

Implementasi Naïve Bayes Untuk Analisis Sentimen Barbershop

Muhammad Rafli Alfiansyah¹

¹Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri
E-mail: [**Abstrak** – Kebangkitan industri barbershop dan interaksinya yang semakin intens dengan pelanggan memunculkan beragam respons dan ulasan publik yang terekam melalui media sosial dan platform digital, sehingga diperlukan analisis sistematis untuk memahami kecenderungan opini tersebut. Penelitian ini menganalisis sentimen masyarakat terhadap layanan dan pengalaman di barbershop dengan menerapkan algoritma Naïve Bayes, yang dikenal efektif dalam mengolah teks berdimensi tinggi. Data diperoleh dari ulasan dan komentar pengguna pada platform digital yang membahas layanan barbershop tertentu, kemudian melalui tahapan text preprocessing dan pembobotan TF-IDF sebelum diklasifikasikan menjadi kategori positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa model Naïve Bayes mampu mengklasifikasikan sentimen ulasan dengan performa yang baik, dengan akurasi tertinggi pada kelas negatif sebesar 93,39%, diikuti kelas netral sebesar 89,26% dan kelas positif sebesar 85,95%. Nilai presisi tertinggi diperoleh pada kelas positif sebesar 96,81%, sementara kelas netral menunjukkan keterbatasan dengan nilai presisi dan recall sebesar 0%, yang mengindikasikan ketidakseimbangan data pada kelas tersebut. Temuan penelitian diharapkan menjadi dasar pertimbangan yang lebih objektif bagi pemilik dan pengelola barbershop dalam meningkatkan kualitas layanan dan merumuskan strategi pemasaran yang lebih responsif terhadap opini dan kebutuhan pelanggan](mailto:*¹raffialfian14@gmail.com</p></div><div data-bbox=)

Kata Kunci — analisis sentimen, naïve bayes, barbershop, TF-IDF

1. PENDAHULUAN

Perkembangan industri barbershop sebagai bagian dari usaha kecil dan menengah (UKM) menunjukkan pertumbuhan yang signifikan seiring perubahan gaya hidup masyarakat dan meningkatnya perhatian terhadap penampilan. Kondisi ini mendorong persaingan antar barbershop menjadi semakin ketat, sehingga kualitas layanan menjadi faktor strategis dalam mempertahankan kepuasan dan loyalitas pelanggan. Studi [1] mengungkap bahwa kualitas pelayanan barbershop dapat dianalisis melalui kesenjangan antara harapan dan persepsi konsumen menggunakan pendekatan *Service Quality* (SERVQUAL) dan *Importance Performance Analysis* (IPA), yang mampu mengidentifikasi atribut layanan prioritas untuk perbaikan.

Selain kualitas layanan, aspek harga juga berperan penting dalam membentuk loyalitas pelanggan. Hasil analisis regresi linier pada studi [2] menunjukkan bahwa kualitas pelayanan dan harga berpengaruh signifikan terhadap loyalitas konsumen barbershop, yang menjadi indikator keberlanjutan usaha dalam jangka panjang. Pendekatan lain melalui *Quality Function Deployment* (QFD) berbasis *Service Quality* dan *Kano Model* pada studi [3] berhasil mengelompokkan atribut layanan berdasarkan dampaknya terhadap kepuasan pelanggan serta menyusun usulan peningkatan kualitas layanan secara terstruktur.

Di sisi aplikatif, pengelolaan hubungan dengan pelanggan juga menjadi perhatian penting. Studi [4] menekankan bahwa penilaian konsumen dan respons aktif barbershop terhadap opini pelanggan, khususnya melalui media digital, berkontribusi dalam meningkatkan citra usaha dan memperluas jangkauan pemasaran. Namun, sebagian kajian pada sektor jasa masih berfokus pada pengembangan sistem secara teknis tanpa disertai evaluasi mendalam terhadap persepsi pengguna atau dampak bisnis yang dihasilkan, sebagaimana ditunjukkan pada pengembangan sistem informasi berbasis web yang lebih menitikberatkan aspek implementasi teknis dibandingkan analisis pengalaman pengguna [5].

