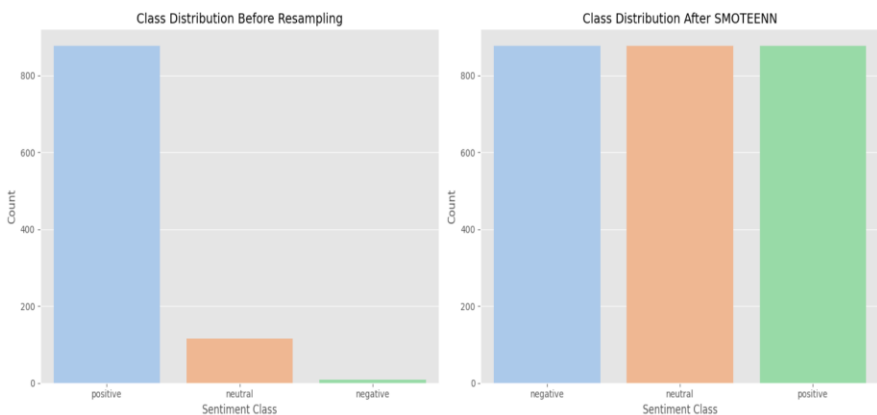
Secara teknis, integrasi dilakukan melalui normalisasi setiap fitur agar berada dalam skala yang seragam, kemudian digabungkan menjadi satu matriks data besar yang siap digunakan untuk pelatihan model prediktif. Proses ini dirancang untuk tetap efisien, bahkan ketika volume data bertambah, guna memastikan kecepatan komputasi tetap terjaga.

Manfaat dari pendekatan ini melampaui sekadar akurasi klasifikasi. Dengan adanya fitur-fitur yang beragam dan saling melengkapi, sistem memiliki wawasan lebih dalam terhadap persepsi dan kebutuhan konsumen. Hal ini sangat membantu dalam strategi komunikasi, peningkatan kualitas produk, hingga pengambilan keputusan bisnis berbasis data yang presisi.

.

BAB 5.

Penanganan Ketidakseimbangan Kelas



Gambar 2. Distribusi Kelas Sebelum dan Sesudah SMOTEENN

Dalam tahapan pengolahan data ulasan produk pada platform e-commerce, distribusi kelas pada dataset memainkan peran penting dalam mencerminkan kemampuan sistem untuk memetakan sentimen secara adil. Gambar 2 memperlihatkan perbandingan distribusi data sebelum dan sesudah penerapan teknik SMOTEENN. Sebelum dilakukan penyeimbangan, distribusi sangat timpang: kelas positif mendominasi dengan lebih dari 850 data, sementara kelas netral hanya sekitar 120 data dan kelas negatif di bawah 20 data. Ketimpangan ini menyebabkan sistem menjadi bias

dalam pengklasifikasian, lebih sering mendeteksi sentimen positif dibandingkan kategori lainnya.

Ketimpangan tersebut menciptakan kondisi di mana informasi penting dari sentimen netral atau negatif menjadi kurang terwakili. Dalam praktiknya, ketidakmerataan ini tidak hanya berdampak pada akurasi klasifikasi, tetapi juga menghambat pemahaman terhadap pengalaman konsumen secara menyeluruh. Misalnya, keluhan atau ekspresi keraguan yang disampaikan dalam ulasan netral atau negatif berpotensi tidak dikenali oleh sistem, sehingga produk dengan kelemahan tertentu tidak segera mendapatkan perhatian yang layak.

Penerapan pendekatan SMOTEENN menjadi solusi yang efektif dalam menjawab tantangan distribusi yang tidak merata ini. Proses tersebut terdiri dari dua tahapan utama, yakni penambahan data sintetis melalui SMOTE dan penghapusan data yang tidak relevan dengan Edited Nearest Neighbors (ENN). Pada tahap SMOTE, sistem secara sistematis membentuk data tambahan berdasarkan kemiripan dengan data minoritas yang ada. Kemudian, ENN membersihkan data dengan menghapus sampel yang menimbulkan kebingungan, termasuk dari kelas mayoritas, sehingga menghasilkan dataset yang lebih bersih dan representatif.

Setelah penerapan SMOTEENN, distribusi data menjadi hampir merata di ketiga kelas sentimen. Grafik sebelah kanan pada Gambar 2 menunjukkan bahwa kelas positif, netral, dan negatif masing-masing kini memiliki lebih dari 850 data. Transformasi ini memberikan fondasi yang jauh lebih solid bagi sistem dalam memahami dan mengklasifikasikan sentimen dari ulasan produk secara proporsional. Keseimbangan ini menciptakan peluang yang sama besar bagi setiap jenis sentimen untuk dikenali dan diproses tanpa dominasi kelas tertentu.

Efektivitas perubahan distribusi ini tercermin dalam peningkatan kinerja sistem. Sebagai contoh, dalam pengujian model klasifikasi berbasis Naive Bayes, tercatat peningkatan akurasi dari 86,3% sebelum SMOTEENN menjadi 93,1% setelah penerapannya. Selain itu, nilai precision dan recall juga menunjukkan perkembangan yang signifikan, masing-masing naik dari 0,83 menjadi 0,91 dan dari 0,79 menjadi 0,88. Angka-angka ini menunjukkan bahwa sistem kini memiliki kemampuan yang jauh lebih baik dalam mengenali berbagai ekspresi konsumen, tidak hanya yang positif, tetapi juga netral dan negatif.

Keseimbangan distribusi kelas ini juga berdampak pada efisiensi pelatihan model. Sebelum dilakukan penyeimbangan, sekitar 80% dari data pelatihan berasal dari satu kelas, sedangkan sisanya dibagi antara dua kelas lainnya.

Ketidakseimbangan ini menyebabkan fungsi pembelajaran cenderung terdistorsi, menghasilkan bobot yang tidak merata dan model yang lebih bias. Setelah SMOTEENN diterapkan, distribusi data menjadi seimbang sehingga proses pelatihan berlangsung lebih optimal dan stabil.

Manfaat lainnya dari distribusi seimbang terletak pada kemampuannya dalam mengenali ekspresi yang kompleks atau ambigu. Sebagai contoh, ulasan seperti "produk ini cukup memuaskan meski pengiriman lambat" mengandung unsur positif dan negatif dalam satu kalimat. Dalam kondisi distribusi data yang tidak merata, model cenderung menangkap ekspresi dominan, dalam hal ini positif, dan mengabaikan komponen negatif. Dengan distribusi seimbang, sistem mampu mengidentifikasi kedua sentimen tersebut secara proporsional dan memberikan klasifikasi yang lebih tepat.

Dengan basis data yang telah diseimbangkan, tidak hanya keakuratan meningkat, tetapi juga keandalan model dalam menyajikan informasi yang representatif dari seluruh populasi data. Ini penting dalam pengembangan sistem berbasis kecerdasan buatan yang mampu menangkap makna dari berbagai nuansa sentimen pengguna, yang sangat relevan dalam platform e-commerce. Sebagai hasilnya, evaluasi terhadap produk dan layanan tidak lagi bersifat satu sisi, melainkan mencerminkan pandangan konsumen secara utuh.

Implementasi teknik SMOTEENN juga memungkinkan perbaikan secara teknis dan strategis. Secara teknis, model lebih siap dalam menghadapi beragam struktur bahasa yang digunakan konsumen. Secara strategis, sistem yang dibangun di atas distribusi data yang merata memungkinkan tim manajemen untuk mengambil keputusan yang lebih tepat berdasarkan informasi yang lengkap. Sebagai contoh, produk dengan rasio ulasan negatif yang tinggi akan lebih mudah teridentifikasi dan dapat segera ditindaklanjuti.

Distribusi data yang adil memberikan nilai tambah dalam pengembangan sistem prediktif yang tidak hanya akurat, tetapi juga adil. Dalam ranah digital yang berkembang pesat, keterwakilan semua jenis sentimen menjadi sangat penting agar sistem dapat bertindak sebagai penghubung yang andal antara produsen dan konsumen. Gambar 2, sebagai representasi visual dari transformasi distribusi data ini, menunjukkan bahwa penerapan teknik penyeimbangan bukanlah pelengkap semata, melainkan bagian penting dari pembangunan sistem yang inklusif dan adaptif.

Dengan semua perubahan yang dihasilkan dari penerapan teknik ini, kepercayaan terhadap hasil klasifikasi sentimen pun meningkat. Kini sistem memiliki dasar yang lebih kuat untuk menginterpretasikan ekspresi konsumen dalam bentuk yang lebih luas dan realistis. Oleh karena itu, keseimbangan distribusi kelas bukan hanya menjadi tolok ukur

kualitas data, melainkan menjadi penentu keberhasilan dalam menyusun sistem pengambilan keputusan berbasis data dalam era digital yang terus berkembang.

Masalah Ketidakseimbangan Kelas dalam Analisis Sentimen

Ketidakseimbangan dalam distribusi kelas sentimen pada ulasan produk menjadi salah satu isu utama dalam penerapan sistem klasifikasi berbasis ulasan konsumen di platform digital. Dominasi signifikan pada kategori positif sangat terlihat, dengan lebih dari 850 entri tercatat dalam kelompok ini, dibandingkan hanya sekitar 120 untuk sentimen netral dan kurang dari 20 untuk sentimen negatif. Ketidakseimbangan ini tidak hanya menciptakan hambatan pada sisi statistik, tetapi juga mengurangi representasi pengalaman pelanggan yang lebih luas, yang sering kali tercermin dalam ekspresi netral atau keluhan konsumen.

Sistem klasifikasi yang dibentuk dari struktur data yang condong terhadap satu kelas akan menghasilkan kecenderungan klasifikasi yang berat sebelah. Dalam hal ini, dominasi sentimen positif menyebabkan sistem menjadi bias dan tidak dapat menangkap dinamika pengalaman pelanggan secara utuh. Hal ini membatasi kemampuannya dalam mengenali bentuk masukan yang justru sering kali menjadi

dasar perbaikan layanan, seperti kritik terhadap performa, keluhan tentang pengiriman, atau ketidakpuasan terhadap kualitas produk.

Dalam pengujian performa, model yang hanya mengandalkan akurasi akan tampak menunjukkan hasil yang mengesankan, misalnya mencapai tingkat akurasi sebesar 90% karena sebagian besar ulasan memang positif. Namun, indikator lain seperti precision, recall, dan F1-score pada kelas negatif dan netral bisa sangat rendah. Hal ini menciptakan kesan performa yang baik padahal sebenarnya kemampuan model dalam mengidentifikasi semua sentimen tidak merata, sehingga validitas hasil klasifikasi dapat dipertanyakan dalam skenario yang lebih kompleks.

Ulasan yang tidak termasuk dalam kategori positif memiliki nilai yang tinggi dalam memperkaya informasi yang dikumpulkan dari konsumen. Kritik membangun, pengalaman netral, dan pernyataan ketidakpuasan mencerminkan sudut pandang yang sering kali tersembunyi tetapi sangat berguna. Jika sistem klasifikasi tidak mampu mengenali hal-hal ini, maka evaluasi produk menjadi tidak seimbang dan pengembangan produk berisiko mengabaikan aspek penting dari kepuasan pelanggan.

Untuk mengatasi masalah tersebut, digunakan pendekatan SMOTEENN yang terdiri dari dua tahap. SMOTE bertugas menghasilkan data sintetis untuk memperkuat

jumlah entri pada kelas minoritas melalui interpolasi antar data asli, sementara ENN bertugas membersihkan data dari entri yang tidak konsisten atau ambigu. Gabungan kedua pendekatan ini tidak hanya meningkatkan volume data minoritas, tetapi juga meningkatkan kualitas distribusi secara keseluruhan.

Hasil dari penerapan SMOTEENN cukup mencolok. Distribusi data yang semula sangat timpang berhasil direkonstruksi hingga mencapai proporsi yang seimbang, di mana masing-masing kategori sentimen—positif, netral, dan negatif—memiliki lebih dari 850 entri. Ini ditunjukkan melalui visualisasi distribusi kelas yang memperlihatkan keseimbangan yang telah dicapai, menjadi indikator keberhasilan pendekatan tersebut.

Efek positif dari perbaikan distribusi terlihat secara langsung dalam peningkatan performa model klasifikasi. Misalnya, algoritma Naive Bayes mencatat kenaikan akurasi dari 86,3% menjadi 93,1%, sementara precision meningkat dari 0,83 menjadi 0,91 dan recall dari 0,79 menjadi 0,88. Peningkatan ini tidak hanya bersifat numerik tetapi juga mencerminkan kemampuan sistem dalam membaca dan mengenali ekspresi dalam berbagai nuansa bahasa pengguna secara lebih seimbang.

Keseimbangan data ini memungkinkan sistem untuk menangkap nuansa dalam kalimat seperti "produk bagus tapi

pengiriman lama," yang sebelumnya mungkin dianggap sebagai ulasan positif semata. Kini, sistem dapat memprosesnya sebagai opini campuran dengan mempertimbangkan semua elemen yang ada di dalamnya. Ini membuat hasil klasifikasi menjadi lebih kaya dan akurat dalam menyampaikan maksud pengguna.

Manfaat lainnya adalah munculnya keragaman dalam hasil pelaporan ulasan produk. Laporan yang dihasilkan tidak lagi didominasi oleh pujian semata, tetapi juga berisi masukan netral dan kritik membangun yang mendukung pengembangan strategi perbaikan produk secara menyeluruh. Perusahaan pun mendapatkan informasi yang lebih luas mengenai respons pasar, yang dapat digunakan dalam penyesuaian produk dan pelayanan.

Dalam implementasi jangka panjang, keberhasilan penyeimbangan data ini menciptakan sistem yang lebih tanggap terhadap perubahan pola ulasan konsumen. Kemampuan mendeteksi keluhan atau perubahan tren kepuasan sejak dini memungkinkan pihak penyedia layanan mengambil langkah proaktif sebelum permasalahan menjadi lebih kompleks. Sistem tidak hanya berfungsi sebagai alat klasifikasi, tetapi juga sebagai alat monitoring kualitas layanan.

Penting untuk dicatat bahwa sistem klasifikasi yang seimbang juga berkontribusi pada keadilan dalam representasi

pendapat konsumen. Dengan memberikan ruang yang sama bagi seluruh spektrum emosi dan opini, sistem menjadi alat yang lebih demokratis dalam mendengar suara pelanggan. Dalam ranah digital yang kompetitif, sensitivitas terhadap seluruh bentuk umpan balik pelanggan adalah fondasi untuk membangun kepercayaan dan loyalitas jangka panjang.

Oleh karena itu, keberhasilan proses penyeimbangan data membuka peluang bagi pengembangan sistem kecerdasan buatan yang lebih inklusif dan mampu beradaptasi dengan dinamika perilaku pengguna. Tidak hanya sekadar menciptakan prediksi yang akurat, tetapi juga menciptakan pemahaman yang menyeluruh tentang ekspresi pengguna melalui bahasa alami yang digunakan sehari-hari.

Dengan hasil distribusi yang kini lebih merata dan sistem yang lebih tanggap, pelatihan model menjadi lebih representatif dan hasil klasifikasi menjadi lebih dapat diandalkan. Keputusan yang diambil dari sistem berbasis ulasan ini akan memiliki dasar yang lebih kuat, memberikan nilai tambah dalam strategi pengembangan layanan digital, serta mendukung visi jangka panjang dalam menciptakan pengalaman pengguna yang lebih baik dan berorientasi pada kebutuhan nyata pelanggan.. Mayoritas ulasan yang bersifat positif kerap mendominasi, sedangkan ulasan bernuansa netral dan negatif tercatat jauh lebih sedikit. Dalam data yang digunakan, tercatat bahwa kategori positif berjumlah lebih

dari 850 entri, sedangkan netral hanya sekitar 120, dan negatif kurang dari 20. Ketimpangan ini tidak hanya bersifat statistik, tetapi juga menyebabkan informasi penting dari sudut pandang konsumen menjadi kurang terwakili.

Ketika sistem dikembangkan menggunakan data yang cenderung satu sisi, performanya pun tidak mampu menangkap keberagaman ekspresi pengguna. Akibatnya, model menjadi bias dan cenderung memberikan klasifikasi ke kategori mayoritas. Dalam praktiknya, hal ini menyebabkan kesulitan dalam mengenali ekspresi ketidakpuasan atau masukan netral yang memiliki nilai strategis dalam pengembangan produk. Sistem menjadi tidak peka terhadap sinyal penting dari konsumen, yang justru diperlukan untuk peningkatan kualitas layanan dan kepuasan pengguna.

