KNOWLEDGE DISCOVERY TERHADAP SENTIMEN PELANGGAN KOPI KENANGAN MENGGUNAKAN NAÏVE BAYES

Zahra Diva Putri Munaspin, Nuke Merisca Titiana, Laila Tsabitah, Dzakiah Aulia Karima, Ken Ditha Tania, Allsela Meiriza
Program Studi Sistem Informasi, Universitas Sriwijaya
Jl. Raya Palembang-Prabumulih No. KM 32 Inderalaya, Kabupaten Ogan Ilir, Sumatera Selatan (30662)
09031182227019@student.unsri.ac.id

ABSTRAK

Perkembangan industri minuman kopi di Indonesia semakin pesat, salah satunya ditandai dengan banyaknya kedai kopi modern yang bermunculan, seperti Kopi Kenangan. Seiring meningkatnya jumlah pelanggan, opini atau sentimen pelanggan terhadap produk dan layanan menjadi aspek penting yang perlu dianalisis untuk menjaga kualitas dan kepuasan pelanggan. Permasalahan yang diangkat dalam penelitian ini adalah banyaknya data ulasan pelanggan di internet yang belum dimanfaatkan secara maksimal untuk mengevaluasi sentimen terhadap Kopi Kenangan. Penelitian ini bertujuan untuk mengklasifikasikan sentimen pelanggan Kopi Kenangan menggunakan algoritma Naïve Bayes, serta memberikan wawasan terhadap kecenderungan opini pelanggan secara otomatis. Metode yang digunakan adalah *knowledge discovery in database* (KDD) yang mencakup tahapan: *data selection*, *preprocessing*, *transformation*, *data mining* menggunakan algoritma Naïve Bayes, dan evaluasi model. Data diperoleh dari ulasan pelanggan Kopi Kenangan melalui *platform Google Review*, kemudian diklasifikasikan menggunakan model Naïve Bayes dengan pembagian data latih dan uji sebesar 80:20. Hasil dari penelitian menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen dengan akurasi sebesar 72,24%, presisi sebesar 75,20%, *recall* 72,24%, dan *F1-score* 72,20%. Model dapat mengenali sentimen positif dan netral dengan cukup baik, meskipun masih mengalami kesulitan dalam mengklasifikasikan sentimen negatif. Hasil ini dapat dijadikan acuan bagi perusahaan dalam meningkatkan kualitas layanan dan produk berdasarkan opini pelanggan.

Kata kunci : *knowledge discovery, analisis sentimen, naïve bayes, ulasan pelanggan, kopi kenangan.*

1. PENDAHULUAN

Pada era digital, peran *knowledge discovery* semakin krusial dalam mengelola dan mendistribusikan pengetahuan secara sistematis, terutama dalam pengolahan data ulasan pelanggan di industri makanan dan minuman untuk mendukung keputusan strategis. Ulasan pelanggan memainkan peran penting dalam membentuk persepsi masyarakat terhadap suatu bisnis atau merek. Ulasan positif dapat meningkatkan loyalitas pelanggan dan menarik calon pelanggan baru, sedangkan ulasan negatif dapat berdampak pada citra merek dan mengurangi kepercayaan konsumen [1] . Mekanisme pengelolaan ulasan yang efektif menjadi aspek krusial bagi perusahaan dalam menjaga kualitas layanan dan produk yang ditawarkan.

Dalam industri minuman siap saji yang semakin kompetitif, memahami kepuasan pelanggan menjadi tantangan bagi perusahaan seperti Kopi Kenangan. Menganalisis setiap ulasan pelanggan secara individual merupakan pekerjaan yang memakan banyak waktu dan kurang efisien. Oleh karena itu, dibutuhkan metode yang lebih terstruktur untuk mengelola informasi dari ulasan pelanggan. Sehingga penerapan *knowledge discovery* dalam analisis sentimen dapat menjadi solusi efektif untuk mengolah data ulasan pelanggan secara otomatis. Analisis sentimen adalah proses yang bertujuan untuk menentukan isi dari data set yang berbentuk teks bersifat positif, negatif atau netral [2] . Dengan pendekatan ini, perusahaan dapat mengidentifikasi tren opini pelanggan dan mengambil tindakan yang lebih proaktif dalam meningkatkan kualitas layanan.

