

Analisis Sentimen Siswa Terhadap Fasilitas Pembelajaran Jurusan TKJ Menggunakan Algoritma *Naïve Bayes*

Nabila Melani Putri¹, Khairunnas², Irma Eryanti Putri³
Fakultas Teknik dan Ilmu Komputer, Program Studi Ilmu Komputer
Universitas Muhammadiyah Bima
Email : nabilamelaniputri87@gmail.com

Abstrak

Pendidikan kejuruan merupakan sistem pendidikan yang berorientasi pada kesiapan kerja, salah satunya melalui Sekolah Menengah Kejuruan (SMK). Jurusan Teknik Komputer dan Jaringan (TKJ) di SMKN 1 Kota Bima menuntut kualitas pembelajaran yang didukung fasilitas dan strategi pembelajaran yang efektif. Penelitian ini bertujuan menganalisis sentimen siswa terhadap pembelajaran TKJ menggunakan algoritma *Naïve Bayes*. Kebaruan penelitian terletak pada penerapan analisis sentimen berbasis *Naïve Bayes* pada data opini siswa TKJ di SMKN 1 Kota Bima yang sebelumnya belum pernah dilakukan, serta penggunaan data primer langsung dari siswa untuk mengevaluasi kualitas pembelajaran dan fasilitas sekolah. Dibandingkan metode lain seperti *Support Vector Machine* (SVM) dan *Deep Learning*, *Naïve Bayes* dipilih karena lebih sederhana, efisien, dan mampu bekerja baik pada jumlah data terbatas. Data penelitian diperoleh melalui penyebaran kuesioner kepada siswa kelas X–XII dengan total 106 data. Tahapan penelitian meliputi preprocessing, pembobotan TF-IDF, dan pengujian menggunakan *confusion matrix*. Hasil penelitian menunjukkan akurasi sebesar 77,27%, dengan performa terbaik pada kelas positif, sedangkan kelas netral memiliki performa terendah akibat jumlah data yang sedikit.

Kata kunci: *Naïve Bayes*, Analisis Sentimen, Data Mining

Abstract

Vocational education is an education system focused on work readiness, one of which is through Vocational High Schools (SMK). The Computer and Network Engineering (TKJ) program at SMKN 1 Kota Bima requires high-quality instruction supported by effective facilities and teaching strategies. This study aims to analyze student sentiment toward TKJ learning using the *Naïve Bayes* algorithm. The novelty of this research lies in the application of *Naïve Bayes*-based sentiment analysis to data on the opinions of TKJ students at SMKN 1 Kota Bima, which has not been done before, as well as the use of primary data directly from students to evaluate the quality of learning and school facilities. Compared to other methods such as *Support Vector Machine* (SVM) and *Deep Learning*, *Naïve Bayes* was chosen because it is simpler, more efficient, and capable of performing well with limited data. Research data was obtained through the distribution of questionnaires to students in grades X–XII, yielding a total of 106 data points. The research stages included preprocessing, TF-IDF weighting, and testing using a *confusion matrix*. The results showed an accuracy of 77.27%, with the best performance in the positive class, while the neutral class had the lowest performance due to the small amount of data.

Keywords: *Naïve Bayes*, Sentimen Analysis, Text Mining

PENDAHULUAN

Pendidikan Kejuruan merupakan suatu bentuk sistem pendidikan Indonesia yang berorientasi pada pekerjaan. Bentuk pendidikan kejuruan yang berada di Indonesia salah satunya yaitu Sekolah Menengah Kejuruan (Umarani et al., 2021). Salah satu indikator utama mutu pendidikan terletak pada kualitas pembelajaran yang berlangsung di sekolah (Sita et al., 2024). Fasilitas sangat penting sebagai penunjang terlaksananya proses pembelajaran, mempermudah guru dan tenaga pendidik dalam menyampaikan materi kepada siswa, dan dapat menambah minat dan perhatian siswa, sehingga saat proses pembelajaran siswa mampu aktif (Pratama et al., 2022).