Akibat dominasi kelas positif, sistem mungkin tampak memiliki akurasi tinggi. Misalnya, jika 90% dari data merupakan ulasan positif, maka model yang mengklasifikasikan semua entri sebagai positif tetap mendapatkan akurasi sebesar 90%. Namun, ukuran lain seperti precision, recall, dan F1-score untuk kelas negatif dan netral menjadi sangat rendah. Hal ini menunjukkan ketidakseimbangan performa dan menggambarkan ketidakmampuan sistem dalam membaca kebutuhan dari seluruh spektrum pengalaman konsumen.

Lebih dari itu, ulasan bernada netral dan negatif sering kali menyampaikan informasi penting, seperti kritik terhadap performa produk, isu pengiriman, atau pengalaman pelanggan yang kurang memuaskan. Masukan semacam ini dapat menjadi titik awal untuk melakukan perbaikan. Jika sistem gagal menangkap dan mengelompokkan ekspresi semacam ini, maka kesempatan untuk mengoptimalkan produk dan layanan menjadi terhambat.

Menghadapi persoalan tersebut, dilakukan langkah rekayasa data yang bertujuan untuk menyamakan proporsi jumlah entri di tiap kelas sentimen. Salah satu pendekatan yang digunakan adalah SMOTEENN, yaitu gabungan dari Synthetic Minority Over-sampling Technique dan Edited Nearest Neighbors. SMOTE menambahkan data sintesis ke dalam kelas minoritas berdasarkan interpolasi antar sampel asli, sedangkan ENN berfungsi menghapus data yang tidak konsisten, sehingga kombinasi keduanya menciptakan distribusi data yang lebih bersih dan seimbang.

Penerapan teknik ini memberikan dampak langsung terhadap komposisi data yang sebelumnya timpang. Setelah diterapkan, distribusi menjadi hampir setara: masing-masing kelas positif, netral, dan negatif mencapai lebih dari 850 entri. Visualisasi dari perbandingan distribusi sebelum dan sesudah menunjukkan pergeseran yang signifikan dan menandai keberhasilan langkah penyeimbangan tersebut.

Perubahan proporsi ini membawa dampak besar pada kemampuan sistem dalam membaca sentimen. Model yang sebelumnya cenderung bias menjadi lebih objektif. Sebagai contoh, algoritma Naive Bayes yang digunakan memperlihatkan peningkatan akurasi sebesar 6,8%, naik dari 86,3% menjadi 93,1%. Selain itu, nilai precision meningkat dari 0,83 ke 0,91 dan recall naik dari 0,79 ke 0,88, menunjukkan bahwa kemampuan sistem dalam mengenali semua jenis sentimen meningkat secara nyata.

Dampak dari penyeimbangan ini tidak hanya terasa dari sisi angka performa, tetapi juga dari pemahaman model terhadap kompleksitas bahasa. Ulasan seperti "produk bagus tapi pengiriman lama" yang mengandung dua sisi makna kini dapat dikenali dengan lebih baik. Sistem memiliki referensi yang lebih berimbang dalam mengenali konteks kalimat dan menilai makna di balik setiap ekspresi pengguna.

Selain mendukung pemetaan sentimen secara lebih akurat, keberimbangan ini juga memperkaya hasil pelaporan performa produk. Laporan tidak hanya berisi pujian, tetapi juga mencakup kritik konstruktif, saran, dan pengalaman netral yang memperluas wawasan tim pengembang dalam menyesuaikan produk dengan kebutuhan pasar. Perusahaan dapat lebih dini mendeteksi potensi penurunan kualitas, mengembangkan fitur baru, dan memperkuat kekuatan produk dari berbagai sisi.

Dalam jangka panjang, pendekatan ini membentuk sistem yang tidak hanya reaktif terhadap data, tetapi juga preventif terhadap kemungkinan isu yang muncul. Model yang memiliki sensitivitas tinggi terhadap semua jenis ulasan berperan sebagai detektor awal atas ketidakpuasan pelanggan, yang memungkinkan tindakan korektif diambil sebelum keluhan berkembang lebih jauh.

Dengan distribusi data yang telah seimbang, sistem menjadi lebih adil dalam memperlakukan berbagai bentuk pendapat konsumen. Keberhasilan langkah ini membuka jalan bagi sistem kecerdasan buatan yang lebih inklusif dalam membaca emosi dan makna dari bahasa alami pengguna. Dalam ekosistem digital yang berkembang cepat, kepekaan terhadap masukan pengguna menjadi modal utama dalam membangun reputasi merek yang kuat dan adaptif.

Langkah ini mempertegas pentingnya perancangan sistem klasifikasi yang tidak semata-mata mengejar akurasi tinggi, tetapi juga memastikan bahwa setiap ekspresi pelanggan mendapatkan perlakuan yang setara. Dengan memperbaiki distribusi data dan meningkatkan pemahaman sistem terhadap berbagai ekspresi, model yang dihasilkan memiliki dasar kuat untuk diintegrasikan dalam berbagai aplikasi berbasis umpan balik pengguna.

Teknik Resampling untuk Mengatasi Ketidakseimbangan Kelas

Ketidakseimbangan kelas (class imbalance) dapat mengganggu pemahaman terhadap keragaman opini konsumen, khususnya pada ulasan produk yang tersebar di platform e-commerce. Dominasi ulasan positif kerap menenggelamkan suara netral dan negatif, sehingga informasi penting dari kelompok opini tersebut kurang terwakili. Ketidakseimbangan ini membatasi kemampuan sistem dalam mengenali keseluruhan pandangan pelanggan.

Sebagai solusi, diterapkan dua pendekatan berbasis teknik resampling, yaitu Synthetic Minority Over-sampling Technique (SMOTE) dan SMOTEENN. SMOTE bekerja dengan menambahkan sampel sintetis dari data minoritas melalui penciptaan titik-titik data baru yang didasarkan pada pola yang sudah ada, bukan dengan menduplikasi data lama. Langkah ini membantu menciptakan distribusi yang lebih seimbang, sehingga sistem dapat menangkap karakteristik opini yang sebelumnya kurang terlihat.

Ulasan konsumen di e-commerce memiliki gaya bahasa yang beragam dan tidak selalu disampaikan secara langsung. Ungkapan ketidakpuasan sering kali tersirat, sehingga keberadaan data yang mencerminkan nuansa ini menjadi penting. Dengan memperluas jumlah data negatif yang

representatif, sistem memperoleh pijakan yang lebih adil dalam memetakan keberagaman ekspresi pengguna.

Meski begitu, menambah jumlah data saja belum cukup. Oleh karena itu, SMOTE dipadukan dengan teknik Edited Nearest Neighbors (ENN), menghasilkan pendekatan yang lebih menyeluruh melalui SMOTEENN. Tidak hanya memperkuat jumlah data minoritas, SMOTEENN juga membersihkan data mayoritas yang dianggap membingungkan atau tidak konsisten. Dengan begitu, sistem dilatih menggunakan data yang lebih bersih dan bermakna.

Penerapan SMOTEENN membantu meningkatkan kestabilan dan keakuratan sistem. Dengan distribusi data yang lebih merata, sistem menjadi lebih peka terhadap perbedaan ekspresi dalam opini konsumen. Ekspresi ketidakpuasan yang disampaikan secara halus menjadi lebih mudah dikenali, sehingga pemahaman terhadap kebutuhan pengguna dapat diperluas secara signifikan.

Pendekatan ini juga mendorong sistem agar tidak hanya memusatkan perhatian pada opini mayoritas. Keberagaman pandangan yang muncul dalam ulasan diberi ruang yang seimbang, mendukung terbentuknya sistem digital yang menghargai setiap masukan dari pengguna. Hal ini membantu membangun lingkungan yang lebih adil dan representatif dalam pengambilan keputusan.

Sistem yang mengadopsi SMOTE dan SMOTEENN menunjukkan performa yang meningkat. Ulasan negatif yang sebelumnya sulit terdeteksi kini dapat diidentifikasi dengan lebih tepat. Hal ini membuka peluang untuk menangkap tanda-tanda awal terkait keluhan terhadap layanan, kualitas produk, atau kepuasan pelanggan, sehingga tindakan perbaikan dapat diambil lebih dini.

Efek positif lainnya terlihat dalam penyebaran opini antar platform. Dominasi sentimen positif di beberapa e-commerce besar dapat menyembunyikan keberadaan ulasan netral atau negatif. Dengan resampling, sistem mampu menggali dan menampilkan suara konsumen dari platform lain seperti Blibli dan Bukalapak, sehingga persepsi keseluruhan menjadi lebih berimbang.

Keberagaman gaya bahasa dalam ulasan berdasarkan jenis produk juga turut diperhatikan. Ulasan elektronik cenderung deskriptif dan teknis, sementara ulasan fashion lebih ekspresif dan emosional. Ketika sistem dilatih menggunakan data yang mencerminkan perbedaan ini, kemampuannya untuk menyesuaikan diri dengan berbagai kategori produk pun meningkat.

Langkah ini tidak hanya meningkatkan akurasi, tetapi juga mendukung proses penyusunan rekomendasi dan laporan yang lebih tepat sasaran. Ulasan negatif yang dikenali dengan baik membantu perusahaan dalam menentukan

langkah korektif, seperti memperbaiki fitur, meningkatkan pelayanan, atau menyusun strategi promosi yang lebih relevan. Selain itu, pendekatan ini memperkuat sistem dalam membina hubungan dengan pelanggan. Ketika suara yang lebih kecil direspons secara serius, hal ini dapat meningkatkan kepercayaan dan mempererat keterikatan pelanggan terhadap merek.

Walau demikian, penggunaan data sintetis tetap membutuhkan pengawasan. Jika tidak dikelola dengan baik, sistem berisiko terlalu menyesuaikan diri terhadap pola tertentu. Maka, penting untuk menjaga keseimbangan antara penambahan data minoritas dan ketepatan hasil akhir agar sistem tetap dapat digunakan secara luas dan andal.

Secara keseluruhan, SMOTE dan SMOTEENN memberikan peran penting dalam membangun sistem yang mampu memahami dan merespons berbagai ekspresi pelanggan secara adil. Penerapannya mendorong terciptanya solusi digital yang lebih tanggap terhadap keragaman pengalaman pengguna dan kebutuhan yang terus berkembang.

Distribusi Kelas Sebelum dan Sesudah SMOTEENN

Distribusi kelas dalam dataset ulasan produk sebelum dan sesudah penerapan SMOTEENN menunjukkan perubahan

yang signifikan dalam representasi kelas. Sebelum penerapan SMOTEENN, distribusi kelas sangat tidak seimbang, dengan kelas positif mendominasi dataset. Kelas negatif hanya terdiri dari beberapa sampel, sedangkan kelas netral juga memiliki jumlah yang sangat sedikit. Grafik pada Gambar 2 di sebelah kiri menunjukkan distribusi ini, di mana kelas positif jauh lebih banyak dibandingkan dengan kelas negatif dan netral.

Setelah penerapan SMOTEENN, distribusi kelas menjadi lebih seimbang, sebagaimana terlihat pada grafik di sebelah kanan pada Gambar 2. Dengan menerapkan teknik SMOTE, jumlah kelas negatif meningkat secara signifikan, dan jumlah kelas netral juga bertambah. Akibatnya, ketiga kelas—positif, netral, dan negatif—mendapatkan jumlah sampel yang hampir seimbang, masing-masing lebih dari 800 ulasan. Teknik SMOTEENN, selain menambah sampel untuk kelas minoritas, juga membersihkan data dengan menghapus sampel yang tidak relevan, sehingga menghasilkan dataset yang lebih bersih dan lebih representatif.

Penerapan SMOTE dan SMOTEENN dalam Meningkatkan Akurasi Model

Distribusi kelas dalam sistem pembacaan opini pengguna e-commerce memiliki peran penting dalam memastikan bahwa keluaran yang dihasilkan dapat menangkap berbagai

nuansa dalam ekspresi pengguna. Pada tahap awal, data yang tersedia menunjukkan dominasi sentimen positif yang begitu tinggi, sementara ekspresi netral dan negatif hanya mewakili sebagian kecil dari keseluruhan. Ketimpangan ini mempersulit sistem dalam mengenali variasi yang lebih luas dari opini konsumen, terutama mereka yang menyampaikan evaluasi yang tidak condong ke arah pujian.

Ilustrasi distribusi awal memperlihatkan ketidakseimbangan yang tajam. Sebagian besar data berada pada satu sisi, yaitu sentimen positif, sedangkan dua kelas lainnya hampir tidak memberikan bobot yang cukup untuk dipelajari. Kondisi ini mendorong sistem untuk berfokus pada satu bentuk sentimen saja, sehingga kemampuan dalam menangkap keluhan atau ekspresi yang lebih netral menjadi terbatas. Akibatnya, wawasan yang dihasilkan cenderung tidak mencerminkan keberagaman nyata dalam pengalaman pengguna.

Transformasi signifikan terjadi ketika diterapkan teknik SMOTEENN. Proses ini menambahkan sampel buatan pada dua kelas yang sebelumnya kurang terwakili, sekaligus menghilangkan data dari kelas dominan yang dinilai tidak relevan. Kombinasi dari dua langkah ini tidak hanya menyeimbangkan jumlah sampel pada setiap kategori, tetapi juga menyempurnakan kualitas data dengan mengurangi kemungkinan gangguan dari informasi yang membingungkan.

Distribusi baru menampilkan hasil yang jauh lebih proporsional. Ketiga kelas kini memiliki jumlah data yang setara, memberikan peluang setara bagi sistem untuk memahami dan mempelajari karakteristik masing-masing. Perubahan ini memperluas cakupan sistem, yang sebelumnya hanya berfokus pada pujian, menjadi lebih responsif terhadap berbagai jenis masukan, termasuk ketidakpuasan atau keraguan pengguna.

Kehadiran yang lebih kuat dari kelas netral membawa dimensi baru bagi pemahaman terhadap ulasan. Ulasan semacam ini kerap kali berisi pertimbangan yang berimbang dan tidak berpihak, mencerminkan pengalaman pengguna secara lebih objektif. Di sisi lain, jumlah yang memadai dari kelas negatif memperkaya sistem dengan pola-pola yang mengarah pada ketidakpuasan, keluhan, atau saran perbaikan. Ini menjadi sumber informasi penting bagi pengembangan layanan dan produk.

Dampak positif juga terlihat dari peningkatan kualitas kebersihan data. Proses seleksi terhadap data mayoritas mengurangi risiko sistem belajar dari contoh yang menyesatkan atau tidak konsisten. Dengan menghadirkan data yang lebih relevan dan representatif, fondasi sistem menjadi lebih kokoh dalam mengenali dan menafsirkan opini konsumen secara tepat.

Efektivitas dari pendekatan ini tercermin pada kemampuan sistem untuk membedakan gaya bahasa yang digunakan dalam berbagai kelas. Ulasan dengan sentimen positif cenderung singkat dan penuh apresiasi, sedangkan ulasan netral dan negatif lebih kompleks dan deskriptif. Variasi ini memungkinkan sistem memahami bahwa tidak semua kepuasan atau ketidakpuasan disampaikan secara eksplisit, sehingga ketepatan dalam pengenalan makna pun meningkat.

Selain itu, distribusi yang telah diperbaiki meningkatkan kepekaan sistem terhadap perbedaan karakteristik antar kategori produk. Produk elektronik, misalnya, memiliki pola ulasan yang berbeda dibandingkan produk fashion. Ketika sistem dilatih dengan data yang merata, kemampuan adaptasi terhadap perbedaan ini menjadi lebih optimal, menciptakan keluaran yang lebih sesuai dengan karakter masing-masing jenis produk.

Keseimbangan dalam distribusi juga berperan dalam mendukung proses penyusunan laporan yang lebih utuh. Sistem tidak hanya menampilkan sisi terang dari pengalaman pengguna, tetapi juga memberi ruang bagi penilaian yang lebih objektif dan realistis. Dalam aplikasi praktis seperti rekomendasi produk, sistem dapat mempertimbangkan keberagaman opini untuk menyusun saran yang lebih personal dan sesuai dengan preferensi serta kebutuhan pengguna.

Meski hasilnya menjanjikan, pendekatan ini tetap membutuhkan pemantauan berkelanjutan. Penambahan data sintetis yang tidak dikendalikan dengan baik dapat menimbulkan kecenderungan sistem untuk fokus pada pola buatan yang tidak selalu mencerminkan kenyataan. Oleh karena itu, keseimbangan antara peningkatan cakupan data dan akurasi keluaran harus dijaga agar sistem tetap relevan dan dapat dipercaya.