Seiring meningkatnya penggunaan media sosial dan platform ulasan daring, pelanggan semakin aktif menyampaikan pengalaman dan penilaian mereka dalam bentuk teks tidak terstruktur. Data ini memiliki potensi besar untuk menggambarkan persepsi publik secara lebih objektif dan real-time. Beberapa penelitian telah menerapkan analisis sentimen berbasis *text mining* menggunakan algoritma Naïve Bayes pada berbagai domain, seperti layanan keuangan digital (paylater), dengan hasil klasifikasi sentimen yang cukup baik [6]. Keunggulan Naïve Bayes terletak pada kesederhanaan model dan efisiensi komputasi, meskipun masih memiliki keterbatasan dalam akurasi yang dapat ditingkatkan melalui optimasi dan seleksi fitur [7].

Penerapan analisis sentimen pada ulasan digital juga telah dilakukan pada sektor layanan publik dan pariwisata. Studi [8] memanfaatkan ulasan Google Maps untuk menganalisis sentimen pengguna rumah sakit

menggunakan *Naïve Bayes*, sedangkan studi [9] menerapkan pendekatan serupa pada ulasan pengunjung wisata heritage. Hasil kedua studi tersebut menunjukkan bahwa ulasan digital mampu merepresentasikan persepsi masyarakat secara lebih luas dibandingkan survei konvensional.

Untuk meningkatkan performa klasifikasi sentimen, beberapa penelitian mengombinasikan *Naïve Bayes* dengan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF). Evaluasi berbagai teknik ekstraksi fitur TF-IDF pada ulasan marketplace berbahasa Indonesia menunjukkan bahwa kombinasi TF-IDF dan *Naïve Bayes* menghasilkan performa klasifikasi yang kompetitif [10]. Pendekatan serupa juga terbukti efektif dalam meningkatkan akurasi analisis sentimen melalui optimasi fitur berbasis TF-IDF [11], termasuk pada klasifikasi ulasan aplikasi digital seperti MyXL [12] dan perbandingan dengan algoritma lain seperti *Support Vector Machine* (SVM) [13].

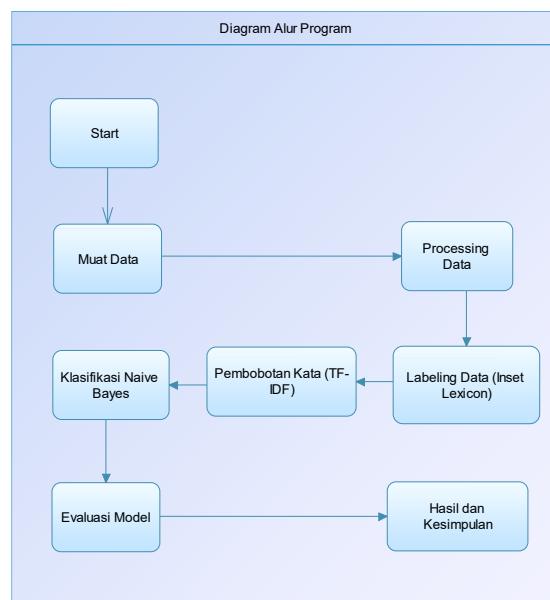
Meskipun analisis sentimen sudah banyak digunakan di berbagai bidang, pemanfaatan ulasan digital pelanggan untuk menilai sentimen terhadap layanan barbershop masih belum banyak dilakukan. Penelitian barbershop sebelumnya umumnya menilai kepuasan pelanggan melalui kuesioner atau penilaian layanan secara langsung, sehingga ulasan pelanggan dalam bentuk teks bebas belum dimanfaatkan secara maksimal. Padahal, ulasan digital dapat memberikan gambaran langsung mengenai pengalaman dan pendapat pelanggan terhadap layanan barbershop. Oleh karena itu, diperlukan penelitian yang mengolah ulasan digital pelanggan menggunakan teknik text mining dan algoritma klasifikasi untuk mengetahui sentimen pelanggan secara lebih jelas.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan algoritma *Naïve Bayes* dengan tahapan *text preprocessing* dan pembobotan TF-IDF untuk mengklasifikasikan sentimen pelanggan barbershop ke dalam kategori positif, negatif, dan netral. Pendekatan ini diharapkan mampu memberikan gambaran yang lebih objektif mengenai persepsi masyarakat terhadap aspek-aspek utama layanan barbershop, seperti kualitas potongan rambut, kenyamanan tempat, dan keramahan staf, sehingga dapat menjadi dasar pengambilan keputusan yang lebih tepat bagi pemilik dan pengelola barbershop dalam meningkatkan kualitas layanan dan strategi pemasaran.