Di era digital ini, pembelajaran di jurusan TKJ (Teknik Komputer dan Jaringan) telah menjadi komponen penting dalam kurikulum sekolah menengah kejuruan (SMK), termasuk di SMKN 1 Kota Bima, yang fokus pada pengembangan keterampilan praktis siswa dibidang teknologi dan inovasi (Wiratama Putra et al., 2022). Fasilitas pembelajar dalam jurusan TKJ (Teknik Komputer dan Jaringan) di sekolah menengah kejuruan (SMK) merupakan langkah penting dalam menentukan masa depan siswa karna sumber daya manusia dapat mempengaruhi kualitas.

Masalah utama yang mendasari penelitian ini adalah kurangnya evaluasi sistematis terhadap persepsi dan sentimen siswa terhadap fasilitas pembelajaran pada jurusan TKJ (Teknik Komputer dan Jaringan) di SMKN 1 Kota Bima. Research gap menunjukkan bahwa dinamika sentimen dalam konteks pembelajan TKJ sering kali belum dieksplorasi secara mendalam di

sekolah – sekolah SMK di Indonesia (Nugraha et al., 2022). Laporan dari Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) pada tahun 2023 menunjukkan bahwa hanya 40% sekolah SMK di Indonesia yang melakukan survei kepuasan siswa secara rutin, dengan fokus utama pada aspek akademik umum, bukan spesifik pada respon terhadap multimedia sebagai fasilitas dalam proses pembelajaran, data ini menegaskan bahwa tanpa analisis sentimen,sekolah seperti SMKN 1 Kota Bima berisiko kehilangan wawasan kritis tentang mutu fasilitas pembelajaran, yang dapat mengakibatkan penurunan motivasi siswa dan kegagalan dalam pencapaian kompetensi vokasi (Sariyusda, 2024).

Analisis sentimen, sebagai cabang dari pemrosesan bahasa alami (Natural language Processing/NLP), memungkinkan pengukuran opini dan emosi siswa melalui teks, seperti ulasan, komentar atau survei. Dengan menggunakan algoritma klasifikasi Naïve Bayes, yang dikenal karena kemudahan dan efisiensinya dalam menangani data berdasarkan analisis probabilitas, sentimen dapat (Varma et al., 2023) mengidentifikasi apakah sentimen siswa terhadap fasilitas pembelajaran pada jurusan teknik komputer dan jaringan bersifat positif, negatif, atau netral (Setups et al., 2021). Urgensi analisis sentimen terletak pada kemampuan untuk mengungkap pola emosi siswa, seperti motivasi belajar atau frustrasi teknis, yang dapat mencegah penurunan kualitas pembelajaran dan mendukung pengembangan kurikulum yang adaptif.

Penelitian ini didorong oleh kebutuhan untuk menjawab kesenjangan penelitian tersebut melalui penerapan

algoritma *Naïve Bayes*, yang bertujuan untuk menganalisis sentimen data dari siswa jurusan TKJ di SMKN 1 Kota Bima, sehingga dapat memberikan wawasan bagi pengambil kebijakan pendidikan untuk meningkatkan strategis fasilitas pembelajaran. Pendekatan ini tidak hanya bersifat teoritis, tetapi juga praktis, karena *Naïve Bayes* telah terbukti efektif dalam analisis sentimen pada skala besar dengan tingkat akurasi yang memadai (Kaka et al., 2023). Dibandingkan dengan SVM yang memerlukan tuning parameter kompleks, *Naïve Bayes* lebih efisien dalam sumber daya komputasi, sementara *Deep Learning* seperti *Reccurent Neural Networks* (RNN) membutuhkan data besar dan waktu pelatihan lebih lama (Rahayu et al., 2022).

Penelitian sebelumnya mengenai analisis sentimen di bidang pendidikan umumnya berfokus pada pembelajaran daring, media sosial, dan evaluasi kurikulum, sedangkan penelitian terkait sentimen siswa terhadap fasilitas pembelajaran pada jurusan Teknik Komputer dan Jaringan (TKJ) di lingkungan SMK masih terbatas. Selain itu, belum banyak penelitian yang memanfaatkan data opini siswa secara langsung sebagai dasar evaluasi fasilitas pembelajaran di sekolah vokasi, khususnya di SMKN 1 Kota Bima.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan analisis sentimen siswa terhadap fasilitas pembelajaran jurusan TKJ menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan data komentar siswa. Kontribusi penelitian ini terletak pada pemanfaatan analisis sentimen berbasis data mining sebagai media evaluasi fasilitas pembelajaran di sekolah vokasi dengan jumlah dataset terbatas.