Penelitian sebelumnya telah banyak menerapkan algoritma Naïve Bayes dalam analisis sentimen ulasan pelanggan di berbagai industri. Sebagai contoh, [3] melakukan penelitian terhadap analisis sentimen ulasan restoran di Singapura dengan menggunakan algoritma Naïve Bayes, yang menghasilkan akurasi sebesar 73% dalam mengelompokkan ulasan ke dalam kategori positif dan negatif. Selain itu, penelitian oleh [4] menggunakan metode yang sama pada ulasan pelanggan TMLST *Coffee*, menghasilkan akurasi sebesar 82%, dengan *recall* positif 76% dan *recall* negatif 89%, menunjukkan kemampuan model dalam mengidentifikasi sentimen negatif dengan baik. Penelitian-penelitian ini memberikan gambaran mengenai efektivitas Naïve Bayes dalam analisis sentimen, namun penerapannya dalam konteks *knowledge discovery* untuk memperbaiki layanan bisnis seperti Kopi Kenangan masih perlu dieksplorasi lebih lanjut. Oleh karena itu, tujuan penelitian ini adalah untuk mengkaji penerapan *knowledge discovery* dalam pengelolaan data pelanggan melalui analisis sentimen menggunakan algoritma Naïve Bayes, guna memberikan pemahaman mendalam dan dapat meningkatkan kualitas layanan dan produk.

2. TINJAUAN PUSTAKA

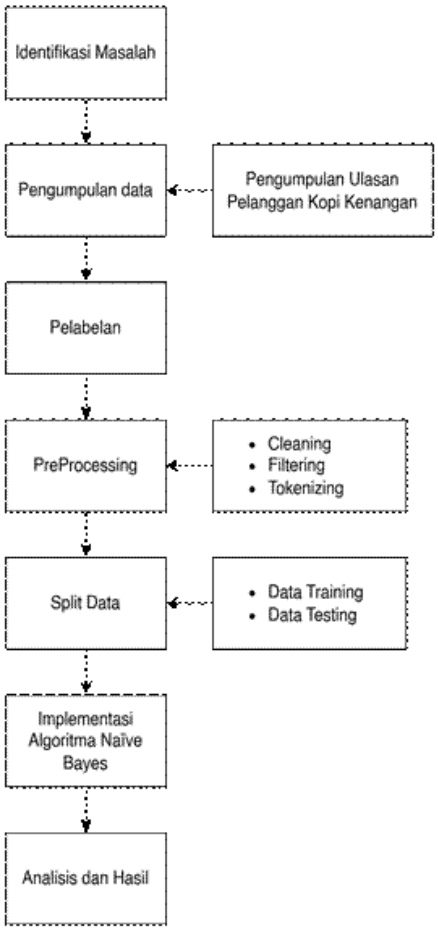
Dalam beberapa penelitian sebelumnya, algoritma Naïve Bayes telah banyak digunakan untuk analisis sentimen pada berbagai konteks, termasuk industri kuliner. Sebagai contoh, penelitian oleh [3] menggunakan Naïve Bayes untuk menganalisis ulasan restoran di Singapura, dan mencapai akurasi sebesar 73% dalam mengklasifikasikan sentimen positif dan negatif . Selain itu, penelitian oleh [4] pada TMLST *Coffee* juga mengadopsi Naïve Bayes sebagai metode klasifikasi utama, dengan hasil akurasi sebesar 82%, *recall* positif sebesar 76%, dan *recall* negatif sebesar 89%. Penelitian ini menyoroti pentingnya pemanfaatan algoritma Naïve Bayes dalam memahami kepuasan pelanggan melalui ulasan yang tersedia di *platform* seperti *Google Maps* dan *ShopeeFood* .
Penelitian serupa dilakukan oleh [5] yang mengkaji ulasan pelanggan Kafe Gamat menggunakan Naïve Bayes, dengan tahapan *preprocessing* data seperti tokenisasi, *stopword removal*, dan *stemming*. Penelitian ini berhasil mencapai akurasi sebesar 72%, meskipun *dataset* yang digunakan tidak seimbang, sehingga diterapkan teknik SMOTE untuk menangani masalah ini . Selain itu, penelitian [6] pada *Plus Coffee and Space* memanfaatkan Naïve Bayes untuk mengklasifikasikan ulasan berdasarkan aspek layanan, makanan, dan lingkungan, dan menghasilkan akurasi rata-rata sebesar 86%. Temuan ini menunjukkan bahwa Naïve Bayes adalah alat yang efisien dan andal dalam melakukan analisis sentimen di industri makanan dan minuman yang Secara khusus, pendekatan ini memprediksi peluang masa depan dengan memanfaatkan data historis [7].