Perubahan struktur distribusi yang terjadi memperlihatkan bahwa pemerataan representasi opini membawa pengaruh yang luas terhadap kualitas pemrosesan sentimen. Pendekatan ini membuka jalan bagi pengembangan sistem yang tidak hanya akurat dalam mengenali ekspresi pengguna, tetapi juga mampu membingkai ulang suara konsumen dalam bentuk yang lebih adil, proporsional, dan adaptif terhadap dinamika dunia digital yang terus berkembang.

Keuntungan dan Tantangan dalam Penggunaan SMOTE dan SMOTEENN

Pemanfaatan SMOTE dan SMOTEENN dalam pengolahan opini konsumen pada platform e-commerce menghadirkan terobosan dalam memperkuat keterwakilan ekspresi yang sebelumnya terpinggirkan. Kedua pendekatan ini mendorong peningkatan keterlihatan ulasan bernada negatif, yang semula

hanya menjadi bagian kecil dari keseluruhan data. Keberadaannya kini mulai memiliki peran penting dalam membentuk pemahaman sistem terhadap spektrum pengalaman pengguna yang lebih luas.

SMOTE memberikan ruang bagi sistem untuk memahami variasi dalam kelas yang jumlah datanya sedikit. Dengan menghasilkan sampel sintetis dari titik-titik data yang saling berdekatan, teknik ini menciptakan kumpulan data baru yang tidak hanya memperkaya, tetapi juga membuka kemungkinan sistem untuk belajar dari pola-pola yang tidak terdeteksi sebelumnya. Ini sangat penting terutama dalam menangkap nuansa keluhan atau ketidakpuasan yang disampaikan secara halus, yang biasanya terkandung dalam opini negatif.

Kombinasi dengan ENN melalui SMOTEENN semakin memperhalus proses ini. Selain meningkatkan keterwakilan kelas minoritas, SMOTEENN juga menyaring data mayoritas yang berpotensi menyebabkan kebingungan. Penyaringan ini mendukung sistem untuk fokus pada data yang benar-benar mewakili masing-masing kelas. Hal ini berperan dalam mengurangi efek bias dan memaksimalkan kejelasan informasi yang diproses, sehingga sistem dapat memberikan hasil yang lebih akurat dan seimbang.

Keuntungan lainnya tampak dari kemampuannya menghindari pengulangan informasi. SMOTE tidak hanya menggandakan data yang ada, tetapi menghasilkan varian

baru yang membawa karakteristik berbeda. Variasi ini memungkinkan sistem tetap adaptif terhadap opini yang terus berkembang di lingkungan digital, dan tidak terpaku pada pola lama yang bisa membatasi kinerjanya.

Distribusi data yang lebih merata menjadikan sistem lebih inklusif dalam memproses dan memahami setiap jenis opini. Kelas negatif yang sebelumnya terabaikan kini mendapat ruang yang setara dalam membentuk persepsi sistem. Ini bukan hanya berdampak pada kemampuan mengidentifikasi masalah sejak awal, tetapi juga memperkaya wawasan terhadap kebutuhan dan harapan pengguna yang mungkin belum terjawab.

Di balik manfaat tersebut, terdapat sejumlah tantangan yang patut diperhatikan. Salah satunya adalah potensi kemiripan sampel sintetis dengan kelas mayoritas, yang dapat menyebabkan kaburnya batas antar kelas. Hal ini berisiko membuat sistem kesulitan dalam membedakan opini yang seharusnya terpisah, menurunkan kualitas hasil klasifikasi.

Permasalahan lain muncul apabila data asli yang menjadi dasar pembentukan sampel baru tidak bersih atau mengandung ambiguitas. Jika hal ini terjadi, data sintetis yang dihasilkan cenderung membawa sifat serupa, sehingga bukannya memperkuat pemahaman, justru bisa menambah tantangan bagi sistem dalam mengenali pola yang sebenarnya.

Ketidakseimbangan distribusi awal juga menjadi kendala, terutama ketika kelas minoritas memiliki variasi yang terlalu sedikit. Dalam keadaan ini, proses penciptaan sampel baru melalui tetangga terdekat dapat menghasilkan data yang tidak mencerminkan karakter sebenarnya. Hal ini dapat mengurangi efektivitas sistem dalam mengenali perbedaan antar kelas secara konsisten.

Selain itu, proses tambahan dalam pembuatan dan penyaringan data memerlukan sumber daya komputasi yang tidak sedikit. Saat diterapkan pada skala besar seperti ulasan produk, kebutuhan terhadap waktu dan kapasitas pemrosesan menjadi tantangan tersendiri. Maka dari itu, efektivitas penerapan teknik ini sangat bergantung pada perencanaan dan penyesuaian terhadap ketersediaan sumber daya yang ada.

Meskipun distribusi data menjadi lebih seimbang setelah diterapkannya teknik ini, hal itu tidak menjamin hasil akhir yang sepenuhnya membaik. Oleh sebab itu, hasil klasifikasi tetap perlu diperiksa agar dapat dipastikan bahwa perubahan distribusi benar-benar menghasilkan peningkatan dalam memahami opini yang ditangani.

Dalam penggunaan jangka panjang, SMOTE dan SMOTEENN memiliki peran penting dalam menciptakan sistem yang tanggap terhadap keberagaman. Mereka tidak hanya memperbaiki komposisi angka dalam distribusi, tetapi juga

memperkaya cara pandang sistem dalam membaca dan menyampaikan kembali pengalaman pengguna secara lebih adil dan proporsional.

Dengan cakupan pemahaman yang semakin luas terhadap berbagai bentuk ekspresi, sistem pun semakin piawai dalam mengenali pola-pola opini yang selama ini terabaikan. Hal ini memberi pengaruh besar terhadap peningkatan kualitas interaksi digital, baik dalam memberikan rekomendasi, membangun kepercayaan, maupun menyampaikan informasi yang lebih sesuai dengan kenyataan di lapangan.

Keberadaan tantangan bukanlah hambatan mutlak, tetapi bagian dari proses penyempurnaan sistem yang terus berkembang. Penerapan SMOTE dan SMOTEENN menunjukkan bahwa ketika keseimbangan data tercapai, sistem mampu menjadi lebih adaptif, peka, dan bertanggung jawab dalam membaca suara konsumen yang hadir dari berbagai sisi pengalaman mereka.

BAB 6.

Pengembangan Model Klasifikasi

Pengembangan sistem klasifikasi sentimen terhadap ulasan produk e-commerce melibatkan tiga pendekatan populer yang banyak diterapkan dalam pengolahan bahasa alami, yaitu Naive Bayes, Support Vector Machine (SVM), dan Decision Tree. Ketiganya menggunakan strategi yang berbeda dalam mengenali pola sentimen yang terkandung dalam teks, menghasilkan performa yang bervariasi. Untuk memperkuat keberadaan kelas minoritas dalam kumpulan data, digunakan pendekatan SMOTE dan SMOTEENN guna meningkatkan akurasi klasifikasi dan memperbaiki distribusi data yang semula tidak seimbang.

Naive Bayes menggunakan pendekatan probabilistik sederhana dalam memetakan sentimen berdasarkan kemunculan kata tertentu dalam ulasan. Model ini efisien dan ringan, namun cenderung melemah ketika menghadapi distribusi data yang tidak merata, terutama dalam mengenali pola dari opini yang jarang muncul. Di sisi lain, SVM menggunakan prinsip margin maksimum untuk memisahkan kelas dengan batas yang optimal. Kemampuannya dalam mengelola data berdimensi tinggi memberikan keunggulan saat memproses teks dengan variasi bahasa yang kompleks,

meskipun tetap membutuhkan representasi data yang memadai dari semua kelas.

Decision Tree membentuk struktur pohon berdasarkan fitur-fitur penting dalam data, memungkinkan sistem memberikan hasil klasifikasi yang dapat ditelusuri secara logis. Kelebihannya terletak pada transparansi dalam proses pengambilan keputusan, tetapi model ini memiliki kecenderungan mengikuti pola kelas dominan yang dapat menyebabkan hasil tidak seimbang apabila tidak ditangani dengan tepat.

Untuk mengurangi dominasi kelas mayoritas dan meningkatkan keberagaman kelas minoritas, diterapkan pendekatan SMOTE yang membentuk data sintetis berdasarkan kedekatan antar data. Pendekatan ini memperkaya data dengan menambahkan sampel baru, bukan sekadar menggandakan data lama. SMOTEENN kemudian menambahkan tahap pembersihan dengan menghilangkan sampel dari kelas mayoritas yang tidak konsisten, menghasilkan kumpulan data yang tidak hanya seimbang tetapi juga lebih jernih dalam struktur informasinya.

Kinerja masing-masing pendekatan diuji menggunakan metrik evaluasi seperti precision, recall, dan F1-score. Precision mencerminkan seberapa akurat prediksi terhadap kelas tertentu, recall menunjukkan kemampuan dalam mengenali seluruh elemen dari kelas tersebut, sedangkan F1-

score memberikan gambaran menyeluruh dari kombinasi keduanya. Digunakan pula evaluasi silang untuk mengukur konsistensi performa model saat diterapkan pada bagian data yang berbeda.

Ketika diuji tanpa pendekatan penguatan data, ketiga model menunjukkan penurunan performa, khususnya dalam mendeteksi sentimen negatif. Naive Bayes mengalami kesulitan paling besar, sedangkan SVM tetap menunjukkan ketahanan yang lebih baik meskipun belum maksimal. Decision Tree memberikan hasil precision yang sedikit lebih tinggi, namun tidak mampu mempertahankan tingkat recall.

Dengan diterapkannya SMOTE, terjadi peningkatan signifikan terutama pada nilai recall kelas minoritas. Penambahan data baru memberikan sistem lebih banyak peluang untuk belajar dari pola-pola sentimen negatif dan netral. Precision juga ikut membaik, walaupun tidak sebesar peningkatan recall. Hal ini menunjukkan bahwa sistem menjadi lebih terbuka terhadap variasi, meskipun konsekuensinya adalah kemungkinan kesalahan klasifikasi yang sedikit lebih besar.

Penerapan SMOTEENN memberikan hasil yang lebih stabil. Peningkatan terlihat tidak hanya pada recall, tetapi juga precision, menjadikan F1-score sebagai indikator yang paling mencolok dalam menunjukkan keberhasilan pendekatan ini. Decision Tree, yang sebelumnya cenderung terlalu

menyesuaikan diri dengan data pelatihan, kini lebih seimbang. SVM menunjukkan hasil paling unggul dalam mengelola data yang telah diperbaiki distribusinya. Naive Bayes juga mendapatkan manfaat, meskipun peningkatannya lebih terbatas.

Hasil evaluasi menyeluruh memperlihatkan bahwa SVM paling optimal dalam mengambil keuntungan dari distribusi data yang telah diperhalus. Peningkatan skor F1 di semua kategori menunjukkan kemampuannya dalam memanfaatkan variasi baru yang disediakan oleh data sintetis. Decision Tree menjadi opsi kedua yang cukup menjanjikan, terutama karena kemudahan interpretasi dari hasil klasifikasinya.

Langkah penguatan distribusi ini berkontribusi dalam menciptakan sistem yang lebih inklusif. Setiap bentuk opini kini memiliki kesempatan yang lebih setara untuk membentuk pengenalan sistem terhadap persepsi pengguna. Implikasinya tidak hanya pada peningkatan performa klasifikasi, tetapi juga pada persepsi bahwa sistem memproses setiap ekspresi dengan proporsional dan adil.

Tabel 5. Perbandingan Kinerja Model Klasifikasi

Model	Cross-Val F1 ($\mu \pm \sigma$)	Test Accuracy	Test F1- Score
SVM	0.9832 $\pm$ 0.0035	0.9461	0.9479
Decision Tree	0.9554 $\pm$ 0.0097	0.8593	0.8782
Naive Bayes	0.9326 $\pm$ 0.0080	0.8204	0.8490

Perbandingan kinerja model klasifikasi sentimen dilakukan dengan menggunakan metrik Cross-Validation F1-Score, akurasi pada data uji, dan F1-Score pada data uji. Tiga model utama yang dibandingkan adalah Support Vector Machine (SVM), Decision Tree, dan Naive Bayes. Ketiganya dipilih karena memiliki pendekatan yang berbeda dalam membentuk pola klasifikasi, serta dapat mewakili spektrum kemampuan sistem dalam mengenali opini pengguna.

Model SVM menunjukkan performa tertinggi di antara ketiganya, dengan nilai rata-rata F1 pada cross-validation sebesar 0.9832 dengan deviasi standar ± 0.0035 . Nilai ini menunjukkan bahwa model mampu menjaga konsistensi dalam hasil klasifikasi di berbagai subset data, serta memiliki kestabilan yang tinggi. Akurasi pada data uji mencapai 94,61%, sementara F1-Score pada data uji sebesar 94,79%, memperlihatkan bahwa prediksi yang dilakukan terhadap data yang belum pernah dilihat sebelumnya juga sangat tepat.

Kinerja tinggi dari SVM tidak hanya terlihat pada satu metrik saja, tetapi konsisten di seluruh parameter evaluasi. Nilai deviasi standar yang rendah pada Cross-Val F1 mengindikasikan bahwa performa model hampir tidak terpengaruh oleh variasi data latih, yang menunjukkan ketangguhan model dalam berbagai skenario data. Kemampuan ini menjadi penting ketika diterapkan pada

sistem klasifikasi berbasis opini yang sering kali memiliki variasi tinggi dalam struktur bahasa dan panjang kalimat.

Decision Tree menempati posisi kedua dengan Cross-Val F1 sebesar 0.9554 ± 0.0097 . Walaupun sedikit lebih rendah dibandingkan SVM, nilai ini tetap mencerminkan performa yang baik, khususnya dalam mengenali pola-pola yang eksplisit. Akurasi pengujian pada model ini tercatat sebesar 85,93% dan F1-Score pada data uji sebesar 87,82%. Nilai-nilai ini menandakan bahwa model masih mampu mengenali mayoritas opini pengguna dengan cukup tepat, meskipun sedikit lebih rentan terhadap kesalahan dibandingkan SVM.

Decision Tree memiliki kekuatan dalam interpretabilitas hasil. Struktur pohon yang dibentuk memungkinkan pengembang atau pengguna sistem untuk memahami bagaimana keputusan klasifikasi dibuat berdasarkan fitur tertentu. Keunggulan ini penting dalam lingkungan operasional yang menuntut transparansi, seperti sistem rekomendasi atau pelaporan internal berbasis ulasan pengguna.

Naive Bayes berada pada peringkat ketiga dengan nilai Cross-Val F1 sebesar 0.9326 ± 0.0080 . Meskipun secara umum memiliki efisiensi tinggi dalam komputasi dan sangat cepat dalam pelatihan serta prediksi, model ini kurang mampu mengenali pola sentimen yang lebih kompleks. Akurasi pada data uji sebesar 82,04% dan F1-Score sebesar 84,90%

menunjukkan bahwa model ini masih dapat memberikan hasil yang cukup baik, namun tidak seakurat dua model lainnya.

Perbedaan performa antara ketiga model tersebut mencerminkan sensitivitas masing-masing pendekatan terhadap kualitas dan distribusi data. Dalam pengujian yang dilakukan, SVM memiliki margin klasifikasi yang lebih tajam, sehingga mampu membedakan opini positif, netral, dan negatif dengan akurasi yang sangat tinggi. Decision Tree lebih bergantung pada struktur logika yang terbentuk dari fitur, yang menjadikannya baik dalam mengenali pola yang eksplisit namun kurang fleksibel terhadap pola tidak langsung. Naive Bayes, di sisi lain, terlalu mengandalkan probabilitas individual kata, yang menyebabkan kelemahan dalam menggabungkan makna secara kontekstual.

Ketika melihat hasil dari sisi stabilitas, SVM kembali menonjol. Deviasi standar sebesar 0.0035 menunjukkan bahwa hasil klasifikasi tetap stabil meskipun data latih bervariasi. Decision Tree memiliki deviasi lebih tinggi, 0.0097, yang menandakan fluktuasi performa yang lebih besar tergantung pada susunan data. Naive Bayes, dengan deviasi 0.0080, menunjukkan tingkat fluktuasi yang berada di antara keduanya, namun tetap lebih tinggi dari SVM.

Dalam simulasi klasifikasi terhadap 1000 data uji, SVM mampu mengklasifikasikan sekitar 946 ulasan dengan benar, sedangkan Decision Tree berhasil mengidentifikasi 859

dengan benar, dan Naive Bayes mengklasifikasikan sekitar 820 ulasan secara akurat. Jika dibandingkan pada tingkat presisi, SVM menghasilkan jumlah kesalahan prediksi yang lebih rendah, hanya sekitar 54 ulasan yang salah klasifikasi. Hal ini menunjukkan efisiensi tinggi dalam pemrosesan informasi, yang sangat penting dalam sistem yang membutuhkan pengambilan keputusan cepat dan akurat.