Penelitian ini diharapkan dapat memberikan kontribusi dalam evaluasi fasilitas pembelajaran berbasis data mining untuk mendukung pengambilan keputusan pada institusi pendidikan vokasi.

Rumusan Masalah

1. Bagaimana mengidentifikasi dan mengklasifikasi sentimen siswa jurusan TKJ di SMKN 1 Kota Bima terhadap fasilitas pembelajaran menggunakan algoritma *Naïve Bayes*?
2. Bagaimana distribusi sentimen positif, negatif, dan netral siswa terhadap fasilitas pembelajaran TKJ di SMKN 1 Kota Bima berdasarkan hasil analisis sentimen?
3. Bagaimana performa algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen siswa terhadap fasilitas pembelajaran berdasarkan hasil akurasi dan *confusion matrix*?

Tujuan Penelitian

1. Mengukur performa algoritma *Naïve Bayes* untuk mengklasifikasikan sentimen siswa jurusan TKJ di SMKN 1 Kota Bima berdasarkan evaluasi *confusion matrix*.
2. Menganalisis persepsi siswa terhadap fasilitas pembelajaran jurusan TKJ melalui klasifikasi sentimen positif, negatif, dan netral menggunakan metode TF-IDF dan *Naïve Bayes*.
3. Memberikan evaluasi terhadap fasilitas pembelajaran berbasis hasil analisis sentimen siswa guna mendukung peningkatan kualitas pembelajaran jurusan TKJ di SMKN 1 Kota Bima.

Manfaat Penelitian

1. Penelitian ini dapat diharapkan memberikan kontribusi akademik untuk pengembangan kajian analisis sentimen dan data mining pada bidang pendidikan, khususnya dalam mengklasifikasikan opini siswa terhadap fasilitas pembelajaran jurusan TKJ menggunakan algoritma *Naïve Bayes*.
2. Penelitian ini diharapkan dapat membantu pihak sekolah dan tenaga pendidik untuk memahami persepsi siswa terhadap fasilitas pembelajaran jurusan TKJ di SMKN 1 Kota Bima guna sebagai bahan evaluasi untuk.

Tinjauan Pustaka

Penelitian yang dilakukan (Pratama et al., 2022) bentuk pendidikan kejuruan yang berada di Indonesia salah satunya yaitu, Sekolah Menengah kejuruan (SMK) dalam konteks pendidikan vokasi, fasilitas pembelajaran teknik komputer dan jaringan memainkan peran krusial dalam mengembangkan keterampilan praktis siswa melalui pembelajaran teori dan praktik.

penelitian yang dilakukan (Alvionita et al., 2025) mengatakan Analisis sentimen merupakan cabang dari Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi, mengekstrak, dan mengklasifikasikan opini, emosi, atau sikap yang terkandung dalam teks. Dan (Harahap et al., 2024) Analisis sentimen telah terbukti efektif dalam mengukur kepuasan pengguna pada berbagai domain, termasuk pendidikan, dan memungkinkan pengambilan kebijakan untuk memahami pola emosi dan sikap siswa.

Selanjutnya penelitian yang dilakukan (Syahar et al., 2024)

menunjukkan bahwa *Naïve Bayes* mampu mengklasifikasikan sentimen fasilitas laboratorium dengan akurasi yang baik, namun penelitian tersebut hanya berfokus pada fasilitas laboratorium tanpa mengevaluasi fasilitas pembelajaran vokasi secara menyeluruh.

Penelitian lain oleh (Alzaid & Fkih, 2023) menunjukkan bahwa analisis sentimen terhadap umpan balik siswa dapat membantu institusi pendidikan memahami tingkat kepuasan siswa pada proses pembelajaran dan fasilitas yang digunakan. Selain itu, (Kastrati et al., 2021) menyatakan metode *Deep Learning* mampu menghasilkan akurasi yang tinggi dalam analisis sentimen pendidikan, namun membutuhkan jumlah data yang besar serta waktu pelatihan yang lebih lama. Dan (Umarani et al., 2021) menegaskan bahwa pemilihan algoritma harus disesuaikan dengan karakteristik data, terutama pada dataset kecil yang lebih cocok menggunakan metode klasifikasi sederhana seperti *Naïve Bayes*.