2. METODE PENELITIAN

2.1 Diagram Alur Program

Program dalam penelitian ini mengikuti alur kerja sistematis yang terdiri dari beberapa tahapan utama, yaitu: muat data, processing data, labelling data, pembobotan kata, klasifikasi *naïve bayes*, evaluasi model, hasil dan kesimpulan. Alur tersebut divisualisasikan dalam Gambar 1 untuk memberikan gambaran menyeluruh mengenai proses program.



Gambar 1 Diagram Alur Program

2.2 Muat Data

Tahap muat data merupakan tahap awal dalam alur program penelitian ini. Pada tahap ini, data ulasan pelanggan dimasukkan ke dalam sistem sebagai bahan analisis. Data yang digunakan berasal dari ulasan Courteous Barbershop pada Google Maps dan diperoleh dengan teknik web *scraping* menggunakan modul *Selenium* pada bahasa pemrograman Python. Ulasan yang berhasil diambil kemudian disimpan dalam bentuk file CSV (*Comma Separated Values*) agar dapat diproses lebih lanjut. Sebanyak 121 data ulasan digunakan sebagai data awal sebelum masuk ke tahap *preprocessing* dan analisis sentimen.

2.3 Text preprocessing

Sebelum dilakukan analisis lebih lanjut, perlu menghilangkan serta mengatasi noisy data agar hasil perhitungan optimal. Maka dilakukan alur proses text preprocessing [14] antara lain, *cleaning*, *case folding*, *stopword removal*, *tokenizing*, dan *stemming*. setelah data melalui proses *preprocessing* maka data tersebut adalah data bersih yang akan dilanjut ke proses selanjutnya yaitu labelling inset *lexicon*. Alur diagram *preprocessing* dapat dilihat sebagai berikut.



Gambar 2 diagram alur *preprocessing* data

a. Cleaning

Cleaning adalah proses pembersihan data dari emoji, url, symbol, dan tanda baca tetapi tidak menghapus huruf atau angka.

b. Case Folding

Case Folding Mengubah huruf menjadi kecil semua (*lower case*).

c. Stopword removal

Stopword removal digunakan untuk menghilangkan kata umum yang tidak diperlukan atau dianggap tidak penting.

d. Tokenizing

Tokenizing adalah proses mengubah kata menjadi unit seperti token dan proses ini penting untuk menentukan polaritas positif, negative, dan netralnya sebuah teks agar bisa dibuat lexicon manual nya.

e. Stemming

Sebagai langkah awal pada *stemming*, digunakan suatu library khusus untuk *stemming* pada bahasa pemrograman python supaya memudahkan dalam mengimplementasikannya. Library khusus yang digunakan adalah library sastrawi dan NLTK dengan algoritma Nazief dan Andriani didalam nya. Cara kerja algoritma ini adalah setiap kata yang ada dilakukan proses stemming [15].

2.4 Labelling Inset Lexicon

Inset lexicon digunakan untuk mengidentifikasi opini tertulis dan mengkategorikannya menjadi opini positif negative, dan netral yang bisa digunakan untuk menganalisis sentimen ulasan publik terhadap topik, acara, atau produk tertentu [16].

2.5 Pembobotan kata TF-IDF

Setelah serangkaian *preprocessing* pada data, setiap kata (*term*) diberi nilai bobot menggunakan metode *Term Frequency — Inverse Document Frequency*. Proses pembobotan ini penting karena nilai yang dihasilkan oleh TF-IDF akan berfungsi sebagai fitur input utama dalam tahap selanjutnya, yaitu proses klasifikasi [17].

$$TF - IDF = TF(t, d) \times IDF(t, D) \dots \dots (1)$$

Keterangan :

$$TF \times IDF = (\text{seberapa sering kata muncul}) \times (\text{seberapa langka kata itu muncul})$$

2.6 Klasifikasi *Naïve Bayes*

Setelah tahap *preprocessing* dan pemberian bobot menggunakan TF-IDF selesai, data teks telah diubah menjadi vektor numerik yang siap diolah. Vektor ini kemudian masuk ke dalam tahap analisis menggunakan algoritma *Naïve Bayes* [18].