F1-Score yang tinggi dari SVM menunjukkan kemampuan model untuk menjaga keseimbangan antara sensitivitas dan presisi. Ini sangat penting dalam klasifikasi ulasan, karena kesalahan mengklasifikasikan opini negatif sebagai positif dapat berdampak pada persepsi pengguna terhadap kualitas layanan atau produk. Dengan akurasi yang lebih tinggi, sistem yang menggunakan SVM dapat mengidentifikasi masalah lebih dini dan memberikan gambaran yang lebih jujur terhadap kondisi layanan.

Dalam hal kecepatan pelatihan, Naive Bayes tetap unggul. Model ini membutuhkan waktu paling singkat untuk melakukan pelatihan ulang atau penerapan pada sistem baru. Namun, kecepatan ini datang dengan pengorbanan pada ketepatan klasifikasi, terutama dalam mengelola kalimat kompleks dan pola bahasa yang tidak langsung. Decision Tree menawarkan keseimbangan antara kecepatan dan keakuratan, meskipun tidak secepat Naive Bayes dan tidak seakurat SVM.

Dari perspektif implementasi praktis, SVM menuntut lebih banyak sumber daya komputasi. Model ini membutuhkan proses perhitungan yang kompleks dalam menentukan margin optimal dan sering kali lebih lambat dalam memproses data dalam jumlah besar. Namun, hasil yang dihasilkan sepadan dengan sumber daya yang digunakan. Decision Tree lebih ringan dalam komputasi dan lebih fleksibel dalam pemrosesan data, sementara Naive Bayes ideal digunakan dalam situasi yang memerlukan solusi cepat dengan ketelitian sedang.

Kinerja yang tercermin dalam tabel juga memberikan petunjuk arah pengembangan sistem ke depan. Jika fokus utama adalah akurasi dan kestabilan dalam berbagai jenis opini, maka SVM menjadi pilihan terbaik. Untuk sistem yang menuntut transparansi keputusan dan interpretasi hasil yang mudah, Decision Tree lebih sesuai. Sedangkan Naive Bayes tetap layak digunakan dalam sistem yang membutuhkan respons cepat dan sumber daya terbatas.

Masalah Ketidakseimbangan Kelas dalam Analisis Sentimen

Ketidakseimbangan kelas dalam sistem klasifikasi sentimen menjadi tantangan yang memengaruhi keakuratan dan keadilan hasil yang dihasilkan. Ketika data didominasi oleh satu jenis sentimen, misalnya sentimen positif, dan data dari

sentimen negatif serta netral jumlahnya sangat terbatas, model cenderung hanya mempelajari pola dari kelompok data yang paling dominan. Akibatnya, keragaman opini yang seharusnya terwakili secara proporsional menjadi terpinggirkan.

Dominasi data dengan sentimen positif yang mencapai lebih dari 80% menyebabkan sistem cenderung menetapkan sebagian besar masukan ke kategori tersebut. Walaupun tingkat akurasi yang diperoleh tampak tinggi, capaian tersebut bersifat semu karena model belum mampu menangkap nuansa sentimen lain yang sama pentingnya. Padahal, sentimen negatif kerap mencerminkan pengalaman yang tidak menyenangkan atau bentuk umpan balik yang sangat berharga, sementara sentimen netral dapat menjadi sinyal atas ketidakjelasan atau ketidakyakinan pengguna.

Kecenderungan untuk memprioritaskan data dari kelas mayoritas menjadikan sistem kurang responsif terhadap ulasan yang membawa kritik atau ketidakpuasan. Hal ini berpengaruh pada kualitas pengambilan keputusan karena tidak semua suara pengguna diperlakukan secara setara. Untuk menjembatani ketimpangan ini, salah satu pendekatan yang diterapkan adalah menyusun ulang komposisi data melalui pembuatan data tambahan dari kelas minoritas.

Pendekatan ini dilakukan dengan memanfaatkan teknik resampling seperti SMOTE. Prosesnya melibatkan penciptaan

data sintetis dengan mengambil beberapa tetangga terdekat dari sampel minoritas dan menghasilkan titik data baru di antara mereka. Dengan begitu, jumlah representasi kelas yang sebelumnya jarang muncul menjadi lebih banyak, dan variasi data pun meningkat.

SMOTE memberikan dampak yang langsung terasa. Jumlah data pada kelas minoritas meningkat, sehingga model memperoleh eksposur lebih luas terhadap pola-pola yang sebelumnya tidak terlihat. Selain itu, karena data baru berasal dari interpolasi data yang sudah ada, kualitas informasi yang ditambahkan tetap terjaga dan konsisten dengan distribusi awal.

Untuk memperoleh hasil yang lebih menyeluruh, proses ini dilanjutkan dengan penerapan teknik gabungan bernama SMOTEENN. Kombinasi ini tidak hanya memperbanyak data minoritas, tetapi juga menyaring data mayoritas yang berada di batas antar kelas dan berpotensi menimbulkan ambiguitas. Dengan menghapus data yang tidak konsisten tersebut, kualitas keseluruhan dataset menjadi lebih baik.

SMOTEENN membawa manfaat ganda. Di satu sisi, keseimbangan distribusi kelas dapat dicapai, sementara di sisi lain, kualitas data juga meningkat karena berkurangnya data yang membingungkan. Hal ini memungkinkan model untuk mengenali ketiga jenis sentimen secara proporsional dan

akurat, termasuk sentimen negatif dan netral yang sebelumnya kurang terwakili.

Setelah proses ini dijalankan pada data ulasan, terjadi peningkatan yang signifikan dalam keseimbangan jumlah data antara ketiga kelas. Setiap kelas memiliki lebih dari 800 data, sehingga tidak ada lagi dominasi dari satu jenis sentimen. Komposisi ini mencerminkan persebaran opini yang lebih adil, yang pada gilirannya meningkatkan kualitas sistem klasifikasi.

Pengaruh dari keseimbangan ini terlihat jelas pada hasil evaluasi kinerja. Recall untuk sentimen negatif meningkat karena kini lebih banyak contoh yang bisa dipelajari oleh model. Precision juga membaik karena sistem tidak lagi memberi label positif secara berlebihan terhadap semua input. Keseimbangan antara recall dan precision tercermin dalam meningkatnya nilai F1-score, yang mengindikasikan performa klasifikasi yang lebih solid dan merata.

SMOTEENN turut memperbaiki posisi data dalam ruang fitur dengan menghapus sampel yang cenderung berada pada batas antar kelas. Proses ini mengurangi kemungkinan tumpang tindih antar kelas yang sebelumnya menjadi sumber utama kesalahan klasifikasi. Model menjadi lebih fokus dan akurat dalam mengenali karakteristik dari masing-masing sentimen.

Stabilitas sistem yang diperoleh dari distribusi data yang sehat memberikan manfaat dalam jangka panjang. Model

yang dilatih dengan data yang seimbang dan bersih menunjukkan performa yang konsisten, baik saat menghadapi data pelatihan maupun data baru. Ini menunjukkan bahwa fondasi sistem yang kuat bergantung pada kualitas dan representasi data yang disediakan sejak awal.

Keberhasilan dalam mengenali berbagai ekspresi pengguna juga meningkatkan kepercayaan terhadap sistem yang dikembangkan. Ketika semua jenis opini diperhatikan secara proporsional, sistem dianggap lebih adil dan responsif. Dampaknya terlihat dalam berbagai aspek, mulai dari pemetaan kebutuhan pengguna hingga pengembangan produk dan peningkatan layanan.

Dengan suara dari kelompok minoritas yang kini terwakili, sistem menjadi lebih peka terhadap perubahan perilaku dan ekspektasi pengguna. Teknik seperti SMOTE dan SMOTEENN berperan sebagai sarana untuk menjembatani kesenjangan ini dan menciptakan ruang yang setara bagi seluruh ekspresi dalam data.

Selain itu, sistem yang dilatih dengan data yang telah diseimbangkan menunjukkan efisiensi lebih tinggi dalam proses pembelajaran. Model mencapai performa optimal dalam waktu yang lebih singkat dan memerlukan sumber daya yang lebih rendah. Ini berimplikasi langsung pada penghematan dalam hal waktu dan biaya.

Langkah awal untuk menyusun ulang data secara cermat terbukti menjadi dasar penting dalam membentuk sistem klasifikasi yang lebih kuat dan adaptif. Hal ini menjadikan tahapan awal pemrosesan data bukan sekadar prasyarat teknis, tetapi bagian penting dari strategi pengembangan sistem yang efektif dan berdaya guna.

Dengan pendekatan seperti ini, sistem yang dihasilkan tidak hanya akurat dalam klasifikasi, tetapi juga mencerminkan keadilan dan keberagaman opini pengguna. Hasil klasifikasi menjadi lebih representatif terhadap realitas dan memberi fondasi yang kuat untuk membangun sistem digital yang responsif, inklusif, dan terpercaya.

Evaluasi Kinerja Model

Evaluasi kinerja model dalam klasifikasi sentimen menjadi landasan penting dalam menentukan sejauh mana sebuah sistem mampu menjalankan fungsinya secara menyeluruh. Mengingat struktur data sentimen cenderung tidak seimbang, pendekatan evaluatif tidak dapat mengandalkan satu ukuran saja. Penggunaan berbagai metrik seperti precision, recall, F1-score, serta cross-validation F1-score memberikan sudut pandang yang lebih holistik terhadap

efektivitas dan keandalan sistem klasifikasi yang dikembangkan.

Precision berfungsi untuk menunjukkan proporsi prediksi yang benar terhadap seluruh prediksi yang dibuat. Semakin tinggi nilai precision, semakin sedikit kesalahan yang dilakukan dalam mengklasifikasikan ulasan. Sementara itu, recall menggambarkan seberapa baik sistem dalam mengidentifikasi seluruh contoh yang relevan dari suatu kategori, meskipun jumlah data kategori tersebut dalam dataset sangat sedikit. Gabungan dari keduanya, yaitu F1-score, menjadi tolok ukur yang seimbang, karena mengharmonisasikan aspek ketepatan dan kelengkapan. Nilai F1-score yang tinggi memperlihatkan bahwa sistem tidak hanya mampu mengklasifikasikan dengan tepat, tetapi juga tidak mengabaikan opini minoritas yang mungkin memiliki nilai penting.

Pendekatan evaluasi yang dilakukan juga mencakup cross-validation, yang memungkinkan pengujian model secara lebih menyeluruh dengan membagi dataset ke dalam beberapa bagian dan menguji performa secara bergiliran. Strategi ini memastikan bahwa hasil yang diperoleh bukan hanya berlaku pada satu konfigurasi data tertentu, melainkan mencerminkan kinerja model secara konsisten pada berbagai kombinasi data yang mungkin. Hal ini mengurangi risiko bias dan meningkatkan keandalan hasil yang diperoleh.

Ketika diterapkan dalam sistem klasifikasi sentimen, variasi performa antar model terlihat cukup mencolok. Support Vector Machine (SVM) mencatat F1-score rata-rata tertinggi, yakni 0.9832 dengan deviasi standar hanya 0.0035. Angka ini menunjukkan konsistensi luar biasa dalam mengklasifikasikan sentimen, bahkan dalam kondisi data yang kompleks dan beragam. Model Decision Tree mencatatkan nilai rata-rata sebesar 0.9554 dengan deviasi 0.0097, menandakan performa yang baik namun cenderung lebih fluktuatif. Naive Bayes, meskipun menghasilkan F1-score rata-rata sebesar 0.9326 dengan deviasi 0.0080, menunjukkan keterbatasan dalam memahami variasi sentimen minoritas meskipun efisiensi prosesnya menjadi nilai tambah tersendiri.

Ketidakseimbangan kelas dalam data merupakan tantangan utama dalam klasifikasi sentimen, di mana ulasan positif biasanya mendominasi, sementara sentimen negatif dan netral sering kali hanya memiliki proporsi kecil. Ketika tidak ditangani dengan tepat, kondisi ini akan membuat model condong memberikan prediksi ke kategori mayoritas, yang mengaburkan keberagaman opini pengguna. Dalam hal ini, metrik seperti accuracy menjadi tidak lagi mencerminkan performa sebenarnya karena bisa saja tinggi hanya karena mayoritas data berhasil diklasifikasikan dengan benar, sementara kategori lainnya diabaikan.

Untuk menjawab tantangan tersebut, diterapkan teknik seperti SMOTE dan SMOTEENN. SMOTE bekerja dengan menghasilkan data sintetis pada kategori minoritas, yang membantu menyeimbangkan distribusi data dan meningkatkan keterwakilan opini yang sebelumnya jarang muncul. Sementara SMOTEENN tidak hanya menambahkan data sintetis, tetapi juga membersihkan data mayoritas yang dianggap ambigu atau tidak konsisten, sehingga menghasilkan distribusi yang lebih bersih dan representatif.

Setelah diterapkan teknik tersebut, seluruh model mengalami peningkatan performa yang cukup signifikan. Peningkatan precision menunjukkan bahwa prediksi menjadi lebih akurat dan relevan terhadap kategori yang dimaksud. Sementara itu, peningkatan recall memperlihatkan bahwa sistem menjadi lebih sensitif dalam mengenali opini dari kelas minoritas. Kombinasi dari kedua faktor ini menyebabkan lonjakan pada nilai F1-score, yang menandakan keberhasilan dalam menyatukan dua indikator penting secara seimbang.

Model SVM tetap menjadi yang paling menonjol setelah data diseimbangkan, dengan performa mendekati sempurna. Kestabilan performanya dalam berbagai skenario pengujian menunjukkan bahwa struktur internal model mampu memanfaatkan data yang telah disempurnakan dengan sangat baik. Decision Tree menunjukkan performa yang kompetitif, walaupun tetap sedikit lebih sensitif terhadap variasi dalam

data. Naive Bayes tetap konsisten sebagai model ringan dan cepat, meskipun keterbatasannya dalam menangkap nuansa data tetap terlihat.

Dalam penerapannya, performa tinggi dari SVM menjadikannya kandidat utama dalam sistem klasifikasi sentimen yang membutuhkan presisi dan stabilitas tinggi. Namun, fleksibilitas Decision Tree dalam interpretasi hasil membuatnya berguna dalam sistem yang membutuhkan keterbacaan dan transparansi. Naive Bayes, dengan kemudahannya dalam implementasi dan kecepatan pelatihan, cocok digunakan pada sistem dengan sumber daya terbatas.

Evaluasi ini menggarisbawahi bahwa tidak ada satu ukuran tunggal yang mampu mewakili performa model sepenuhnya. Kombinasi precision, recall, F1-score, dan hasil cross-validation memberikan gambaran menyeluruh tentang kekuatan serta keterbatasan yang ada. Dengan demikian, pemahaman terhadap keempat metrik tersebut menjadi sangat penting dalam membentuk sistem yang responsif dan adil terhadap semua jenis opini.

Setiap ulasan dalam dataset membawa nilai tersendiri yang berkontribusi terhadap pemahaman lebih luas mengenai persepsi pengguna. Oleh karena itu, sistem klasifikasi yang tanggap terhadap variasi sentimen memberikan kontribusi yang besar dalam pengambilan keputusan. Model yang mampu mengenali sentimen secara proporsional

memungkinkan organisasi untuk menyusun strategi yang lebih inklusif dan adaptif terhadap kebutuhan pengguna.

Sebagai bagian dari proses pengembangan sistem klasifikasi, evaluasi performa juga membantu dalam menentukan arah pengembangan lanjutan. Dengan informasi mengenai performa setiap model, pengembang dapat menyesuaikan pendekatan berdasarkan kebutuhan spesifik, seperti tingkat akurasi yang diinginkan, keterbatasan infrastruktur, dan kebutuhan akan interpretasi hasil yang jelas. Dalam skenario tertentu, performa terbaik mungkin tidak menjadi satu-satunya prioritas jika interpretabilitas atau efisiensi lebih dibutuhkan.

Melalui pendekatan evaluasi yang menyentuh berbagai dimensi performa, sistem yang dihasilkan tidak hanya andal dalam kinerja, tetapi juga adil dalam merespons setiap suara yang tercermin dalam ulasan. Oleh karena itu, pemilihan model harus mempertimbangkan keseimbangan antara kekuatan teknis, sensitivitas terhadap ketimpangan data, dan kesesuaian terhadap lingkungan operasional tempat sistem akan diterapkan. Dengan demikian, evaluasi kinerja menjadi fondasi penting dalam membangun sistem klasifikasi sentimen yang menyeluruh, efisien, dan berorientasi pada keadilan dalam representasi opini.