Berdasarkan hasil penelitian terdahulu tersebut, masih terdapat keterbatasan penelitian yang secara khusus menganalisis sentimen siswa terhadap fasilitas pembelajaran pada jurusan TKJ di lingkungan SMK menggunakan data primer siswa. Oleh karena itu, penelitian ini menerapkan algoritma *Naïve Bayes* untuk menganalisis sentimen siswa TKJ SMKN 1 Kota Bima terhadap fasilitas pembelajaran dengan dataset berjumlah 106 data opini siswa.

LANDASAN TEORI

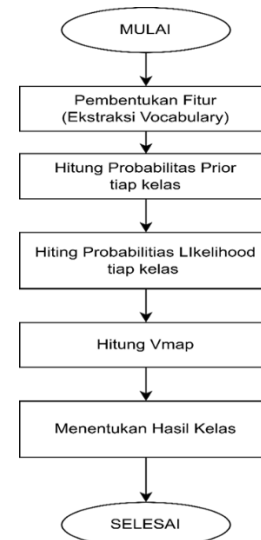
Menurut (Lende et al., 2023) Evaluasi kepuasan siswa SMK di Sumatra menggunakan text mining dan clustering, berhasil mengidentifikasi 5 faktor

kepuasaan utama, namun tidak menggunakan klasifikasi sentimen dan lebih fokus pada clustering.

Menurut (Pratama et al., 2022) Teori Teknik Komputer dan Jaringan (TKJ) merupakan program keahlian pada pendidikan vokasi yang menekankan penguasaan keterampilan teknis dibidang komputer dan jaringan melalui perpaduan antara pembelajaran teori dan praktik. Pembelajaran TKJ dirancang untuk membekali siswa dengan kompetensi kerja yang relevan dengan kebutuhan dunia industri khususnya yang berkaitan dibidang perakitan komputer, instalasi perangkat lunak, serta pengelolaan jaringan komputer.

Menurut (Pebdika et al., 2023) Analisis sentimen mengidentifikasi opini positif, negatif, atau netral dari teks sebagai cabang pemrosesan bahasa alami (NLP), Teori ini mendukung aktivitas preprocessing, seperti tokenisasi dan penghapusan stop words, untuk klasifikasi sentimen yang akurat.

Menurut (Febriyani & Februariyanti, 2022) Naive Bayes digunakan untuk mengambil keputusan dengan melakukan prediksi suatu kasus berdasarkan hasil dari klasifikasi yang telah diperoleh, Algoritma ini menggunakan perhitungan probabilitas dan statistik dalam pengoperasiannya. Alur algoritma *Naive Bayes* sebagai berikut:



Gambar 1. Alur Algoritma Naive Bayes

Bedasarkan Gambar 1. pada tahapan algoritma *Naive Bayes* dapat dijelaskan sebagai berikut:

- Pembentukan Fitur (*Ekstraksi Vocabulary*) merupakan proses ekstraksi dan identitas fitur (kata-kata unik) dari data yang telah melalui preprocessing.
- Hitung Probabilitas Prior Tiap Kelas $P(c)$ menghitung probabilitas prior untuk setiap kelas sentimen berdasarkan distribusi data latih.

Rumus yang digunakan :

$$P(C) = \frac{N_c}{N} \quad (1)$$

Penjelasan :

N_c = jumlah dokumen pada kelas c .

N = jumlah seluruh dokumen.

- Hitung Probabilitas Likelihood Tiap Kelas $P(X_i|C)$ dengan menghitung probabilitas likelihood, dan kemunculan probabilitas setiap kata $\{(X)_i\}$ disetiap kelas sentimen (C), menggunakan rumus Laplace Smoothing:

$$P(w | C) = \frac{n_{w,C} + 1}{n_C + |V|} \quad (2)$$

Penjelasan :

$n_{w,C}$ = frekuensi kata w pada kelas C .

n_C = total keseluruhan kata pada kelas C .

$|V|$ = jumlah vocabulary.

- d. Hitung V_{map} dengan menghitung Maximum A Posterior (MAP) untuk menentukan kelas dengan probabilitas tertinggi bagi setiap dokumen uji.