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \dots \dots (2)$$

2.7 Evaluasi

Tahap evaluasi dalam penelitian ini bertujuan mengukur performa model yang diusulkan. Alat utama yang digunakan untuk evaluasi adalah Confusion Matrix, sebuah tabel yang membandingkan hasil klasifikasi prediksi sistem dengan hasil klasifikasi yang sebenarnya (aktual). Dari *Confusion Matrix* ini [19], empat metrik dasar True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) akan diperoleh. Nilai-nilai ini selanjutnya dipakai untuk menghitung metrik kinerja utama seperti *akurasi*, *presisi*, *recall*, dan *f1-score*.

$$Accuracy = \frac{(TP+FN)}{(TP+FP+FN+TN)} \dots\dots\dots (3)$$

$$Precision = \frac{(TP)}{(TP+FP)} \dots\dots\dots (4)$$

$$Recall = \frac{(TP)}{(TP+FN)} \dots\dots\dots (5)$$

$$f1-score = 2 * \frac{precision*recall}{precision+recall} \dots\dots\dots (6)$$

3. HASIL DAN PEMBAHASAN

3.1 Preprocessing

a) Cleaning

Cleaning dilakukan untuk pembersihan data dari emoji, url, symbol, dan tanda baca seperti yang dapat dilihat pada table 1.

Tabel 1 Cleaning

Ulasan	Cleaning
potongan oke, tapi hasil wavy permnya bikin nangis 🥺 belum keluar dari barbershopnya aja rambut bener bener kering ga berbentuk padahal sebelumnya rambut normal. asli jauh banget dari ekspektasi.... Potongannya rapi bgt gilss 🔥 🔥 🔥 Baru disini yg cocok buat cowo gw, yg lain selalu kependekan dan ga rapi, cuma ya emang ga free cukur jenggot wkwk ...	potongan oke tapi hasil wavy permnya bikin nangis belum keluar dari barbershopnya aja rambut bener bener kering ga berbentuk padahal sebelumnya rambut normal asli jauh banget dari ekspektasi Potongannya rapi bgt gilss Baru disini yg cocok buat cowo gw yg lain selalu kependekan dan ga rapi cuma ya emang ga free cukur jenggot wkwk

b) Case Folding

Proses *case folding* dilakukan Untuk menghilangkan perbedaan representasi kata yang disebabkan oleh penggunaan huruf kapital. Mengubah semua yang merupakan huruf kapital menjadi huruf kecil. Hasil dari *case folding* dapat dilihat pada Tabel 2.

Tabel 2 Case folding

Ulasan	Case Folding
Jangan Lupa Buat Follow IG Atau TikTok nya Buat Tau Update an Dan Stay Tuned Tentang Barber Ini Dan Juga Untuk Pastikan Reservasi Dan Konsul Di Admin Yaa Kmrin nyobain treatment keratin smooth dsni dari cut rambut dan keratinnya bagus bngt hasil nya rapi sangat detail sekali dan tenaga barbernya satu persatu menunjukkan keahlian mereka didepan customer nya dngan sangat komunikatif	jangan lupa buat follow ig atau tiktok nya buat tau update an dan stay tuned tentang barber ini dan juga untuk pastikan reservasi dan konsul di admin yaa kmrin nyobain treatment keratin smooth dsni dari cut rambut dan keratinnya bagus bngt hasil nya rapi sangat detail sekali dan tenaga barbernya satu persatu menunjukkan keahlian mereka didepan customer nya dngan sangat komunikatif

c) Stopword removal

Proses ini bertujuan untuk membuang kata-kata yang tidak relevan atau tidak memengaruhi polaritas sentimen. Hal ini dicapai melalui pencocokan kata-kata dalam data *tweet* dengan koleksi *stopword* yang sudah ada dalam basis data. Hanya kata-kata yang terdeteksi sama dengan *stopword* yang akan dihilangkan, dan hasilnya dapat dilihat dalam Tabel 3.