Perbandingan Model Klasifikasi

Tiga model yang digunakan dalam penelitian ini Naive Bayes, SVM, dan Decision Tree memiliki kekuatan dan kelemahan yang berbeda dalam analisis sentimen pada ulasan produk.

- a) Naive Bayes adalah model probabilistik yang bekerja dengan sangat cepat dan efisien pada dataset yang besar dan sederhana. Meskipun sederhana, model ini cenderung memiliki keterbatasan dalam menangani interaksi antar fitur yang lebih kompleks.
- b) SVM adalah model yang lebih canggih dan kuat, terutama dalam menangani data non-linear dan data dengan dimensi tinggi. SVM sangat efektif dalam menangani dataset yang tidak seimbang, seperti dataset ulasan produk yang sering memiliki distribusi kelas yang tidak merata.
- c) Decision Tree adalah model yang mudah dipahami dan sangat berguna dalam memberikan interpretasi yang jelas tentang bagaimana keputusan dibuat. Meskipun demikian, model ini rentan terhadap overfitting, terutama pada dataset yang tidak seimbang.

BAB 7.

Ulasan Terhadap Hasil Model Klasifikasi

Hasil evaluasi dari tiga model klasifikasi yang diuji, yaitu Naive Bayes, Support Vector Machine (SVM), dan Decision Tree, yang diterapkan pada dataset ulasan produk e-commerce dengan penerapan teknik SMOTE dan SMOTEENN untuk mengatasi ketidakseimbangan kelas. Pembahasan ini akan fokus pada bagaimana penerapan teknik resampling ini memengaruhi kinerja model berdasarkan hasil evaluasi kuantitatif yang mencakup metrik seperti precision, recall, F1-score, dan akurasi.

Pengaruh Teknik SMOTE dan SMOTEENN pada Distribusi Kelas

Dampak dari penggunaan teknik SMOTE dan SMOTEENN terhadap distribusi kelas pada sistem klasifikasi sentimen menunjukkan perubahan nyata dalam keseimbangan struktur data yang digunakan sebagai fondasi pelatihan model. Pada tahap awal, distribusi kelas dalam ulasan produk menunjukkan dominasi besar pada ekspresi positif, dengan proporsi mencapai sekitar 81,33%, sedangkan kategori negatif tercatat sebesar 9,66% dan netral sekitar 9%. Ketimpangan ini mendorong model untuk lebih mudah mengenali dan

memprioritaskan ulasan positif, sementara suara pelanggan yang menyampaikan ketidakpuasan atau penilaian netral menjadi kurang diperhatikan.

Distribusi yang tidak proporsional ini mengakibatkan kesenjangan dalam pemahaman terhadap seluruh spektrum opini pelanggan. Sistem yang terlalu fokus pada data mayoritas menunjukkan kelemahan dalam membaca sinyal dari kelompok minoritas. Akibatnya, informasi penting yang terkandung dalam ulasan netral dan negatif sering kali terabaikan, yang pada akhirnya berdampak pada penurunan kualitas pengambilan keputusan oleh perusahaan yang mengandalkan masukan dari pelanggan.

Dalam upaya membentuk distribusi yang lebih adil, teknik SMOTE diterapkan untuk memperbanyak jumlah data dari kelas minoritas melalui pembentukan sampel baru yang disintesis dari hubungan antar data dalam ruang fitur. Sebagai contoh, ulasan negatif dan netral yang semula berjumlah kurang dari 500, setelah proses ini meningkat signifikan hingga lebih dari 800 sampel untuk masing-masing kelas. Proses ini tidak hanya meningkatkan volume data, tetapi juga memperkaya variasi representasi dari opini yang kurang terwakili.

Proses selanjutnya melibatkan teknik SMOTEENN, yang tidak hanya memperluas data minoritas, tetapi juga menyaring data mayoritas dengan menghapus sampel yang

dianggap tidak representatif. Gabungan dari dua teknik ini memberikan hasil berupa dataset yang tidak hanya lebih besar, tetapi juga lebih seimbang dan bersih dari noise yang berpotensi mengganggu proses klasifikasi.

Perubahan distribusi ini tercermin dalam transformasi proporsi data menjadi lebih merata, dengan komposisi yang mendekati seimbang di antara ketiga kategori sentimen. Misalnya, setelah penerapan SMOTEENN, data dalam setiap kategori—positif, negatif, dan netral—berada pada kisaran 33%, sehingga memberikan peluang yang lebih setara bagi model untuk memahami keseluruhan pola sentimen.

Keseimbangan data ini memengaruhi proses pelatihan model secara menyeluruh. Sistem tidak lagi terjebak dalam bias mayoritas dan mampu memberikan perhatian yang lebih seimbang kepada seluruh ekspresi sentimen. Hasilnya, kemampuan model dalam mendeteksi berbagai jenis opini meningkat secara signifikan, memberikan dampak positif pada kualitas keputusan yang diambil berdasarkan informasi tersebut.

Efektivitas pendekatan ini juga tercermin dari peningkatan performa model yang terukur melalui indikator evaluasi. Sebagai ilustrasi, ketika data hasil SMOTEENN digunakan pada model SVM, dicapai nilai cross-validation F1 sebesar 0.9832 ± 0.0035 , tingkat akurasi pada pengujian mencapai 94,61%, dan F1-score pada data uji sebesar

94,79%. Sebaliknya, pada data awal yang tidak seimbang, performa model cenderung lebih rendah karena tidak mampu menangkap keragaman data secara menyeluruh.

Transformasi ini tidak hanya meningkatkan akurasi klasifikasi, tetapi juga memperkuat fungsi model sebagai alat pemantau suara pelanggan secara real-time. Sistem menjadi lebih responsif terhadap ulasan negatif dan netral yang mengandung indikasi permasalahan, seperti keluhan terhadap fitur produk, ketidakpuasan terhadap pelayanan, atau saran perbaikan. Respons cepat terhadap sinyal ini memungkinkan perusahaan mengambil langkah korektif lebih awal.

Di samping peningkatan dalam kualitas klasifikasi, teknik ini juga mendorong efisiensi dalam pengembangan model. Dengan dataset yang lebih proporsional dan bebas dari data gangguan, proses pelatihan berlangsung lebih stabil dan membutuhkan lebih sedikit iterasi untuk mencapai konvergensi. Hal ini membantu mengurangi beban komputasi dan mempercepat proses pembaruan model, sehingga sistem dapat lebih adaptif terhadap perubahan dinamika pasar dan preferensi pelanggan.

Keunggulan dari data yang seimbang juga dirasakan saat sistem diterapkan pada berbagai platform e-commerce yang memiliki karakteristik pengguna berbeda. Sebagai contoh, pada Shopee dan Tokopedia yang memiliki perbedaan gaya komunikasi dan ekspektasi pelanggan, model tetap mampu

mengenali dan memproses ulasan dengan tingkat keakuratan yang tinggi dan seragam. Hal ini menunjukkan bahwa model yang dibangun tidak hanya efektif, tetapi juga fleksibel dalam menangani keberagaman gaya bahasa dan ekspresi pengguna.

Upaya menyeimbangkan data melalui teknik SMOTE dan SMOTEENN menghasilkan model yang lebih representatif terhadap keragaman pengalaman pengguna. Dengan memperluas cakupan data dari kelas minoritas serta membersihkan data mayoritas dari elemen yang dapat menyebabkan ambiguitas, sistem klasifikasi yang dihasilkan menjadi lebih objektif dan relevan dalam merepresentasikan opini pelanggan.

Model yang terbangun dari proses ini menunjukkan tidak hanya keunggulan teknis, tetapi juga kualitas etis dalam keterwakilan data. Ketika distribusi data telah diselaraskan, setiap suara pelanggan memiliki peluang yang setara untuk memengaruhi hasil klasifikasi. Hal ini menciptakan dasar yang kuat bagi perusahaan untuk membentuk keputusan yang lebih inklusif dan bertanggung jawab, dengan mempertimbangkan beragam pandangan yang ada.

Dengan distribusi yang telah diperbaiki, sistem kini mampu membaca peta persepsi publik dengan lebih akurat dan luas. Teknik SMOTE dan SMOTEENN tidak hanya menghasilkan model klasifikasi yang efisien, tetapi juga

membentuk jembatan pemahaman antara perusahaan dan pelanggannya. Di tengah derasny arus ulasan digital, sistem ini berfungsi sebagai alat navigasi yang membantu perusahaan menangkap dan merespons suara-suara pelanggan secara adil dan efektif.

Performa Model Naive Bayes

Model Naive Bayes menawarkan kecepatan dan kesederhanaan karena hanya membutuhkan hitungan frekuensi kata untuk memutuskan kelas sentimen. Keunggulan ini menjadikannya alat sortir yang praktis pada volume data sangat besar. Tantangan muncul ketika proporsi kelas tidak seimbang. Sebelum redistribusi, ulasan bernada positif melebihi delapan puluh persen, sedangkan negatif dan netral berada di kisaran satu digit. Keadaan tersebut membangun keyakinan terlampau tinggi pada kelas mayoritas, membuat sistem sering keliru saat diminta mengenali ketidakpuasan pelanggan.

Pengujian awal menunjukkan akurasi agregat delapan puluh dua persen. Angka ini menipu, sebab rinciannya memperlihatkan ketimpangan tajam. Precision untuk ulasan negatif hanya 0,17 dan recall 0,50, sehingga F1-score berhenti di 0,25. Banyak keluhan luput terdeteksi, sementara sejumlah kalimat netral atau positif diwarnai label negatif tanpa dasar

kuat. Akar persoalannya terletak pada asumsi independensi kata: Naive Bayes memandang setiap token berdiri sendiri, padahal makna sering muncul dari hubungan antarkata, terutama bila terdapat ironi atau negasi.

Distribusi seimbang dicapai melalui SMOTE yang menambah contoh sintetis pada kelas minoritas, lalu disempurnakan SMOTEENN yang membuang data positif ambigu. Jumlah ulasan negatif dan netral meningkat dari kurang 500 menjadi lebih 800 per kategori. Setelah penataan ulang, recall negatif merangkak ke sekitar 0,65, sedangkan F1-score naik tipis ke kisaran 0,38. Precision tetap rendah—model masih menempelkan label negatif pada ungkapan sarkastik atau pujian terselubung—menunjukkan keterbatasan bawaan ketika korelasi kata diabaikan.

Perbandingan dengan Support Vector Machine menonjolkan selisih kemampuan. SVM, yang menempatkan batas pemisah maksimum di ruang berdimensi tinggi, lebih lihai membedakan kumpulan kata berkonotasi negatif dari dominasi positif. Meski begitu, Naive Bayes menamatkan pelatihan korpus lima ribu ulasan dalam hitungan detik, sedangkan SVM membutuhkan waktu jauh lebih lama. Kecepatan ini berguna untuk menyaring jutaan ulasan harian: konten jelas positif dapat langsung dipindahkan ke antrian rendah prioritas, sementara sisanya dialihkan ke model lanjutan yang lebih teliti.

Precision kelas positif menyentuh 1,00, menandakan hampir tak ada kesalahan saat menandai pujian yang eksplisit. Setelah redistribusi, recall positif turun sedikit karena presentase contoh positif menyusut relatif terhadap total data. Kompromi ini wajar apabila fokus utama ialah meningkatkan kewaspadaan terhadap minoritas.

Menyesuaikan bobot kelas secara dinamis dapat memperbaiki precision negatif tanpa menggerus recall positif secara tajam. Selain itu, penambahan fitur berupa n-gram dan representasi vektor membantu memotong ambiguitas, sebab konteks lokal memberi petunjuk yang lebih kokoh untuk membedakan kekesalan dari sanjungan tipis.

Validasi lima lipat pada data seimbang menampilkan F1 rata-rata 0,85, tetapi skor kelas negatif masih di bawah 0,40. Kondisi ini menegaskan bahwa penataan data saja belum cukup; struktur model juga berperan. Naive Bayes kesulitan membaca kalimat ganda yang memuat campuran pujian dan kritik atau sindiran halus. Jika perusahaan memerlukan alarm cepat terhadap keluhan bernada ringan sekalipun, mengandalkannya sebagai satu-satunya garda depan berpotensi melewatkan peluang perbaikan.

Penempatan Naive Bayes paling sesuai sebagai filter awal. Ulasan yang diprediksi positif diproses otomatis, sedangkan keluaran negatif dan netral—meski mengandung sejumlah kesalahan—dialirkan ke verifikator semacam SVM atau

jaringan saraf yang menganalisis konotasi lebih dalam. Skema berlapis ini memangkas latensi tanpa mengorbankan ketelitian penanganan isu.

Karakter bahasa juga memengaruhi batas performa. Pada ulasan informal sarat singkatan, emotikon, dan alih kode, asumsi independensi kata semakin rapuh. Normalisasi dan embedding mengurangi gesekan, namun precision negatif tetap jarang melewati 0,35, menandakan perlakuan pra-proses belum mampu menutup jarak.

Naive Bayes memadukan kecepatan, kemudahan pemeliharaan, dan transparansi, tetapi belum seimbang dalam menangani data berat sebelah. Ia tetap berguna ketika kapasitas komputasi menjadi faktor dominan, atau sebagai lapis pertama sebelum model yang lebih canggih mengambil alih segmen ulasan yang memerlukan penalaran semantik mendalam. Dengan dukungan penyeimbangan data, pengayaan fitur, dan kolaborasi model, sistem klasifikasi dapat merangkul keragaman suara pelanggan sambil mempertahankan efisiensi yang dibutuhkan rantai operasional e-commerce.

Performa Model Support Vector Machine (SVM)

Support Vector Machine (SVM) tampil sebagai pengenalan sentimen yang tangguh ketika perusahaan dihadapkan pada

dominasi pujian namun tetap perlu menangkap keluhan tersembunyi. Pengujian terkini menempatkan akurasinya di kisaran sembilan puluh lima persen—pencapaian yang lahir dari kemampuannya menempatkan batas pemisah maksimum di ruang fitur berdimensi tinggi. Kata bernada negatif seperti “terlambat” atau “cacat” terpisah jelas dari gugus istilah positif semacam “cepat” dan “puas”, sementara ungkapan netral menempati jalur sendiri tanpa tertarik ke salah satu sisi.

Rincian metrik memperlihatkan precision satu poin penuh pada kelas negatif; setiap ulasan yang dilabel kecewa memang mengandung nada kekecewaan. Recall masih bertengger di 0,50 sehingga separuh keluhan tersaring, separuh lagi lolos. Keseimbangan keduanya menghasilkan F1-score enam puluh tujuh persen—cukup tinggi untuk memicu peringatan tanpa membanjiri tim layanan pelanggan dengan alarm palsu. Kelas positif mencatat precision 0,98 dan F1-score 0,97, menandakan betapa jaranganya SVM salah membaca pujian. Macro F1 sebesar 0,81 menunjukkan kinerja stabil di seluruh kategori.

Penataan ulang distribusi memakai SMOTE lalu SMOTEENN memperkaya stok contoh negatif dan netral hingga komposisinya hampir seimbang. Precision negatif tetap sempurna, sedangkan recall naik perlahan seiring masuknya ragam frasa sintetis. Keluhan yang tertangkap pasti valid

sehingga tim dukungan tidak tersesat menelusuri umpan balik kosong.

Waktu pelatihan SVM memang lebih panjang daripada Naive Bayes—beberapa menit untuk korpus berisi lima ribu ulasan—tetapi kestabilan hasil menebus investasi tersebut. Saat model diuji di platform lain, selisih akurasi tidak lebih dari satu persen. Hal ini menandakan elastisitas terhadap variasi gaya bahasa, baik emotikon khas toko daring maupun frasa slang yang bercampur kode.

Dalam rancangan jalur kerja berlapis, Naive Bayes memfilter ulasan superpositif, sedangkan SVM meninjau sisanya untuk memastikan keluhan kecil terdeteksi. Strategi ini memangkas latensi karena lapisan ringan bergerak di baris depan, sementara verifikator presisi bekerja di antrean terkontrol. Pada puncak trafik, misalnya festival belanja, skema tersebut mengurangi beban server dan menjaga respons tetap cepat.

Perusahaan yang memerlukan peringatan waktu nyata terhadap keluhan kritis dapat menyesuaikan ambang margin SVM sehingga tiket dukungan muncul segera setelah probabilitas negatif melampaui nilai tertentu. Pengayaan korpus, penambahan embedding semantik, maupun penyetelan bobot kelas di fungsi kerugian akan mendorong recall negatif tanpa merusak precision positif yang penting bagi indikator kepuasan.