Rumus yang digunakan:

$$V_{map} = P(C) \times \prod_{i=1}^n P(w_i | C) \quad (3)$$

Penjelasan :

w_i = kata ke- i dalam dokumen.

n = jumlah kata dalam dokumen.

- e) Menentukan hasil akhir Kelas dengan nilai V_{map} terbesar dipilih sebagai hasil klasifikasi:

$$mC^* = \arg \max_C \left(P(C) \times \prod_{i=1}^n P(w_i | C) \right)$$

Penjelasan :

C = himpunan kelas yang tersedia.

$P(c)$ = probabilitas prior dari kelas C .

w_i = kata ke- i dalam dokumen yang diklasifikasikan.

n = jumlah kata dalam dokumen.

$P(w_i | C)$ = Probabilitas likelihood .

$Argmax$ = fungsi yang memiliki kelas dengan nilai probabilitas yang besar.

METODE PENELITIAN

Pengumpulan data dilakukan menggunakan data primer melalui penyebaran kuisioner pada siswa jurusan TKJ kelas X, XI, dan XII di SMKN 1 Kota Bima, total data yang berhasil dikumpulkan dari keseluruhan kelas sebanyak 106 data. Setiap komentar kemudian dilakukan proses labeling sentimen secara manual kedalam tiga kategori yaitu positif, negatif dan netral berdasarkan opini siswa terhadap fasilitas pembelajaran, tahapan labeling dilakukan untuk memastikan data kelas yang jelas sebelum memasuki proses klasifikasi dengan algoritma *Naïve Bayes*.

Algoritma *Naïve Bayes* dipilih karena memiliki keunggulan dalam menangani data teks dengan jumlah dataset relatif kecil serta proses komputasi yang lebih cepat dan

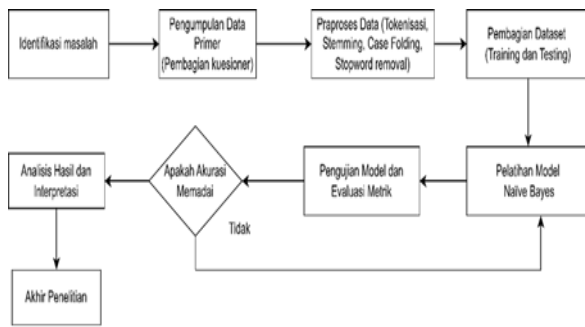
sederhana. Dibandingkan Support Vector Machine (SVM), *Naïve Bayes* tidak memerlukan proses tuning parameter yang kompleks. Sementara Random Forest dan XGBoost membutuhkan sumber daya komputasi lebih besar serta proses pelatihan yang lebih lama. Oleh karena itu, *Naïve Bayes* dinilai lebih sesuai untuk penelitian ini yang menggunakan dataset terbatas sebanyak 106 data komentar siswa.

Pengujian model dilakukan menggunakan *Confusion Matrix* untuk mengukur nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Selain itu, penelitian ini juga menerapkan metode *K-Fold Cross Validation* sebanyak 5-fold untuk meningkatkan validitas evaluasi model.

Metode *K-Fold Cross Validation* dilakukan dengan membagi dataset menjadi lima bagian, dimana setiap bagian digunakan secara bergantian sebagai data testing dan sisanya sebagai data training. Penggunaan metode ini bertujuan untuk mengurangi bias pembagian data dan menghasilkan evaluasi model yang lebih stabil dibandingkan hanya menggunakan satu kali pembagian data training dan testing.

Alur Penelitian

Penelitian ini merangkai beberapa tahapan penelitian yang disusun secara sistematis agar tujuan penelitian tercapai, Bagan alir penelitian ini dirancang untuk menggambarkan urutan metodologis yang sistematis dalam penerapan algoritma *Naïve Bayes* untuk sentimen siswa terhadap fasilitas pembelajaran TKJ di SMKN 1 Kota Bima. Gambar. 2 merangkum tahapan alur penelitian sebagai berikut:



Gambar 2. Diagram alur penelitian

Identifikasi Masalah

Dimulai dengan identifikasi masalah penelitian, yaitu analisis sentimen siswa terhadap fasilitas pembelajaran TKJ di SMKN 1 Kota Bima. Fasilitas pembelajaran TKJ efektif untuk meningkatkan pemahaman, namun persepsi siswa perlu di analisis menggunakan metode kuantitatif seperti analisis sentimen.