Meskipun berbagai penelitian telah membuktikan efektivitas algoritma Naïve Bayes dalam analisis sentimen ulasan pelanggan, belum ada penelitian yang secara eksplisit mengaitkan hasil analisis sentimen ini dengan penerapan *knowledge discovery*. *Knowledge discovery* dapat membantu mengelola dan memanfaatkan pengetahuan yang terkandung dalam ulasan pelanggan untuk mendukung pengambilan keputusan strategis. Penelitian ini bertujuan untuk mengisi celah tersebut dengan menerapkan Naïve Bayes dalam analisis sentimen ulasan pelanggan Kopi Kenangan, sekaligus menghubungkannya dengan penerapan *knowledge discovery*, sehingga hasil yang diperoleh dapat mendukung pengambilan keputusan yang lebih baik di bidang layanan dan pengembangan produk.

3. METODE PENELITIAN

3.1. Skema Alur Penelitian

Penelitian ini memanfaatkan pendekatan kuantitatif dengan metode klasifikasi berbasis algoritma Naïve Bayes untuk menganalisis sentimen ulasan pelanggan terhadap Kopi Kenangan. Algoritma Naive bayes merupakan salah satu Algoritma yang sering digunakan pada penelitian dalam upaya mengatasi klasifikasi pada teks dengan mengunakan perhitungan probabilitas [1] . Dalam penelitian ini, pendekatan *Knowledge Discovery* digunakan untuk mengelola data dan informasi secara terstruktur. Dengan pendekatan ini, proses ekstraksi pengetahuan dari data ulasan dapat dilakukan secara lebih efisien, sehingga memberikan wawasan yang lebih mendalam mengenai pemahaman sentimen pelanggan. Berikut adalah diagram metode penilitian:



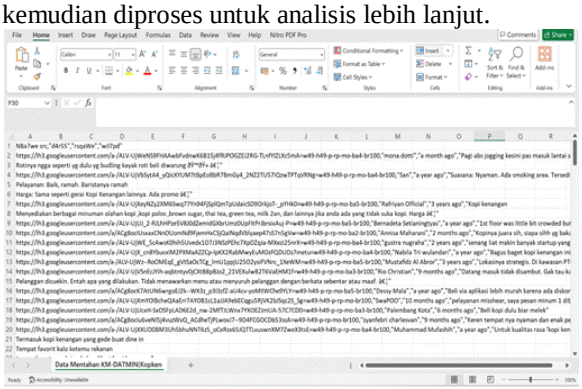
Gambar 1. Diagram Tahapan Penelitian

Penelitian ini diawali dengan mengidentifikasi permasalahan utama yaitu sulitnya menganalisis dan memanfaatkan ratusan ulasan pelanggan Kopi Kenangan secara manual. Untuk mengatasi hal tersebut, peneliti mengumpulkan data ulasan dari *platform Google Review*, kemudian memberikan label sentimen pada setiap ulasan sebagai positif, negatif, atau netral. Setelah proses pelabelan, data melalui tahap *preprocessing* yang terdiri dari tiga langkah yaitu pembersihan data (*cleaning*), penyaringan (*filtering*), dan tokenisasi (*tokenizing*). Data yang telah diproses selanjutnya dibagi menjadi data latih dan data uji dengan rasio 80:20. Representasi kata diubah menjadi bentuk vektor menggunakan metode TF-IDF dan model klasifikasi Naïve Bayes diterapkan untuk mengidentifikasi sentimen. Pada tahap akhir, performa model dianalisis menggunakan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score* pada data uji. Hasil evaluasi ini menjadi acuan bagi perusahaan