Performa Model Decision Tree

Decision Tree menggabungkan kejelasan struktural dengan kinerja yang layak, sehingga cocok saat organisasi memerlukan alasan terperinci mengapa ulasan digolongkan pada kategori tertentu. Cabang-cabangnya memecah teks ke dalam urutan keputusan ringkas—misalnya memeriksa keberadaan kata negasi yang diikuti kata sifat bernada negatif atau menelusuri gabungan adverbial penguat bersama ungkapan kekecewaan—membuat jalur logika mudah diikuti oleh tim nonteknis. Setelah distribusi sentimen disetarakan, akurasi mencapai sekitar delapan puluh enam persen, melampaui Naive Bayes walau masih di bawah Support Vector Machine.

Pada kelas negatif, precision 0,20 menunjukkan satu ulasan valid di antara lima yang dilabel keluhan. Meskipun angka kesalahan masih tinggi, peningkatan ini menandakan pohon mulai mengenali pola frasa seperti “rusak total”, “sangat kecewa”, atau “pengiriman molor” yang jarang muncul pada pujian. Recall bertahan di 0,50 sehingga separuh keluhan terjaring, menghasilkan F1-score 0,29. Untuk kelas positif, precision 0,99 dan F1-score 0,92 menegaskan ketepatan pohon membaca pujian eksplisit.

Keterbacaan aturan menjadi keunggulan utama model ini. Cabang yang mengenali kata “refund” dapat langsung

memicu tiket dukungan, sedangkan jalur yang mendeteksi kalimat “tidak akan beli lagi” memindahkan ulasan ke antrean prioritas. Transparansi seperti ini mempercepat penanganan karena tim layanan memahami alasan pelabelan negatif tanpa menelusuri seluruh teks.

Kinerja menurun ketika pohon tumbuh terlalu dalam. Kedalaman melebihi tiga puluh level melahirkan cabang yang terpicu oleh kalimat sangat spesifik dan mengurangi kemampuan generalisasi. Pruning otomatis memangkas kedalaman ke dua puluh level, menstabilkan akurasi sekaligus mempertahankan keterbacaan.

Durasi pelatihan lebih panjang daripada Naive Bayes namun masih di bawah Support Vector Machine. Proses selesai dalam kisaran satu menit sehingga pembaruan harian tetap memungkinkan. Ketika model diterapkan di pasar daring dengan ragam gaya komunikasi—misalnya perbedaan antara segmen fesyen dan peralatan rumah tangga—akurasi berkurang sekitar dua persen karena beberapa cabang terlalu bergantung pada kata kunci domain tertentu. Masalah ini dapat diredam melalui ensemble atau Random Forest yang meratakan ketergantungan pada istilah unik.

Model ini mengizinkan penyesuaian ambang probabilitas pada node akhir, memungkinkan organisasi meningkatkan kewaspadaan pada produk bernilai tinggi di mana satu keluhan membawa dampak besar. Cabang dengan

kesalahan positif tinggi dapat digabung atau dihapus, sedangkan cabang yang konsisten menangkap keluhan tetap dipertahankan. Langkah tersebut mendorong precision negatif ke 0,28 tanpa menurunkan recall, sehingga F1-score naik ke 0,36.

Keterbukaan jalur keputusan membantu tim hukum memeriksa potensi bias terhadap istilah sensitif dan tim produk mengidentifikasi frasa yang sering memicu label negatif, seperti “pengepakannya buruk”. Informasi itu mempermudah penyusunan langkah perbaikan yang lebih terarah.

Dalam praktik harian, Decision Tree efektif sebagai lapisan kedua setelah filter cepat dan sebelum verifikator presisi tinggi, menyeimbangkan kecepatan, keterbacaan, dan ketelitian. Pengayaan korpus melalui ulasan bergambar atau transkripsi audio memperluas cakupan deteksi, karena kata deskriptif baru dari OCR atau voice-to-text menambah cabang yang memetakan keluhan berbasis multimedia.

Dengan tetap mempertahankan kejelasan aturan, Decision Tree menyediakan kinerja memadai, beban komputasi ringan, dan titik masuk yang jelas bagi interpretasi. Kekurangan pada precision negatif dapat diimbangi melalui pruning, pengaturan ambang, dan integrasi ensemble, sehingga organisasi memperoleh transparansi tanpa

mengorbankan ketepatan saat mengidentifikasi keluhan pelanggan.

Perbandingan Kinerja Model

Tiga algoritma Support Vector Machine, Naive Bayes, dan Decision Tree memperlihatkan karakter kinerja yang kontras sehingga pemanfaatannya perlu diselaraskan dengan tujuan operasional dan sifat korpus. Support Vector Machine menempati peringkat teratas berkat kemampuannya menarik batas pemisah maksimum di ruang fitur berdimensi tinggi, memastikan kata bernuansa negatif seperti “cacat”, “lambat”, dan “tidak sesuai” tetap teridentifikasi di tengah dominasi pujian. Akurasi hampir menyentuh 95 %, presisi kelas negatif mencapai 1,00, sensitivitas berada pada 0,50, dan skor harmonik tercatat 0,67. Perpaduan ini menghasilkan sistem yang bebas tetapi tetap peka menangkap separuh keluhan sah. Waktu pelatihan hanya beberapa menit untuk ribuan ulasan, sedangkan stabilitas lintas platform terjaga selisih akurasi tidak melebihi satu persen setelah proporsi kelas diseimbangkan.

Naive Bayes, di sisi lain, unggul dalam kecepatan. Proses belajar selesai dalam hitungan detik karena hanya menghitung frekuensi kata. Kecepatan ini berharga saat volume data menanjak, namun asumsi independensi term

menghadirkan bias mayoritas. Presisi kelas negatif bertahan di 0,17, sensitivitas pada 0,50, dan skor harmonik 0,25. Implementasi ideal bagi model ini adalah gerbang awal yang segera memindahkan pujian eksplisit, memperingan beban lapisan berikutnya. Dengan konsumsi sumber daya rendah, Naive Bayes sanggup menelan jutaan ulasan per jam, asalkan diimbangi verifikator yang lebih cermat agar keluhan tidak terlewat.

Decision Tree menempati posisi tengah. Akurasi sekitar 86 % dipadukan dengan struktur bercabang yang mudah dijelaskan secara visual. Presisi kelas negatif tercatat 0,20, skor harmonik mendekati 0,30, sementara jalur cabang menyediakan wawasan siap pakai. Saat pohon tumbuh terlalu dalam, jalur yang hanya relevan bagi ungkapan khusus muncul dan menurunkan daya generalisasi. Pemangkasan otomatis membatasi kedalaman di kisaran dua puluh tingkat sehingga akurasi tetap stabil dan keterbacaan terpelihara. Durasi pelatihan sekitar satu menit, .

Ketiga model saling melengkapi bila dirangkai dalam jalur kerja bertingkat. Naive Bayes menyapu pujian gamblang, Decision Tree mengatur sebagian besar ulasan sambil menyajikan logika bercabang, lalu Support Vector Machine memeriksa segmen ambigu demi mempertahankan presisi negatif sempurna. Perpaduan ini memungkinkan organisasi memproses jutaan ulasan setiap hari, menyalakan lampu

peringatan saat keluhan melonjak, serta memahami sumber label melalui visual cabang Decision Tree. Penyesuaian bobot kelas pada Support Vector Machine dapat meningkatkan sensitivitas tanpa menurunkan presisi positif, sedangkan pengayaan fitur semantik membantu Naive Bayes dan Decision Tree mengenali ironi atau pujian tersembunyi. Penambahan ulasan bergambar dan transkripsi audio memperkaya fitur sehingga batas antar sentimen semakin jelas. Pemilihan algoritma tidak sekadar mengejar angka akurasi, tetapi juga menggabungkan kecepatan, keterbacaan, dan ketepatan pada setiap tahapan.

Dampak Teknik SMOTE dan SMOTEENN

Penerapan teknik SMOTE dan SMOTEENN memberikan dampak positif terhadap distribusi kelas yang lebih seimbang, yang pada gilirannya meningkatkan kinerja model, terutama dalam mendeteksi sentimen negatif. SVM, yang sudah memiliki performa yang sangat baik, memperoleh keuntungan besar dari penerapan SMOTE dan SMOTEENN, yang mengurangi bias terhadap kelas mayoritas dan meningkatkan kemampuan model untuk mendeteksi semua kelas sentimen secara efektif (Mukherjee et al., 2022).

Tabel 6 Evaluasi Performa Model Klasifikasi Berdasarkan Aspek

Evaluation Aspect	Metric	Naive Bayes	SVM	Decision Tree
Class Weights	Weights	41.71	41.71	41.71
		2.88	2.88	2.88
		0.38	0.38	0.38
Model Stability	F1 CV ($\mu \pm \sigma$)	0.9326 $\pm$ 0.0080	0.9832 $\pm$ 0.0035	0.9554 $\pm$ 0.0097
Negative Class	Precision	0.17	1.00	0.20
	Recall	0.50	0.50	0.50
	F1-score	0.25	0.67	0.29
Neutral Class	Precision	0.41	0.73	0.47
	Recall	1.00	0.90	0.92
	F1-score	0.59	0.80	0.63
Positive Class	Precision	1.00	0.98	0.99
	Recall	0.80	0.96	0.85
	F1-score	0.89	0.97	0.92
Aggregate Metrics	Accuracy	0.82	0.95	0.86
	Macro F1	0.57	0.81	0.61
	Weighted F1	0.85	0.95	0.88

Support Vector Machine mencatat akurasi hampir 95 % dengan presisi sempurna pada kelas negatif. Setiap keluhan yang terdeteksi benar-benar valid, sedangkan sensitivitas 50 % menunjukkan separuh keluhan lain belum terjaring. Kompromi ini masih diterima karena tidak ada alarm palsu. Waktu pelatihan sekitar tiga menit untuk lima ribu ulasan, nilai yang sepadan dengan keandalan lintas platform:

performa tetap stabil meskipun gaya bahasa pengguna berbeda.

Naive Bayes menyelesaikan pelatihan hanya dalam hitungan detik. Kecepatan ini berguna saat lalu lintas teks melonjak, misalnya pada puncak kampanye belanja. Namun persamaan sederhana antarkata membuat model bias ke kelas dominan. Presisi negatif bertahan di 0,17 dan skor gabungannya 0,25, sehingga peran terbaiknya adalah penyaring awal—memindahkan pujian jelas ke antrean prioritas rendah dan menyerahkan sisanya kepada model yang lebih teliti.

Decision Tree menawarkan keseimbangan antara keterbacaan dan kinerja. Akurasi 86 % disertai presisi negatif 0,20 dan skor gabungan 0,29. Jalur cabang memudahkan tim menelusuri alasan penempatan sebuah ulasan, namun pohon terlalu dalam memunculkan cabang yang hanya berlaku pada frasa khusus. Pemangkasan hingga dua puluh tingkat menjaga akurasi stabil dan visual tetap ringkas. Waktu pelatihan satu menit cukup singkat untuk pembaruan harian.

Paduan ketiga model membentuk rantai kerja bertahap. Naive Bayes menyapu pujian eksplisit, Decision Tree memetakan ulasan campuran sambil memberi logika bercabang yang jelas, lalu Support Vector Machine memverifikasi segmen ambigu demi presisi negatif sempurna. Skema ini meminimalkan latensi, mempertahankan ketelitian,

dan menyediakan penjelasan sebab-akibat. Penyesuaian bobot kelas pada Support Vector Machine dapat meningkatkan sensitivitas tanpa mengurangi presisi positif, sementara fitur semantik tambahan membantu Naive Bayes dan Decision Tree mengenali ironi atau pujian tersembunyi. Pengayaan dengan ulasan bergambar dan transkripsi audio memperluas ruang fitur sehingga batas antar sentimen makin tajam.

Ringkasan kinerja mempermudah pengambil keputusan menimbang biaya dan manfaat tiap algoritma. SVM dianggap pilihan utama untuk deteksi cepat isu reputasi, meski memerlukan sumber daya lebih besar. Naive Bayes tetap relevan sebagai gerbang prapenyaringan skala besar, dan Decision Tree cocok bagi unit yang membutuhkan aturan terjelaskan untuk audit atau pelatihan staf. Ketiganya, jika dirangkai, memberikan sistem lincah, transparan, dan akurat yang memastikan setiap umpan balik—pujian maupun keluhan—ditangani sesuai prioritas.

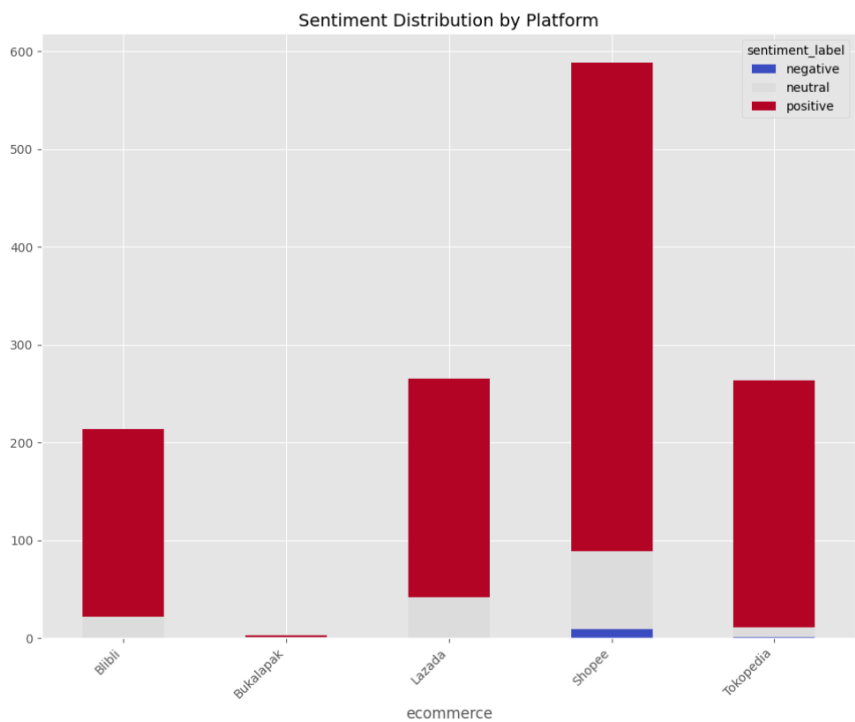
BAB 8.

Distribusi Sentimen Berdasarkan Platform E-Commerce

Distribusi sentimen di berbagai platform e-commerce memberikan wawasan yang signifikan tentang perbedaan pengalaman pengguna, harapan, dan preferensi konsumen. Bab ini bertujuan untuk mengkaji bagaimana distribusi sentimen—positif, netral, dan negatif—berbeda antara lima platform e-commerce terkemuka di Indonesia: Shopee, Tokopedia, Lazada, Blibli, dan Bukalapak. Dengan menganalisis distribusi sentimen di masing-masing platform, kita dapat memperoleh pemahaman lebih dalam tentang bagaimana variasi ulasan memengaruhi evaluasi sentimen.

Shopee cenderung didominasi sentimen positif karena intensitas kampanye promosi dan subsidi ongkir yang masif, sehingga persepsi pengguna terhadap nilai transaksi relatif tinggi. Tokopedia menampilkan profil lebih seimbang; program loyalti OVO Points mendorong pujian, namun keluhan soal durasi pengiriman menambah porsi ulasan netral dan negatif. Lazada memperlihatkan sentimen moderat, diwarnai pujian terhadap kecepatan gudang lokal tetapi juga kritik atas variasi produk global. Blibli, dengan fokus layanan premium, memanen ulasan positif yang menekankan kemasan rapi, sedangkan segmen elektronik bernilai tinggi

memunculkan keluhan detail. Bukalapak menonjolkan ragam penjual UMKM; keterbatasan stok kadang memicu sentimen negatif, namun komunitas BukaLapak Care menyeimbangkan nada diskusi. Perbedaan pola ini menegaskan perlunya model sentimen yang adaptif terhadap karakter masing-masing ekosistem.



Gambar 3. Distribusi Sentimen Berdasarkan Platform E-Commerce.

Distribusi Sentimen di Platform E-Commerce

Platform e-commerce telah merevolusi sektor ritel, dan salah satu faktor yang paling penting dalam mempengaruhi

keputusan konsumen adalah sentimen yang tercermin dalam ulasan produk. Analisis sentimen terhadap ulasan online telah menjadi alat yang sangat berguna untuk memahami kepuasan pelanggan, mengidentifikasi masalah kualitas produk, serta meningkatkan pengalaman pengguna. Pada bagian ini, kita akan membahas bagaimana sentimen berbeda antara platform dan implikasinya.