Pengumpulan Data

Pengumpulan data ini diperoleh dari data primer ini dikumpulkan secara langsung dari responden siswa melalui penyebaran kuensioner pada siswa X, XI, dan XII. Penyebaran koensioner pada setiap kelas jurusan TKJ. Dari keseluruhan data yang didapat sebanyak 106 data komentar siswa, contoh data komentar siswa dapat dilihat pada tabel 1 berikut.

Tabel 1. Komentar mentah siswa

Kelas	Komentar	Label Sentiemn
X TKJ	SANGAT BURUK,Kurang memadai, WIFI NGELEK,BERSIH TAPI TIDAK NYAMAN KARENA PANAS,BURUK SEMUA,tingkatkan pembelajarannya karena sering guru tidak masuk ke kelas pembelajaran berantakan.	NEGATIF

XI TKJ beberapa mapel POSITIF jurusan sebaiknya materi langsung selanjutnya praktek, bukan materi dan materi setelah materinya di tumpuk baru praktek satu persatu, karna kalau hanya media catat saja mungkin akan lebih sulit untuk mengingat nya kembali, beda dengan baru saja di jelaskan materinya dan di ikuti oleh praktek nya

XII TKJ menurut saya, NEGATIF fasilitas pembelajaran di tkj sudah baik dan mendukung proses belajar mengajar.saya berharap fasilitas ini dapat terus di pertahankan dan di tingkatkan agar pembelajaran semakin optimal

Pre-processing

Untuk tahapan preprocessing data mentah hasil komentar siswa akan melewati beberapa tahapan yaitu dengan data cleaning, case folding, tokenizing, normalisasi, stopword removal, stemming, dan TF-IDF untuk mencapai hasil data final text (Nugraha et al., 2022).

1. Data Cleaning

Tabel 2. Proses data *cleaning*

Komentar	Cleaning
saya harap fasilitas nya lebih di perhatikan lagi,tidak cukup baik kurang, BERSIH TAPI NGGAK NYAMAN KARENA PANAS,kurang baik dimohon kepada pihak sekolah agar	saya harap fasilitas nya lebih di perhatikan lagitidak cukup baik kurang BERSIH TAPI NGGAK NYAMAN KARENA PANAS kurang baik dimohon kepada pihak sekolah

memperbaiki fasilitas untuk menunjang pembelajaran.	agar memperbaiki fasilitas untuk menunjang pembelajaran
SANGAT BURUK, Kurang memadai, WIFI NGELEK, BERSIH TAPI TIDAK NYAMAN KARENA PANAS, BURUK SEMUA, tingkatkan pembelajarannya karena sering guru tidak masuk ke kelas pembelajaran berantakan.	SANGAT BURUK Kurang memadai WIFI NGELEK BERSIH TAPI TIDAK NYAMAN KARENA PANAS, BURUK SEMUA tingkatkan pembelajarannya karena sering guru tidak masuk ke kelas pembelajaran berantakan

Cleaning merupakan tahapan pembersihan data dalam preprocessing data yang berfungsi untuk menghapus karakter tanda baca seperti tanda koma ataupun tanda titik.

2. Case Folding

Tabel 3. Proses case folding

Cleaning	Case Folding
saya harap fasilitas nya lebih di perhatikan lagi tidak cukup baik kurang BERSIH TAPI NGGAK NYAMAN KARENA PANAS kurang baik dimohon kepada pihak sekolah agar memperbaiki fasilitas untuk menunjang pembelajaran	saya harap fasilitas nya lebih di perhatikan lagi tidak cukup baik kurang bersih tapi nggak nyaman karena panas kurang baik dimohon kepada pihak sekolah agar memperbaiki fasilitas untuk menunjang pembelajaran
SANGAT BURUK Kurang memadai WIFI NGELEK BERSIH TAPI TIDAK NYAMAN KARENA PANAS BURUK SEMUA tingkatkan pembelajarannya karena sering guru tidak masuk ke kelas pembelajaran berantakan	sangat buruk kurang memadai wifi ngelek bersih tapi tidak nyaman karena panas buruk semua tingkatkan pembelajarannya karena sering guru tidak masuk ke kelas pembelajaran berantakan

case folding tahapan untuk melakukan keseragaman huruf dimana tahapan ini akan mengubah huruf yang awalnya huruf kapital menjadi huruf kecil.