Tabel 3 Stopword removal

Ulasan	Stopword
barber yang profesional dan hasilnya gacor k doang ga rugi sih malah untung banyak top tempatnya nyaman baguus dan sang paling aku suka adalah bisa konsultasi masalah rambut	barber profesional hasilnya gacor k doang ga rugi untung top tempatnya nyaman baguus sang suka konsultasi rambut

d) Tokenizing

Tokenizing adalah langkah awal paling mendasar dalam memproses teks. Proses ini pada dasarnya memecah-mecah kalimat atau seluruh dokumen teks menjadi unit-unit yang lebih kecil dan diskrit, yang disebut "token". Token yang digunakan adalah kata. Fungsinya sangat penting karena komputer tidak dapat memproses kalimat secara utuh dengan tokenisasi, setiap kata atau unit dapat diperlakukan sebagai elemen data individual. Hal ini dapat dilihat pada tabel 4.

Tabel 4 Tokenize

Ulasan	Tokenize
harga ribu worth it si barber dengann harga jangkau kualitas nggak asal ku coba langgan si	harga, ribu, worth, it, si, barber, dengann, harga, jangkau, kualitas, nggak, asal, ku, coba, langgan, si
nunggu jam sampe jam set ga kelar cukur nya kali kalo barber nya ga pingin cukur ngomong aja ga lama in cukur nya	nunggu, jam, sampe, jam, set, ga, kelar, cukur, nya, kali, kalo, barber, nya, ga, pingin, cukur, ngomong, aja, ga, lama, in, cukur, nya

e) Stemming

Stemming adalah proses untuk menyederhanakan kata-kata dalam teks. Fungsinya adalah memotong semua imbuhan (seperti awalan atau akhiran) pada kata agar kata tersebut kembali ke bentuk dasarnya. Hasilnya dapat dilihat pada tabel 5.

Tabel 5 Stemming

Ulasan	Stemming
harga ribu worth it si barber dengann harga terjangkau kualitas nggak asalan ku coba langganan si	harga ribu worth it si barber dengann harga jangkau kualitas nggak asal ku coba langgan si
tempatnya oke konsultasi bab rambut langsung akar beda dg penjelasan yg	tempat oke konsultasi bab rambut langsung akar beda dg jelas yg

3.2 Pembobotan TF-IDF

Pemberian nilai atau bobot kepada setiap kata (*term*) dilakukan setelah tahap *preprocessing* selesai. Untuk melakukan pembobotan ini, digunakanlah metode TF-IDF. Inti dari pembobotan ini adalah menetapkan nilai relevansi bagi setiap *term*, di mana nilai bobot tersebut nantinya akan digunakan sebagai input penting dalam menjalankan proses klasifikasi.

3.3 Klasifikasi dan pengujian

Proses klasifikasi menggunakan pendekatan *lexicon-based* melibatkan skoring sentimen pada teks dengan mencocokkan setiap kata dengan daftar kata berlabel positif, netral, dan negatif. Sentimen teks ditentukan oleh dominasi salah satu kategori: teks akan diberi label positif jika memiliki skor positif yang lebih tinggi, dan sebaliknya. Dengan kata lain, pengkategorian sentimen dicapai hanya melalui pemeriksaan langsung terhadap frekuensi dan makna kata-kata yang terdaftar dalam kamus sentimen, jika nilai *sentiment_score* sama dengan 0 (netral). Jika nilai *sentiment_score* lebih dari 0, maka sentimen ditetapkan "Positif". Sebaliknya, jika nilai *sentiment_score* kurang dari 0, maka sentimen ditetapkan sebagai "Negatif".

Tabel 6 Hasil Klasifikasi

Index	Preprocessing	Score	Sentiment
0	tempat nyaman bagus sang suka konsultasi rambut	1	Positif
1	nyaman potong oke karyawan ramah	1	Positif
2	layan buruk antri jam ga tangan tinggal main hp barber	-1	Negatif
3	tolong tingkat lage kerja kasir nya masak kerja nya main hp aja lantai kotor sapu emang gitu kerja an nya kasir kalo ketemu temen nya kerja bongkar aib orang custamer kasir nya	-1	Negatif
4	kesini kesini capster tinggal antre kapok	0	Netral

3.4 Evaluasi

Pengujian kinerja dari Naïve Bayes dilakukan dengan metode *confusion matrix* berdasarkan data *testing* yang akan di proses oleh model Klasifikasi Naïve Bayes yang telah dibuat sebelumnya. Metode *confusion matrix* menghasilkan empat nilai yaitu *True Positive*, *True Negative*, *False Positive* dan *False Negative*. Selanjutnya dilakukan perhitungan mulai dari *accuracy*, *precision*, *recall*, dan *F1-score*. Berikut adalah hasil dari pengujian dengan data *training* 80% dan data *testing* 20%, Hasil dari evaluasi algoritma Naïve bayes disajikan pada tabel berikut.