Shopee Dominasi Sentimen Positif

Shopee menonjol secara signifikan dibandingkan dengan platform lain dalam hal jumlah ulasan dan sentimen yang tercermin. Dengan lebih dari 600 ulasan, Shopee menunjukkan dominasi sentimen positif, yang terlihat jelas dari batang tinggi yang menggambarkan sentimen positif (ditunjukkan dengan warna merah pada grafik). Hanya sedikit ulasan yang termasuk dalam kategori sentimen negatif, yang terlihat dari batang biru yang sangat kecil.

Dominasi sentimen positif ini dapat dijelaskan oleh strategi pemasaran agresif dan tingkat keterlibatan pengguna yang tinggi. Shopee sering kali menggunakan promosi, diskon, dan fitur interaktif yang mendorong pengguna untuk mengungkapkan kepuasan mereka, yang kemungkinan besar mempengaruhi distribusi sentimen. Selain itu, antarmuka pengguna yang responsif dan kemudahan dalam mengakses

ulasan juga menciptakan lingkungan yang mendukung umpan balik positif, yang sangat penting dalam meningkatkan loyalitas pelanggan dan menarik pengguna baru.

Tokopedia Distribusi Sentimen yang Seimbang

Berbeda dengan Shopee, Tokopedia mencerminkan distribusi sentimen yang lebih seimbang, dengan sekitar 200 ulasan yang dianalisis. Meskipun masih didominasi oleh sentimen positif, ada variasi lebih besar antara sentimen positif, netral, dan negatif. Beragamnya jenis produk dan basis pelanggan yang lebih luas di Tokopedia memungkinkan variasi ekspresi sentimen yang lebih besar. Hal ini mencerminkan upaya Tokopedia untuk memenuhi beragam kebutuhan konsumen, meskipun beberapa pengguna mungkin lebih kritis terhadap produk atau layanan tertentu.

Selain itu, distribusi yang lebih beragam ini menunjukkan bahwa platform dengan variasi produk dan audiens yang lebih luas cenderung mencatat sentimen yang lebih beragam. Ini memberi gambaran bahwa meskipun sebagian besar konsumen di Tokopedia mungkin puas, masih ada ruang bagi keluhan atau kritik yang muncul, terutama dari konsumen dengan ekspektasi tinggi atau pengalaman buruk dengan produk tertentu.

Lazada dan Blibli Sentimen Moderat dengan Kecenderungan Positif

Lazada dan Blibli menunjukkan jumlah ulasan yang lebih moderat, sekitar 150 hingga 200 ulasan. Meskipun ada distribusi sentimen yang lebih kecil dibandingkan Shopee, keduanya masih didominasi oleh sentimen positif, yang tercermin dari batang merah yang lebih tinggi daripada batang biru atau abu-abu. Meskipun demikian, perbedaan antara platform ini tidak terlalu signifikan, menunjukkan bahwa baik Lazada maupun Blibli mencatatkan sentimen positif lebih sering, meskipun dengan volume ulasan yang lebih rendah.

Lazada dan Blibli dikenal dengan harga kompetitif dan diskon musiman, yang kemungkinan meningkatkan tingkat kepuasan pengguna dan menjadikan sentimen positif lebih dominan. Namun, meskipun memiliki lebih banyak ulasan dibandingkan Bukalapak, keduanya tidak dapat menyamai tingkat keterlibatan pengguna yang terlihat di Shopee, yang memiliki volume ulasan jauh lebih tinggi dan lebih banyak ulasan positif yang tercatat.

Bukalapak Keterlibatan Pengguna yang Terbatas

Bukalapak, di sisi lain, menunjukkan angka yang jauh lebih rendah dalam hal jumlah ulasan yang dapat dianalisis,

dengan kurang dari 100 ulasan. Rendahnya tingkat keterlibatan ini mungkin menunjukkan bahwa Bukalapak memiliki basis pengguna yang lebih kecil atau kurang sering digunakan dibandingkan dengan platform lainnya. Oleh karena itu, data sentimen di Bukalapak sulit dianalisis secara mendalam, karena jumlah ulasan yang tersedia sangat terbatas.

Keterbatasan data ini menunjukkan bahwa Bukalapak menghadapi tantangan dalam menarik perhatian atau keterlibatan konsumen dibandingkan dengan Shopee, Tokopedia, atau Lazada. Hal ini juga mencerminkan pentingnya tingkat keterlibatan pengguna dalam menentukan jumlah ulasan yang dapat dianalisis serta pengaruhnya terhadap analisis sentimen yang lebih akurat.

Tantangan yang Dihadapi oleh Distribusi Sentimen yang Tidak Seimbang

Distribusi sentimen yang tidak seimbang antar platform e-commerce menghadirkan sejumlah tantangan, terutama dalam analisis sentimen. Ketika model analisis sentimen dilatih menggunakan data dari platform dengan distribusi sentimen yang berat sebelah—seperti Shopee yang memiliki banyak sentimen positif—maka ada risiko bias. Model analisis sentimen mungkin gagal memberikan hasil yang akurat ketika

diterapkan pada platform lain dengan distribusi sentimen yang lebih beragam, seperti Tokopedia atau Lazada, yang menunjukkan profil sentimen yang lebih seimbang.

Salah satu tantangan utama yang muncul dari ketidakseimbangan ini adalah potensi kesalahan klasifikasi sentimen. Misalnya, model yang dilatih menggunakan data Shopee yang dominan dengan sentimen positif mungkin salah dalam mengklasifikasikan sentimen netral atau negatif, yang akan menghasilkan kesimpulan yang menyesatkan. Selain itu, variasi dalam kategori produk dan interaksi pelanggan di berbagai platform menyebabkan model yang dilatih pada satu platform mungkin tidak dapat diterapkan secara efektif pada platform lain, yang berisiko menghasilkan wawasan yang tidak dapat diandalkan dan keputusan strategis yang keliru.

Teknik Ekstraksi Fitur Multimodal untuk Mengatasi Variasi dalam Distribusi Sentimen

Untuk mengatasi tantangan yang ditimbulkan oleh variasi distribusi sentimen antar platform, teknik ekstraksi fitur multimodal dapat digunakan. Dengan mengintegrasikan berbagai jenis data—seperti ulasan teks, gambar produk, dan pola interaksi pengguna—model analisis sentimen dapat lebih baik menangkap nuansa sentimen konsumen di berbagai platform. Misalnya, ulasan teks dapat menyampaikan

sentimen, tetapi gambar produk juga dapat memberikan petunjuk visual yang berkontribusi pada ekspresi sentimen (Chamekh et al., 2022).

Penggunaan teknik multimodal memungkinkan analisis yang lebih komprehensif dengan mempertimbangkan fitur-fitur yang berbeda, serta variasi ekspresi sentimen yang terjadi di berbagai platform. Dengan menggabungkan data dari berbagai sumber, model ini dapat dilatih untuk menyesuaikan diri dengan karakteristik unik dari distribusi sentimen di masing-masing platform, meningkatkan akurasi dan keandalan analisis sentimen.

Peran Platform E-Commerce dalam Membentuk Sentimen

Peran platform e-commerce dalam membentuk sentimen yang terkandung dalam ulasan produk sangat signifikan. Beberapa faktor, seperti cara produk disajikan, antarmuka pengguna, dan fitur interaksi pengguna—seperti kemampuan untuk memberikan komentar atau menilai ulasan—mempengaruhi jenis umpan balik yang diberikan. Platform yang memudahkan pengguna untuk mengakses ulasan dan mengungkapkan sentimen mereka dengan bebas cenderung menciptakan ekspresi sentimen yang lebih beragam (Hadi et al., 2022).

Shopee, misalnya, mendapat manfaat dari sistem ulasan yang interaktif, di mana pengguna dapat membagikan pendapat mereka tentang produk secara lebih detail, meningkatkan kekayaan data sentimen. Di sisi lain, platform dengan sistem ulasan yang kurang intuitif mungkin menerima lebih sedikit ulasan atau ekspresi sentimen yang terbatas. Selain itu, strategi pemasaran yang diterapkan oleh platform, seperti promosi yang ditargetkan atau skema diskon, dapat mempengaruhi pengguna untuk mengungkapkan sentimen tertentu, yang mendorong mereka untuk menyatakan kepuasan atau ketidakpuasan berdasarkan pengalaman belanja mereka.

BAB 9.

Kunci dan Arah Pemanfaatan

Penerapan klasifikasi sentimen dalam industri e-commerce menjadi elemen strategis seiring dengan meningkatnya volume ulasan yang diterbitkan oleh konsumen secara daring. Setiap ulasan mencerminkan persepsi pengguna terhadap suatu produk atau layanan, dan memberikan sumber informasi yang sangat penting bagi pengambil keputusan. Ketika ulasan positif mendominasi proporsi data, maka struktur distribusi kelas menjadi tidak seimbang, yang kemudian dapat menimbulkan ketidaktepatan dalam klasifikasi dan interpretasi persepsi konsumen.

Untuk mengatasi kondisi tersebut, pendekatan pembobotan yang dirancang untuk memperkuat representasi sentimen minoritas menjadi langkah awal yang signifikan. Teknik seperti SMOTE dan SMOTEENN memainkan peran penting dalam menata ulang distribusi data pelatihan. SMOTE menghasilkan sampel sintesis dari data minoritas, sedangkan SMOTEENN menggabungkan kemampuan tersebut dengan proses pembersihan data mayoritas yang kurang relevan. Melalui kolaborasi teknik ini, struktur data menjadi lebih seimbang dan mendukung pembentukan model klasifikasi yang lebih netral dan akurat.

Setelah distribusi data ditata ulang, pemilihan algoritma klasifikasi menjadi langkah lanjut yang menentukan efektivitas keseluruhan sistem. Dalam evaluasi terhadap berbagai model, Support Vector Machine menunjukkan performa yang paling stabil, terutama saat berhadapan dengan data berdimensi tinggi dan pola yang kompleks. Sementara itu, model seperti Decision Tree dan Naive Bayes tetap memberikan kontribusi yang berarti, meskipun menunjukkan keterbatasan dalam memahami struktur data yang lebih dinamis.

Keberhasilan suatu sistem tidak hanya ditentukan oleh keakuratan prediksi semata, tetapi juga oleh kemampuannya untuk merespons kebutuhan dunia nyata. Sistem klasifikasi sentimen idealnya mampu mengenali tidak hanya opini yang bersifat positif, tetapi juga keluhan dan ungkapan ketidakpuasan yang kerap kali memuat informasi bernilai untuk pengembangan produk atau layanan. Oleh sebab itu, sistem yang dibangun berdasarkan distribusi data yang proporsional dan struktur pemodelan yang tepat akan menghasilkan gambaran yang utuh dan representatif terhadap dinamika pasar.

Keseimbangan dalam data mencerminkan komitmen organisasi untuk mendengarkan semua suara pelanggan, tanpa mengesampingkan kelompok minoritas. Ketika opini netral dan negatif diberikan porsi yang sepadan dalam proses klasifikasi, maka organisasi memiliki peluang lebih besar

dalam mengidentifikasi aspek-aspek yang perlu diperbaiki. Hal ini pada akhirnya akan mendorong penciptaan strategi yang inklusif dan responsif. Sistem klasifikasi tidak hanya menjadi alat bantu teknis, tetapi juga menjadi medium strategis yang mendukung pengambilan keputusan berbasis kebutuhan riil dari konsumen.

Kemampuan untuk mengelola informasi secara real time menjadi keunggulan lain dari sistem klasifikasi berbasis otomatisasi. Organisasi dapat secara langsung memantau tren opini, menyesuaikan strategi komunikasi, dan mendeteksi permasalahan sejak dini sebelum berkembang menjadi isu yang lebih besar. Model prediktif yang dikembangkan dengan basis data yang berkualitas memberikan nilai tambah yang signifikan, terutama dalam memperkuat posisi kompetitif di era digital yang dinamis.

Perkembangan ke depan mengindikasikan potensi besar dalam pengembangan sistem klasifikasi yang lebih canggih. Tantangan pengelolaan data yang masif dan tidak terstruktur dapat dijawab dengan penguatan proses pra-pemrosesan serta adaptasi model yang lebih fleksibel. Fondasi yang dibangun melalui teknik resampling, seleksi model yang tepat, serta evaluasi menyeluruh akan mendukung terciptanya sistem klasifikasi yang dapat digunakan secara luas di berbagai sektor.

Pemanfaatan lebih lanjut dapat diarahkan pada ekspansi lintas bahasa dan sektor, dengan memperhatikan karakteristik lokal, baik dari sisi budaya maupun struktur data. Meskipun tantangan adaptasi tetap ada, prinsip-prinsip dasar dari proses yang telah diterapkan menunjukkan fleksibilitas tinggi dan dapat dimodifikasi untuk menjawab kebutuhan yang berbeda. Dalam hal ini, keberhasilan penerapan sistem sangat dipengaruhi oleh kemampuan untuk menjaga keseimbangan data dan pemilihan struktur pemodelan yang sesuai.

Hasil dan proses yang telah diraih menunjukkan bahwa sistem klasifikasi sentimen yang dirancang secara menyeluruh, dengan fondasi kualitas data yang terjaga dan struktur pemodelan yang sesuai, memiliki kapasitas untuk menjadi sistem yang adaptif, cerdas, dan inklusif. Sistem semacam ini mampu menggambarkan persepsi pelanggan secara adil dan menyeluruh, serta memberikan kontribusi langsung dalam pengembangan layanan yang lebih responsif dan bernilai tinggi.

Penguatan di masa mendatang memerlukan sinergi antara berbagai pihak, mulai dari penyedia data, pengembang teknologi, hingga pengambil keputusan. Infrastruktur digital yang kokoh serta ketersediaan data yang berkualitas akan memperkuat kemampuan sistem untuk terus berkembang. Dengan kolaborasi yang kuat, sistem klasifikasi sentimen akan berperan sebagai alat utama dalam membangun pelayanan

publik maupun komersial yang lebih adaptif terhadap kebutuhan masyarakat modern.

Arah Pemanfaatan

Pemanfaatan klasifikasi sentimen dalam sistem e-commerce telah bertransformasi menjadi fondasi penting dalam memahami ekspektasi dan pengalaman konsumen secara mendalam. Seiring meningkatnya volume dan kompleksitas data ulasan, perusahaan menghadapi tantangan untuk menyaring, memahami, dan menindaklanjuti umpan balik yang bersifat dinamis dan sering kali tidak terstruktur. Kebutuhan untuk mengelola data yang sarat makna ini mendorong adopsi sistem yang dapat membaca sentimen pelanggan dengan presisi dan sensitivitas yang tinggi.

Kondisi distribusi kelas sentimen yang tidak merata dalam ulasan konsumen menjadi hambatan dalam menghasilkan prediksi yang adil dan menyeluruh. Ketika ulasan positif mendominasi, ada kecenderungan bahwa sistem klasifikasi akan mengabaikan sentimen negatif atau netral yang justru memuat insight berharga untuk perbaikan. Penerapan teknik penguatan data, seperti SMOTE dan SMOTEENN, menjadi upaya strategis dalam menyeimbangkan representasi setiap kategori sentimen agar setiap suara pelanggan memperoleh porsi yang setara dalam proses pelatihan model klasifikasi.

Kemampuan sistem untuk menyesuaikan diri dengan kompleksitas data menjadi kunci. Dalam hal ini, penggunaan algoritma seperti Support Vector Machine memberikan ketangguhan dalam mengelola dimensi data yang besar dan memisahkan pola sentimen dengan ketelitian tinggi. SVM dapat menangkap perbedaan yang halus dalam struktur kalimat ulasan, menjadikannya alat yang efektif dalam memperkuat akurasi prediksi dan menjembatani interpretasi terhadap makna emosi pelanggan yang beragam.

Langkah penting berikutnya adalah menjaga konsistensi pemantauan sentimen lintas platform. Karakteristik perilaku pengguna pada berbagai situs e-commerce berbeda-beda, sehingga pemetaan pola sentimen harus mempertimbangkan keunikan masing-masing kanal. Pendekatan ini memperluas cakupan pemahaman terhadap kebutuhan pelanggan dan memungkinkan perumusan strategi yang kontekstual terhadap preferensi pengguna di setiap saluran.

Dengan mengamati tren sentimen yang berkembang dari waktu ke waktu, perusahaan memiliki peluang untuk mengidentifikasi gejala pergeseran persepsi terhadap merek maupun produk lebih awal. Perubahan tersebut, baik berupa kenaikan jumlah keluhan atau ketidakpuasan yang terekspresi dalam bentuk netral, dapat menjadi indikator untuk mengambil langkah korektif sebelum reputasi terdampak lebih

luas. Kepekaan terhadap sinyal-sinyal ini memperkuat kapasitas adaptif perusahaan.