3. Tokenizing

Tabel 4. Proses tokenizing

Case Folding	Tokenizing
--------------	------------

saya harap fasilitas nya lebih di perhatikan lagi tidak cukup baik kurang bersih tapi nggak nyaman karena panas kurang baik dimohon kepada pihak sekolah agar memperbaiki fasilitas untuk menunjang pembelajaran	[saya, harap, fasilitas, nya, lebih, di, perhatikan, lagi, tidak, cukup, baik, kurang, bersih, tapi, nggak, nyaman, karena, panas, kurang, baik, dimohon, kepada, pihak, sekolah, agar, memperbaiki, fasilitas, untuk, menunjang, pembelajaran]
sangat buruk kurang memadai wifi ngelek bersih tapi tidak nyaman karena panas buruk semua tingkatkan pembelajarannya karena sering guru tidak masuk ke kelas pembelajaran berantakan	[sangat, burukkurang, memadai, wifi, ngelek, bersih, tapi, tidak, nyaman, karena, panas, buruk, semua, tingkatkan, pembelajarannya, karena, sering, guru, tidak, masuk, ke, kelas, pembelajaran, berantakan]

Proses tokenizing dilakukan setelah proses case folding, pada tahapan ini teks hasil case folding dipecah menjadi unit-unit kata tunggal yang disebut token.

4. Normalisasi

Tabel 5. Proses normalisasi

Tokenizing	Normalisasi
[saya, harap, fasilitas, nya, lebih, di, perhatikan, lagi, tidak, cukup, baik, kurang, bersih, tapi, nggak, nyaman, karena, panas, kurang, baik, dimohon, kepada, pihak, sekolah, agar, memperbaiki, fasilitas, untuk, menunjang, pembelajaran]	[saya, harap, fasilitas, nya, lebih, di, perhatikan, lagi, tidak, cukup, baik, kurang, bersih, tapi, nggak, nyaman, karena, panas, kurang, baik, dimohon, kepada, pihak, sekolah, agar, memperbaiki, fasilitas, untuk, menunjang, pembelajaran]
[sangat, buruk, kurang, memadai, wifi, ngelek, bersih, tapi, tidak, nyaman, karena, panas, buruk, semua, tingkatkan, pembelajarannya, karena, sering, guru, tidak, masuk, ke, kelas, pembelajaran, berantakan]	[sangat, buruk, kurang, memadai, wifi, ngelek, bersih, tapi, tidak, nyaman, karena, panas, buruk, semua, tingkatkan, pembelajarannya, karena, sering, guru, tidak, masuk, ke, kelas, pembelajaran, berantakan]

Tahapan ini setiap kata yang singkatan disesuaikan dalam kata-kata baku. Normalisasi dilakukan untuk mengurangi variasi kata yang memiliki

makna sama namun ditulis dalam bentuk yang berbeda.

5. Stopword Removal

Tabel 6. Proses *stopword removal*

Normalisasi	Stopword Removal
[saya, harap, fasilitas, nya, lebih, di, perhatikan, lagi, tidak, cukup, baik, kurang, bersih, tapi, nggak, nyaman, karena, panas, kurang, baik, dimohon, kepada, pihak, sekolah, agar, memperbaiki, fasilitas, untuk, menunjang, pembelajaran]	[harap, fasilitas, nya, perhatikan, lagi, tidak, bersih, nggak, nyaman, panas, kurang, dimohon, sekolah, memperbaiki, fasilitas, menunjang, pembelajaran]
[sangat, buruk, kurang, memadai, wifi, ngelek, bersih, tapi, tidak, nyaman, karena, panas, buruk, semua, tingkatkan, pembelajarannya, karena, sering, guru, tidak, masuk, ke, kelas, pembelajaran, berantakan]	[buruk, kurang, memadai, wifi, ngelek, bersih, nyaman, panas, buruk, semua, tingkatkan, pembelajarannya, guru, masuk, kelas, pembelajaran, berantakan]

Proses *stopword removal* merupakan tahapan yang dilakukan setelah melakukan proses normalisasi. Tujuan dari proses *stopword removal* adalah untuk mengurangi jumlah kata yang tidak relevan.