Table 7 Hasil Evaluasi

Kelas	Akurasi	Presisi	Recall	F1-Score	Support	TP	TN	FP	FN
Positif	85.95%	96.81%	86.67%	91.46%	105	91	13	3	14
Netral	89.26%	00.00%	00.00%	00.00%	4	0	108	9	4
Negatif	93.39%	61.11%	91.67%	73.33%	12	11	102	7	1

4. SIMPULAN

Berdasarkan hasil dan pembahasan dapat ditarik kesimpulan mengenai analisis sentiment ulasan *barbershop* di google maps menggunakan *scraping* dengan *selenium* menggunakan metode Naïve Bayes sebagai berikut:

- Penerapan metode Naïve Bayes menghasilkan performa klasifikasi yang bervariasi antar kelas sentimen. Model menunjukkan akurasi tertinggi pada kelas Negatif (93,39%), diikuti oleh Netral (89,26%) dan Positif (85,95%). Nilai presisi mencapai 96,81% untuk kelas Positif, mengindikasikan ketepatan prediksi yang sangat tinggi untuk ulasan positif. Namun, model mengalami kesulitan dalam mengenali kelas Netral dengan presisi dan recall 0%, yang menunjukkan perlunya peningkatan pada data pelatihan untuk kelas tersebut. Nilai F1-score terbaik dicapai pada kelas Positif (91,46%), sementara kelas Negatif memperoleh 73,33%
- Berdasarkan hasil evaluasi membuktikan bahwa metode Naïve Bayes efektif diterapkan untuk analisis sentimen ulasan barbershop, khususnya dalam mengidentifikasi ulasan Positif dan Negatif. Meskipun terdapat tantangan dalam klasifikasi Netral, performa keseluruhan model menunjukkan kemampuan yang memadai untuk memberikan wawasan sentimen dari data ulasan digital.

5. SARAN

Penelitian selanjutnya dapat berfokus pada peningkatan akurasi model dengan menerapkan algoritma *deep learning* yang lebih kompleks seperti LSTM atau BERT untuk menangani variasi bahasa ulasan secara lebih efektif, serta melakukan analisis sentimen berbasis aspek ABSA untuk mengidentifikasi sentimen spesifik terhadap berbagai komponen layanan *barbershop* untuk memberikan wawasan yang lebih detail dan implementatif bagi pemilik usaha.