dalam meningkatkan kualitas layanan dan produk berdasarkan opini pelanggan.

3.2. *Web Scrapping*

Web scrapping merupakan metode yang digunakan untuk mengambil data dari sebuah situs web. Dalam penelitian ini, pengambilan data dilakukan secara manual dengan mengekstrak ulasan pelanggan Kopi Kenangan dari *platform Google Review* ke dalam *file* Excel. Data yang diperoleh kemudian diproses untuk analisis lebih lanjut.



Gambar 2. *Web Scrapping*

3.3. Studi Literatur

Studi literatur dilakukan dengan mengumpulkan dan menganalisis berbagai sumber literatur yang relevan, seperti jurnal ilmiah, artikel penelitian, sumber internet lainnya yang berkaitan dengan metode penelitian, tujuan serta rumusan masalah yang berkaitan dengan penelitian serupa yang dilakukan.

3.4. Analisa Data

Dalam penelitian ini, proses *Knowledge Discovery* diterapkan untuk menganalisis sentimen ulasan pelanggan Kopi Kenangan yang mencakup beberapa tahapan, mulai dari pemilihan data (*Selection*), pemrosesan awal (*Preprocessing*), transformasi data (*Transformation*), penerapan teknik penambangan data (*Data Mining*), serta evaluasi hasil (*Evaluation*).

3.5. *Data Selection*

Data ulasan pelanggan Kopi Kenangan dikumpulkan dari platform *Google Review*. Pemilihan data dilakukan berdasarkan relevansi dan kelengkapan informasi, sementara ulasan yang tidak relevan, kosong, atau berisi *spam* dihapus untuk memastikan hasil analisis lebih akurat. Data hasil seleksi ini kemudian disimpan dalam *dataset* khusus yang terpisah dari basis data operasional sebelum memasuki tahap pemrosesan lebih lanjut.

3.6. *Preprocessing*

Preprocessing merupakan salah satu Teknik dalam *data mining* untuk melakukan transformasi pada data mentah agar menjadi format yang bisa lebih mudah dimengerti [8]. Pemrosesan awal mencakup cleaning untuk menghapus *noise*, *filtering* untuk menyaring kata tidak relevan, dan *tokenizing* untuk memecah teks ulasan. Terakhir, *enrichment* menambahkan informasi seperti tanggal ulasan dan *rating* pelanggan guna meningkatkan akurasi analisis.

3.7. *Transformation*

Pencarian fitur yang relevan dilakukan agar data lebih representatif terhadap tujuan analisis sentimen. Data ulasan dikonversi ke dalam format numerik menggunakan metode CSV (*Comma-Separated Values*) untuk mempermudah pemrosesan oleh model Naïve Bayes. Konversi ini memastikan bahwa data terstruktur dengan baik dan siap untuk diklasifikasikan. Selain itu, standarisasi diterapkan untuk menjaga konsistensi setiap fitur, sehingga hasil analisis lebih akurat dan tidak terpengaruh oleh perbedaan skala.

3.8. *Data Mining*

Pada tahap ini, data yang telah dikumpulkan digunakan untuk membentuk pola klasifikasi melalui *data training* menggunakan metode klasifikasi. Pola tersebut kemudian diterapkan pada *data testing* untuk mengevaluasi performa metode yang digunakan. Penelitian ini menggunakan algoritma Naïve Bayes untuk mengklasifikasikan sentimen ulasan pelanggan Kopi Kenangan menjadi positif dan negatif berdasarkan probabilitas kemunculan kata.