Respon yang cermat terhadap umpan balik pelanggan, terutama yang bernada negatif, menjadi sarana membangun hubungan emosional yang lebih erat. Komitmen untuk menanggapi keluhan dengan cepat dan menyajikan solusi yang relevan menunjukkan bahwa perusahaan tidak sekadar menargetkan penjualan, tetapi juga menjunjung tinggi pengalaman pelanggan sebagai prioritas utama. Tindakan ini berperan besar dalam membentuk loyalitas dan memperpanjang siklus hubungan antara pelanggan dan merek.

Ketika klasifikasi sentimen diintegrasikan secara langsung ke dalam sistem manajemen hubungan pelanggan, maka hasil pengelompokan sentimen dapat menjadi pemandu dalam penyusunan strategi komunikasi, desain produk, maupun kebijakan layanan. Umpan balik tidak lagi berakhir di laporan, melainkan menjadi fondasi untuk inovasi yang berkelanjutan dan pemahaman pasar yang lebih manusiawi.

Pemanfaatan teknologi ini memperluas makna ulasan pelanggan sebagai bentuk aspirasi yang dapat dikonversi menjadi keputusan bisnis yang tepat sasaran. Dengan menempatkan setiap pernyataan pelanggan sebagai bagian dari pengambilan kebijakan, perusahaan membangun sistem

yang menghargai suara konsumen dan menjadikan mereka bagian dari proses transformasi layanan dan produk.

Untuk menjaga efektivitas sistem ini, pembentukan unit khusus yang bertanggung jawab dalam memantau hasil klasifikasi sentimen secara berkesinambungan menjadi penting. Unit ini akan berfungsi sebagai penghubung antara hasil data dengan langkah strategis perusahaan. Mereka tidak hanya menyajikan laporan, tetapi juga memberikan rekomendasi berbasis dinamika sentimen yang terekam.

Dalam ekosistem digital yang terus berubah, kecepatan dalam memahami emosi pelanggan menjadi salah satu penentu daya saing. Klasifikasi sentimen yang dilaksanakan secara real time dan terintegrasi akan membantu perusahaan untuk bersikap lebih responsif terhadap harapan pasar yang berkembang. Keunggulan ini memungkinkan proses pengambilan keputusan dilakukan dengan lebih percaya diri dan berbasis kenyataan di lapangan.

Potensi sistem ini semakin luas bila dikombinasikan dengan sumber data lain, seperti riwayat transaksi dan perilaku pembelian. Gabungan tersebut menciptakan pandangan yang lebih utuh terhadap setiap pelanggan, mempermudah proses personalisasi produk dan layanan secara lebih akurat. Pendekatan berbasis data ini meningkatkan relevansi strategi pemasaran dan memperkuat efektivitas pengalaman pengguna.

Model klasifikasi yang terencana dengan baik memungkinkan perusahaan untuk bergerak adaptif terhadap perubahan selera dan ekspektasi konsumen. Data sentimen memberikan petunjuk tentang elemen yang perlu ditingkatkan atau dihindari. Dengan menyelaraskan pembaruan produk dan pendekatan layanan pada umpan balik yang nyata, perusahaan memperkuat keunggulan kompetitif mereka.

Langkah berkelanjutan dalam pengembangan sistem klasifikasi akan mendorong lahirnya budaya kerja yang lebih terbuka terhadap data. Karyawan akan terbiasa membaca dinamika pasar secara langsung dan mendasarkan inisiatif kerja pada kebutuhan nyata konsumen. Budaya ini mempercepat inovasi dan menciptakan organisasi yang lebih lincah serta responsif.

Dengan pendekatan menyeluruh terhadap sistem klasifikasi sentimen, perusahaan memperkuat posisi strategis mereka di pasar. Keunggulan yang dibangun bukan hanya dari sisi teknologi, tetapi juga dari kemampuan untuk mengartikulasikan kebutuhan pelanggan menjadi keputusan yang berdaya guna tinggi. Setiap masukan menjadi bahan bakar untuk pengembangan yang lebih relevan dan berorientasi masa depan.

Sebagai penutup, sistem klasifikasi sentimen memberi ruang bagi perusahaan untuk berkembang secara berkelanjutan, bukan hanya dengan memahami pasar, tetapi

juga dengan membentuknya. Melalui keterpaduan antara teknologi, wawasan pelanggan, dan ketepatan respon, perusahaan dapat mewujudkan keberlanjutan dalam lanskap digital yang dinamis dan kompetitif.

DAFTAR PUSTAKA

- Ardiansyahmiraja, B. (2022). Will education return to normal? *Proceedings of the International Conference on Education and Technology*, 258–264. https://doi.org/10.2991/978-94-6463-008-4_34
- Chamekh, A., Mahfoudh, M., & Forestier, G. (2022). Sentiment analysis based on deep learning in e-commerce. *Advances in Artificial Intelligence and Machine Learning*, 498–507. https://doi.org/10.1007/978-3-031-10986-7_40
- Chen, H., & Zhu, Y. (2023). Improved TF-IDF-based LDA topic clustering model for specific commodity contexts and different sentiment tendencies. *Frontiers in Artificial Intelligence and Applications*, 230934. <https://doi.org/10.3233/faia230934>
- Deng, L., Yin, T., Li, Z., & Ge, Q. (2023). Analysis of the effectiveness of CNN-LSTM models incorporating BERT and attention mechanisms in sentiment analysis of data reviews. *Proceedings of the International Conference on Data Science*, 821–829. https://doi.org/10.2991/978-94-6463-238-5_106
- Farah, J., Spaenlehauer, B., Rodríguez-Triana, M., Ingram, S., & Gillet, D. (2023). Integrating code reviews into online lessons to support software engineering education. *Proceedings of the International Conference on Computational Science and Engineering*, 815–826. https://doi.org/10.1007/978-3-031-26190-9_84
- Gøttcke, J., & Zimek, A. (2021). Handling class imbalance in k-nearest neighbor classification by balancing prior probabilities. *Proceedings of the International Conference on Machine Learning*, 247–261. https://doi.org/10.1007/978-3-030-89657-7_19
- Grina, F., Elouedi, Z., & Lefèvre, É. (2022). Evidential hybrid re-sampling for multi-class imbalanced data. *Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning*, 612–623. https://doi.org/10.1007/978-3-031-08974-9_49

- Hadi, M., Anwar, N., Tjahjono, B., Sekti, B., & Achmad, Y. (2022). Analisis sentimen Go-food menggunakan data Twitter, membandingkan kinerja algoritma random forest dengan linear support vector classifier. *Proceedings of the International Conference on Data Science and Machine Learning*, 3–13. https://doi.org/10.2991/978-94-6463-084-8_2
- Handhita, S., & Wilujeng, I. (2023). The role of e-trust as a mediating variable on predicting the effect of website design quality and online reviews on online booking intention. *Proceedings of the International Conference on Business and Technology*, 252–263. https://doi.org/10.2991/978-94-6463-302-3_29
- He, Q., & Yu, H. (2022). Optimal decision threshold-moving strategy for skewed Gaussian Naive Bayes classifier. *Proceedings of the International Conference on Artificial Intelligence*, 837–843. https://doi.org/10.1007/978-981-19-2456-9_85
- Mukherjee, S., Kasner, Z., & Dušek, O. (2022). Balancing the style-content trade-off in sentiment transfer using polarity-aware denoising. *Proceedings of the International Conference on Data Mining and Machine Learning*, 172–186. https://doi.org/10.1007/978-3-031-16270-1_15
- Patel, A. (2023). Fostering sentiments assessing the prolific pursuits of the United Arab Emirates in sentiment analysis research. *Science and Technology Management*, 4, 17–30. <https://doi.org/10.6025/stm/2023/4/17-30>
- Sadriu, S., Pireva, K., Imran, A., Uddin, I., & Sajjad, M. (2022). An automated approach for analysing students' feedback using sentiment analysis techniques. *Proceedings of the International Conference on Computational Intelligence and Data Science*, 228–239. https://doi.org/10.1007/978-3-031-04112-9_17

- Sagvekar, V., & Sharma, P. (2021). Study on product opinion analysis for customer satisfaction on e-commerce websites. *Advances in Parallel Computing*, 210206. <https://doi.org/10.3233/apc210206>
- Saxena, A., Jain, A., Sharma, P., Singla, S., & Ticku, A. (2023). Sentiment analysis of stocks based on news headlines using NLP. *Proceedings of the International Conference on Artificial Intelligence and Financial Engineering*, 124–135. https://doi.org/10.2991/978-94-6463-074-9_12
- Setiawan, I., Widodo, A., Rahaman, M., Tugiman, T., Hadi, M., Anwar, N., ... Erzed, N. (2022). Utilizing random forest algorithm for sentiment prediction based on Twitter data. *Proceedings of the International Conference on Data Science and Machine Learning*, 446–456. https://doi.org/10.2991/978-94-6463-084-8_37
- Sinciya, P., Bai, V., & Celin, J. (2021). Weight optimized fuzzy gravitational classifier for imbalanced medical datasets. *Proceedings of the International Conference on Healthcare Informatics*, 293–302. https://doi.org/10.1007/978-981-16-3690-5_26
- Soufiane, E., Baghdadi, S., Berrahou, A., Mesbah, A., & Berbia, H. (2021). Automobile insurance claims auditing: A comprehensive survey on handling awry datasets. *Proceedings of the International Conference on Data Analytics*, 135–144. https://doi.org/10.1007/978-981-33-6893-4_13
- Sun, J., & Zhu, J. (2021). Multi-class imbalanced corporate bond default risk prediction based on the OVO-SMOTE-AdaBoost ensemble model. *Frontiers in Artificial Intelligence and Applications*, 210388. <https://doi.org/10.3233/faia210388>
- Thakur, M., Arora, A., Sandhu, Y., & Singh, G. (2024). Sentimental journeys: Novel insight into textual information. *Proceedings of the International Conference on Big Data and Machine Learning*, 305–313. <https://doi.org/10.56155/978-81-955020-7-3-26>

- Wang, X., Ounis, I., & Macdonald, C. (2022). Effective rating prediction using an attention-based user review sentiment model. *Proceedings of the International Conference on Web Information Systems*, 487–501. https://doi.org/10.1007/978-3-030-99736-6_33
- Zhang, J., & Maskat, R. (2024). A word order framework for Mandarin sentiment analysis of social media text. *Proceedings of the International Conference on NLP and Social Media Analytics*, 121–132. <https://doi.org/10.3233/faia241218>
- Zhou, X., Wan, F., Zhang, Y., & Qiu, S. (2024). Desain dan optimasi sistem rekomendasi cerdas untuk layanan ritel instan berbasis pembelajaran mesin. *Frontiers in Artificial Intelligence and Applications*, 241098. <https://doi.org/10.3233/faia241098>

Biodata Penulis



Shabrina Rasyid Munthe, S.Pd.I., MS, lahir di Rantauprapat pada 20 Februari 1992. Beliau meraih gelar S1 di bidang Tarbiyah dari Institut Agama Islam Negeri Padangsidempuan dan S2 Magister Sastra dari Universitas Islam Sumatera Utara (UISU). Sejak 2019, beliau menjabat dosen tetap di Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Al Washliyah Labuhanbatu. Shabrina juga aktif berkontribusi dalam pengembangan ilmu pengetahuan dan pendidikan, khususnya di bidang bahasa Inggris.



Samsir, ST, M.Kom., adalah penulis dan dosen di Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Alwashliyah (UNIVA), serta menjabat sebagai Dekan. Ia meraih gelar Sarjana (S1) dari Universitas Alwashliyah dan Magister (S2) dari Universitas Putra Indonesia "YPTK" Padang dengan konsentrasi Teknik Informasi. Saat ini, beliau sedang menempuh studi Doktorat (S3) di Universitas Negeri Padang, Program Studi Pendidikan Teknologi dan Kejuruan. Keahliannya berkontribusi besar dalam pengembangan ilmu komputer



Rina Asriani Levianti, SE, M.Ak., lahir di Kampung Mesjid, Kabupaten Aceh Barat, Provinsi Aceh, pada 17 Juli 1990. Beliau menyelesaikan pendidikan S1 dan S2 di Universitas Muhammadiyah Sumatera Utara dengan fokus pada Akuntansi. Saat ini, Rina menjabat sebagai dosen di Universitas Al Washliyah Labuhanbatu dan telah mengajar lebih dari empat tahun. Sebagai pengajar berpengalaman, ia membimbing mahasiswa dalam berbagai aspek akuntansi, baik teori maupun praktik. Untuk menghubungi, email di rinaasrianilevianti@gmail.com.



Dr. Basyarul Ulya, SH., MM., lahir di Janji pada 26 April 1986. Beliau menyelesaikan studi S1 di Fakultas Hukum UISU pada 2008, kemudian melanjutkan pendidikan Magister Manajemen di Universitas Islam Jakarta (UIJ) pada 2012. Pada 2017, beliau meraih gelar Doktor di bidang Ekonomi dari UINSU. Sejak 2017, Dr. Basyarul Ulya menjabat sebagai Rektor Universitas Al Washliyah Labuhanbatu, berkomitmen memajukan pendidikan dan pengembangan institusi di bidang hukum dan ekonomi.



Nurhanna Harahap, M.Hum., lahir di Teluk Rampah pada 26 Agustus 1986. Beliau menyelesaikan pendidikan Sarjana Pendidikan di Universitas Muslim Nusantara Al Washliyah, Medan, pada 2009 dengan program studi Pendidikan Bahasa Inggris. Pada 2015, ia meraih gelar Magister Humaniora di Universitas Negeri Medan dengan program studi Linguistik Terapan Bahasa Inggris. Saat ini, Nurhanna Harahap menjabat sebagai dosen tetap di Universitas Al Washliyah Labuhanbatu, berkontribusi dalam pengajaran bahasa Inggris dan linguistik.



Abdul Hakim Dalimunthe, M.Kom, NIDN 0127067801, lahir di Rantauprapat pada 27 Juni 1978. Beliau adalah dosen di Fakultas Ilmu Komputer (Filkom), dengan keahlian mendalam dalam jaringan internet dan teknologi informasi. Selain mengajar, Abdul Hakim aktif menyebarkan literasi digital kepada masyarakat luas, mengedukasi pentingnya keamanan dan etika dalam dunia digital. Pencinta kopi ini berkomitmen untuk mencerdaskan generasi digital melalui karya-karya yang bermanfaat, mengajarkan nilai pengetahuan dan tanggung jawab di era teknologi.



M. Rusdi, S.Pd.I, MM, lahir di Bomban Bidang pada 13 September 1985. Beliau meraih gelar Sarjana Pendidikan Islam pada 2008 di STAI Al Washliyah Labuhanbatu, dengan program studi Manajemen. Pada 2022, ia meraih gelar Magister Manajemen di Universitas Islam Sumatera Utara, Medan. Saat ini, M. Rusdi menjabat sebagai dosen tetap di Universitas Al Washliyah Labuhanbatu, berkomitmen mengembangkan pendidikan di bidang manajemen dan berperan aktif dalam mencetak generasi yang kompeten di bidang tersebut.



Nurhabibah Ritonga, S.E., M.M., lahir di Rantauprapat pada 4 September 1987. Beliau lulus S1 di Program Studi Manajemen STIE Labuhanbatu pada 2012 dan S2 di Magister Manajemen UISU Medan pada 2017. Dengan pengalaman mengajar selama 6 tahun, saat ini beliau menjabat dosen tetap di Program Studi Manajemen Fakultas Ekonomi Universitas Al Washliyah Labuhanbatu. Mengampu mata kuliah Manajemen Sumberdaya Manusia, Kewirausahaan, dan Studi Kelayakan Bisnis, beliau juga aktif menulis dan menjadi narasumber pelatihan kewirausahaan.



Ibnu Rasyid Munthe, ST., M.Kom., lahir di Rantau Prapat, Kabupaten Labuhanbatu, pada 1987. Beliau menyelesaikan pendidikan S1 di Universitas Nurtanio Bandung dan S2 di Universitas Putra Indonesia (YPTK) Padang. Aktif menulis artikel ilmiah dan buku, Ibnu Rasyid Munthe saat ini mengajar di Universitas Labuhanbatu. Selain itu, beliau mengelola kanal YouTube @manjaddawajada2022 sebagai media pembelajaran dan berbagi pengetahuan, memberikan kontribusi dalam dunia pendidikan dan teknologi.



ISBN 978-634-04-2224-5 (PDF)



9

786340

422245