DAFTAR PUSTAKA

- [1] F. A. Yeskia, F. Usnen, M. Hafidz, R. Rinaldi, R. Efendi, and R. B. Putra, “Analisis Kepuasan Konsumen terhadap Pelayanan Barbershop dengan Metode Service Quality (Serverqual) dan Impotence Performance Analysi (IPA) (Studi Kasus Pada Barbershop The Gold 74 Andaleh),” *Innov. J. Soc. Sci. Res.*, vol. 2, no. 1, pp. 394–399, 2022, doi: 10.31004/innovative.v2i1.3596.
- [2] K. et al 2023, “PENGARUH KUALITAS PELAYANAN DAN HARGA TERHADAP LOYALITAS KONSUMEN PADA BARBERSHOP KANG CUKUR DI PAGEDANGAN KABUPATEN TANGERANG,” vol. 32, no. 3, pp. 167–186, 2021.
- [3] D. N. Fauzy, K. Muhammad, and A. Sofiana, “USULAN PENINGKATAN KUALITAS PELAYANAN MENGGUNAKAN QUALITY FUNCTION DEPLOYMENT BERBASIS SERVICE QUALITY DAN KANO MODEL PADA BARBERSHOP HAX,” *JUSTI (Jurnal Sist. Dan Tek. Ind.)*, vol. 4, no. 4, pp. 425–432, 2023.
- [4] R. Herdianto, L. Kurniawati, U. Mercu, B. Yogyakarta, and C. A. Pemasaran, “Usaha Pengembangan The Cut Room Barbershop Melalui Penilaian Konsumen,” vol. 3, 2023, [Online]. Available: <http://journal.undiknas.ac.id/index.php/parta>.
- [5] Latifah and Rony Setiawan, “Sistem Informasi Perpustakaan Berbasis Web Dengan Menggunakan Metode Waterfall,” *Informatech J. Ilm. Inform. dan Komput.*, vol. 1, no. 2, pp. 134–141, 2024, doi: 10.69533/52r6tc18.
- [6] A. Safira, F. N. Hasan, K. J. Timur, and N. B. Classifier, “ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER,” *Zo. J. Sist. Inf.*, vol. 5, no. 1, pp. 59–70, 2023.
- [7] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, “Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, p. 180, 2017, doi: 10.25126/jtiik.201743251.
- [8] A. F. RIZKI, W. Prihartono, and F. Rohman, “Analisis Sentimen Ulasan Google Maps Rumah Sakit Khalishah Di Cirebon Dengan Algoritma Naive Bayes,” *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 2, pp. 728–738, 2025, doi: 10.23960/jitet.v13i2.6309.
- [9] C. P. Helena Panjaitan and C. Supriadi, “Analisis Sentimen Pengunjung Wisata Heritage Kota Semarang Menggunakan Naive Bayes Pada Ulasan Google Maps,” *J. Elektron. dan Teknol. Inf.*, vol. 4, no. 2, pp. 2721–9380, 2023.
- [10] V. B. Lestari and C. A. Hutagalung, “Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naive Bayes,” *J-KOMA J. Comput. Sci. Appl.*, no. 021, pp. 22–2025, 2025, [Online]. Available: <https://doi.org/10.21009/j-koma.v02i01.021>
- [11] A. Ardi and Kurniawan, “Optimasi Metode Naïve Bayes Classifier Menggunakan Pendekatan Term Frequency-Inverse Document Frequency (TF-IDF) Pada Analisis Sentimen,” *JSAI (Journal Sci. Appl. Informatics)*, vol. 7, no. 3, pp. 458–463, 2024, doi: 10.36085/jsai.v7i3.7153.
- [12] L. Hartimar, Y. Manza, and K. P. Siregar, “Text Classification Using TF-IDF and Naïve Bayes : Case Study of MyXL App User Review Data,” vol. 2, no. 2, pp. 100–108, 2025.
- [13] M. F. Madjid, D. E. Ratnawati, and B. Rahayudi, “Sentiment Analysis on App Reviews Using Support Vector Machine and Naïve Bayes Classification,” *Sinkron*, vol. 8, no. 1, pp. 556–562, 2023, doi: 10.33395/sinkron.v8i1.12161.
- [14] H. Ma’rifah, A. Prasetya Wibawa, and M. I. Akbar, “Sains, Aplikasi, Komputasi dan Teknologi Informasi Klasifikasi artikel ilmiah dengan berbagai skenario preprocessing,” *Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 70, 2020.
- [15] A. L. Maheswari and E. Krisnanik, “Jurnal Sistem Informasi dan Aplikasi,” *J. Sist. Inf. dan Apl.*, vol. 1, no. 1, pp. 50–64, 2024, [Online]. Available: <https://ejournal.upnvj.ac.id/jgia/article/view/5907>
- [16] all Desi Musfiroh etc, “Sentiment Analysis of Online Lectures in Indonesia from TwitterDataset Using InSet Lexicon,” *J. Homepage*, vol. 1, no. April, pp. 24–33, 2021, [Online]. Available: <https://journal.irpi.or.id/index.php/malcom/article/view/20>

- [17] D. Septiani and I. Isabela, “Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks,” *Sintesia*, vol. 1, pp. 81–88, 2022.
- [18] M. D. Aulia and Y. Akbar, “Penerapan Algoritma Naive Bayes dalam Sistem Analisis Sentimen Media Sosial X terhadap Film Agak Laen,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 6, no. 3, pp. 1958–1967, 2025, doi: 10.63447/jimik.v6i3.1608.
- [19] F. Yudistira and A. R. Isnain, “Analisis Sentimen Terhadap Seleksi CPNS Tahun 2024 Berbasis Media Sosial X Menggunakan Algoritma Naïve Bayes Program Studi Informatika , Fakultas Teknik dan Ilmu Komputer , Universitas Teknokrat Indonesia , SENTIMENT ANALYSIS OF THE 2024 CPNS SELECTION BASED ON SOCIAL MEDIA X USING THE NAÏVE BAYES ALGORITHM,” vol. 5, no. 3, pp. 887–897, 2025.