3.9. *Evaluation*

Tahap ini melibatkan pemeriksaan apakah pola atau informasi yang diperoleh sesuai atau bertentangan dengan fakta atau hipotesis yang telah ada sebelumnya. Terdapat tiga nilai yang digunakan untuk menilai kemampuan sistem klasifikasi yang dibangun, yaitu *precision*, *recall*, dan *accuracy*. Ada beberapa metode yang dapat digunakan untuk mengukur performa klasifikasi, namun dalam evaluasi ini, akan dilakukan penghitungan akurasi, *precision*, *recall*, dan *F1-Score* untuk menyeimbangkan nilai presisi dan juga *recall*. Akurasi adalah persentase keseluruhan klasifikasi yang berhasil diidentifikasi dengan benar. Hasil performa perhitungan ini akan disajikan melalui *Confusion Matrix*. Akurasi dihitung dengan membagi jumlah data yang berhasil diprediksi dengan benar oleh model dengan total data uji. Nilai akurasi dapat diperoleh menggunakan suatu persamaan, seperti yang ditunjukkan dalam Persamaan (1).

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Selain akurasi, kinerja pengklasifikasi juga dapat diukur menggunakan presisi dan *recall*. Presisi mengukur sejauh mana hasil prediksi sesuai dengan kejadian sebenarnya, memastikan relevansi data yang ditemukan. Presisi dihitung sebagai rasio antara data relevan yang berhasil diidentifikasi dengan jumlah keseluruhan data yang ditemukan, sesuai dengan Persamaan (2).

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (2)$$

Recall merupakan metrik lain yang digunakan untuk menilai kinerja pengklasifikasi. *Recall* mengukur seberapa banyak data relevan yang berhasil ditemukan dibandingkan dengan total data relevan yang ada. Nilainya dihitung dengan membagi jumlah data benar bernilai positif dengan total data benar bernilai positif ditambah data salah bernilai negatif, sesuai dengan Persamaan (3).

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

F1-score merupakan nilai rata-rata dari penjumlahan presisi dan *recall*. *F1-score* digunakan untuk menyeimbangkan nilai presisi dan juga *recall*. Persamaan 4 dapat digunakan untuk menghitung *F1-score*.

$$F1\text{-}Score = 2\,\frac{Precision\ x\ Recall}{Precision + Recall} \quad (4)$$

Keterangan:

a. TP (*True Positive*) adalah jumlah data dengan nilai aktual positif dan diklasifikasikan dengan benar sebagai positif.
b. TN (*True Negative*) adalah jumlah data dengan nilai aktual negatif dan diklasifikasikan dengan benar sebagai negatif.
c. FN (*False Negative*) adalah jumlah data dengan nilai aktual negatif, tetapi salah diklasifikasikan sebagai positif.
d. FP (*False Positive*) adalah jumlah data dengan nilai aktual positif tetapi salah diklasifikasikan sebagai negatif.

4. HASIL DAN PEMBAHASAN

4.1. *Data Selection*

Kumpulan data ulasan pelangaan diambil dari *platform* Google Review. Kopi Kenangan mendapatkan sejumlah data ulasan dari pelanggan yang telah mengunjungi gerainya, namun hanya terdapat 235 ulasan yang berisi teks dan digunakan dalam penelitian ini. Ulasan tersebut mencakup sentimen positif dan negatif, yang selanjutnya diinput dan diberi label secara manual. Setelah proses pelabelan, data ulasan disimpan dalam Microsoft Excel. Tabel 1 memperlihatkan hasil penginputan data yang diambil dari Google Review.

Tabel 1. *Data Selection*

No	Text
1	Kopinya & Cookies nya enak, pelayanan juga ramah & cepat. Parkiran tidak begitu luas. Suasananya cukup nyaman meskipun tempatnya kecil, tersedia ruangan bagi perokok di lantai 2.
2	Tempatnya luas dan banyak sitting area. Nyaman dan kopinya pasti enak
3	Nyaman, rasanya jg enak, Rapi jg Barista nya juga Ramah
4	Pelayanan: Baik, ramah. Baristanya ramah
5	Tempat ngopi dan minum coklat dengan harga murah. Enak buat nongkrong sambil nyantai

4.2. *Preprocessing*

Dalam tahap *preprocessing*, data ulasan yang telah dikumpulkan dan diberi label diproses melalui empat langkah agar siap digunakan pada tahap berikutnya. Proses *preprocessing* ini dilakukan dengan bantuan *tools* Google Colaboratory menggunakan bahasa pemrograman *Python*. Langkah-langkah tersebut dimulai dengan normalisasi, yaitu proses mengubah teks menjadi bentuk standar agar lebih mudah dianalisis. Selanjutnya dilakukan *stopword* yaitu, proses yang akan menghapus kata umum yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna. Pada tahap ini menggunakan *library* Sastrawi untuk melakukan *Stopword Removal* [9]. Setelah itu dilakukan langkah *tokenizing* yang merupakan tahapan membuat setiap kata menjadi berurutan dari suatu kalimat, setiap kata akan di pisah-pisah atau dilakukan separasi [10]. Dan langkah *stemming*, yaitu tahapan yang digunakan untuk mengkonversikan kata-kata yang berimbuhan menjadi kata dasar atau menghilangkan imbuhan [5]. Hasil *preprocessing* dapat dilihat pada Tabel 2.

Tabel 2. Hasil *Preprocessing*

Text	Sentiment
Untuk kualitas rasa kopi kenangan selalu konsisten disetiap kedai nya	normalisasi : kualitas rasa kopi kenangan selalu konsisten disetiap kedai nya
Untuk kualitas rasa kopi kenangan selalu konsisten disetiap kedai nya	*stopwords* : kualitas rasa kopi kenangan selalu konsisten disetiap kedai nya
Untuk kualitas rasa kopi kenangan selalu konsisten disetiap kedai nya	*tokenize* : [kualitas, rasa, kopi, kenangan, selalu, konsisten, disetiap, kedai, nya]
Untuk kualitas rasa kopi kenangan selalu konsisten disetiap kedai nya	*stemming* : kualitas rasa kopi kenang selalu konsisten tiap kedai nya

4.3. *Transformation*

Pada tahap ini, dilakukan transformasi data teks ulasan pelanggan menggunakan metode *Term Weighting*, yakni *Term Frequency-Inverse Document Frequency* (TF-IDF).

Tabel 3. Menghitung TF-IDF

Kosakata	TF			df	N/df	Idf
	d1	d2	d3			
enak	1	1	1	3	1	0
kopi	0	1	0	1	3	0,477
ramah	1	0	1	2	1,5	0,176
nyaman	1	1	1	3	1	0
pelayanan	1	0	0	1	3	0,477
luas	1	1	0	2	1,5	0,176
tempatnya	1	1	0	2	1,5	0,176

Data ulasan pelanggan diubah menjadi nilai numerik dengan menggunakan metode TF-IDF untuk menilai pentingnya setiap kata dalam dokumen. Bobot dihitung berdasarkan seberapa sering kata muncul dalam dokumen tertentu (TF) dan seberapa jarang kata tersebut muncul di seluruh kumpulan dokumen (IDF). Pada Tabel 3, disajikan contoh perhitungan manual TF-IDF untuk kata-kata seperti "enak," "kopi," dan "ramah," guna mengetahui relevansi setiap kata dalam dokumen.

4.4. *Data Mining*

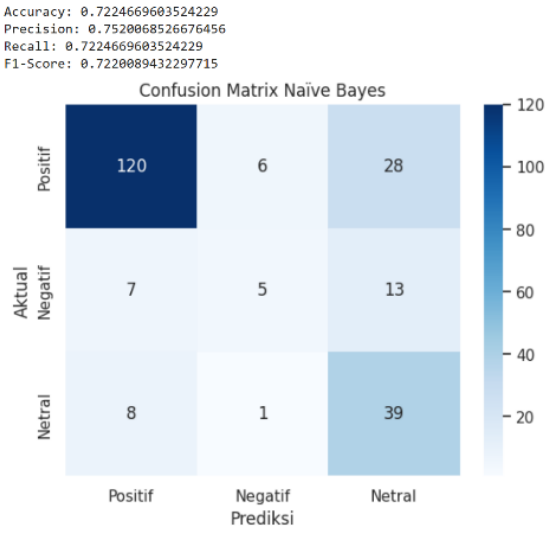
Pada tahap ini, data yang telah menjalani proses *preprocessing* atau pembersihan akan dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan rasio 80:20 dengan total data 227 ulasan. Data latih digunakan untuk membangun model *machine learning*, sementara data uji berfungsi untuk mengevaluasi kinerja model yang telah dibuat dirincikan pada Tabel 4.

Tabel 4. Pembagian *Split Data*

Data	Rasio	Total
Data Training	80%	182
Data Testing	20%	45

4.5. *Evaluation*

Evaluasi performa model *Naïve Bayes* dalam klasifikasi sentimen dilakukan dengan menggunakan beberapa metrik, yaitu *accuracy*, presisi, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa model memiliki akurasi sebesar 72,24%, yang mengindikasikan bahwa 72,24% dari total data uji diklasifikasikan dengan benar oleh model. Selain itu, model memperoleh nilai presisi sebesar 75,20%, yang berarti dari seluruh prediksi kelas positif yang dihasilkan oleh model, sekitar 75,20% merupakan prediksi yang benar. Adapun nilai *recall* sebesar 72,24% menunjukkan bahwa model berhasil mengidentifikasi 72,24% dari total sampel yang seharusnya termasuk dalam kelas positif. Kombinasi antara presisi dan *recall* menghasilkan *F1-score* sebesar 72,20%, yang mencerminkan keseimbangan antara kedua metrik tersebut dan menunjukkan tingkat keandalan model dalam klasifikasi sentimen.



Gambar 3. Hasil Pengujian *Confusion Matrix*

Berdasarkan *confusion matrix* pada Gambar 3, model berhasil mengklasifikasikan 120 sampel kelas positif dengan benar. Namun, terdapat 6 sampel kelas positif yang salah diklasifikasikan sebagai negatif dan 28 sampel lainnya yang dikategorikan sebagai netral. Pada kelas negatif, model hanya mampu mengidentifikasi 5 sampel dengan benar, sementara 7 sampel justru diklasifikasikan sebagai positif dan 13 sampel lainnya sebagai netral. Adapun untuk kelas netral, model dapat mengklasifikasikan 39 sampel dengan tepat, meskipun masih terdapat 8 sampel yang keliru dikategorikan sebagai positif dan 1 sampel lainnya sebagai negatif.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa model *Naïve Bayes* memiliki performa yang cukup baik dalam mengklasifikasikan sentimen. Model telah mampu mengidentifikasi sebagian besar sampel dengan benar, terutama pada kelas positif dan netral. Hasil analisis evaluasi akurasi, presisi, *recall* dengan metode Naïve Bayes dapat dilihat pada Tabel 5.

Tabel 5. Hasil Data Uji Evaluasi Naïve Bayes

Keterangan	Hasil Analisis
Data Training	80%
Data Testing	20%
True Positive	120
False Positive	15
False Negative	34
True Negative	58
Precision	75,20%
Recall	72,24%
F1-Score	72,20%
Accuracy	72,24%

4.6. *Visualization*

Visualisasi menggunakan fitur *Wordcloud* pada *Google Colab* menampilkan gambaran kata-kata yang paling sering muncul dalam ulasan pelanggan mengenai Kopi Kenangan. *Wordcloud* ini dibuat

dengan menggunakan pustaka *wordcloud* dan *matplotlib* dalam *Python*, yang memungkinkan representasi visual dari keseluruhan komentar pelanggan. Dengan teknik ini, kata-kata yang memiliki frekuensi tinggi akan ditampilkan dengan ukuran lebih besar, sehingga dapat memberikan wawasan mengenai aspek-aspek yang paling banyak dibahas oleh pelanggan. Hasil visualisasi dari data ulasan tersebut dapat dilihat pada Gambar 4.

Gambar 4. Hasil *Wordcloud*

5. KESIMPULAN DAN SARAN

Hasil penelitian menunjukkan bahwa metode Naïve Bayes efektif dalam melakukan analisis sentimen terhadap ulasan pelanggan Kopi Kenangan di Google Review, dengan tingkat akurasi dan *recall* sebesar 72,24%. Metode ini mampu mengklasifikasikan opini menjadi sentimen positif dan negatif serta unggul dalam kecepatan pemrosesan data berukuran besar. Namun, kelemahan masih ditemukan dalam menangani teks yang kompleks serta keterbatasan data dari satu sumber yang mempengaruhi representasi opini pelanggan secara menyeluruh. Oleh karena itu, disarankan agar penelitian selanjutnya mengeksplorasi metode lain seperti SVM atau *Random Forest*, serta memanfaatkan data dari berbagai platform media sosial untuk meningkatkan keberagaman data. Pengembangan sistem analisis sentimen berbasis *real-time* dan penerapan pendekatan *Deep Learning* juga direkomendasikan untuk meningkatkan pemahaman konteks serta akurasi klasifikasi opini pelanggan.

DAFTAR PUSTAKA

[1] P. Hartini *et al.*, "Naive Bayes Classifier untuk Analisis Sentimen Ulasan Pelanggan pada Domo Coffee and Resto," *Jurnal Informatika dan Rekayasa Perangkat Lunak* , vol. 6, no. 1, 2024.

[2] A. Saputra and F. Noor Hasan, "ANALISIS SENTIMEN TERHADAP APLIKASI COFFEE MEETS BAGEL DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 2, pp. 465–474, Jan. 2023, doi: 10.54443/sibatik.v2i2.579.

[3] V. A. Permadi, "Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Review Restoran di Singapura 141," *Jurnal Buana Informatika*, vol. 11, no. 2, pp. 141–151, 2020, [Online]. Available: https://www.kaggle.com/hj5992/restaurantreviews

[4] D. A. Hamidah, R. Salkiawati, and R. Sari, "Analisis Sentimen Ulasan Customer Kopi TMLST Menggunakan Algoritma Naïve Bayes," *Journal of Students' Research in Computer Science*, vol. 5, no. 1, pp. 27–40, May 2024, doi: 10.31599/mrm89y71.

[5] A. S. Adhityas, S. Satya Nugraha, and M. A. Syuhada, "Analisis Sentimen Ulasan Pelanggan pada Kafe Gamat Bendungan Hilir Menggunakan Algoritma Naive bayes," *CENTIVE*, vol. 4, no. 1, pp. 776–785, 2024.

[6] D. A. Firmansah, N. Y. Setiawan, and D. E. Ratnawati, "Analisis Ulasan Pelanggan Menggunakan Naïve Bayes Classifier pada Plus Coffee and Space," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 3, 2025, [Online]. Available: http://j-ptiik.ub.ac.id

[7] Hajaroh, T. Suprapti, and R. Narasati, "IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN ULASAN PRODUK MAKANAN DAN MINUMAN DI TOKOPEDIA," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 1, 2024.

[8] N. A. Agustina, D. H. Citra, W. Purnama, C. Nisa, and A. R. Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM:Indonesian Journal of Machine Learning and Computer Science*, vol. 2, pp. 47–54, 2022.

[9] Nurzaman, N. Suarna, and W. Prihartono, "ANALISIS SENTIMEN ULASAN APLIKASI THREADS DI GOOGLE PLAYSTORE MENGGUNAKAN ALGORITMA NAÏVE BAYES," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 1, 2024.

[10] R. Maulana, A. Voutama, and T. Ridwan, "ANALISIS SENTIMEN ULASAN APLIKASI MYPERTAMINA PADA GOOGLE PLAY STOREMENGGUNAKAN ALGORITMA NBC," *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 42–48, 2023