6. Stemming

Tabel 7. Proses *stemming*

Stopword Removal	Stemming
[harap, fasilitas, nya, perhatikan, lagi, tidak, bersih, nggak, nyaman, panas, kurang, dimohon, sekolah, memperbaiki, fasilitas, menunjang, pembelajaran]	[harap, fasilitas, nya, perhati, lagi, tidak, bersih, nggak, nyaman, panas, kurang, mohon, sekolah, baik, fasilitas, tunjang, ajar]
[buruk, kurang, memadai, wifi, ngelek, bersih, nyaman, panas, buruk, semua, tingkatkan, pembelajarannya, guru, masuk, kelas, pembelajaran, berantakan]	[buruk, kurang, pada, wifi, ngelek, bersih, nyaman, panas, buruk, semua, tingkatkan, ajar, guru, masuk, kelas, ajar, beranta]

Proses *stemming* merupakan tahapan yang dilakukan setelah proses *stopword removal*. Pada tahap ini, setiap kata dikembalikan ke bentuk kata dasarnya dengan cara menghilangkan imbuhan

seperti awalan, akhiran, sisipan, maupun gabungan imbuhan. Proses *stemming* bertujuan untuk menyatukan berbagai bentuk kata yang memiliki makna yang sama ke dalam satu representasi kata dasar.

Pembagian Dataset

Membagi dataset menjadi data training 70% dan data testing 30% secara acak untuk menghindari *overfitting*, pembagian dataset ini untuk mengevaluasi apakah pembagian ini seimbang berdasarkan distribusi kelas sentimen Negatif, Positif, Netral (Adelia et al., 2023).

Pelatihan Model *Naïve Bayes*

Pelatihan model *Naïve Bayes* dengan menggunakan data training untuk menghitung probabilitas prior dan memungkinkan untuk setiap kata kunci dalam kelas sentimen (Syahar et al., 2024).

Pengujian Model

Untuk pengujian model menggunakan metode *Confusion Matrix* yang merupakan sebuah metode yang dapat mengevaluasi kinerja dan menentukan seberapa mungkit tingkat akurasi dari metode klasifikasi yang dipakai untuk memprediksi hasil *True Positive*, *False Positive*, *True Negative* atau *False Negative* (Lende et al., 2023).

IMPLEMENTASI DAN PEMBAHASAN

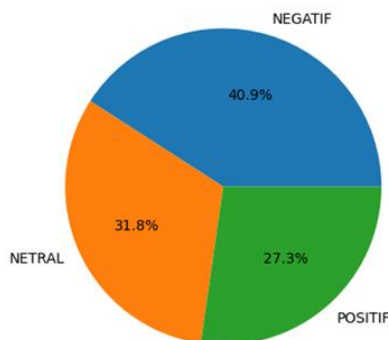
Berdasarkan hasil pengujian model nilai akurasi yang diperoleh sebesar 77,27%, nilai tersebut menunjukkan bahwa model memiliki kemampuan kinerja yang cukup baik dalam mengidentifikasi sentimen siswa terhadap fasilitas pembelajaran TKJ di SMKN 1 Kota Bima.

Pada penelitian ini analisis sentimen dengan data yang relatif kecil dan menghasilkan akurasi diatas 70% sudah

termasuk kategori baik karena model mampu mengenali pola sentimen secara konsisten. Hasil penelitian ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa algoritma *Naïve Bayes* efektif digunakan pada klasifikasi sentimen berbasis teks dengan kompleksitas data yang sederhana dan ukuran data yang terbatas.

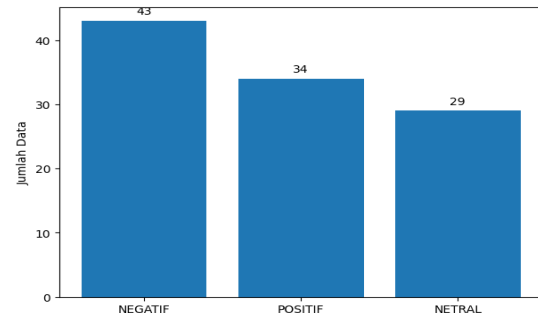
Hasil Visualisasi

Pada penelitian ini visualisasi representasi merupakan hasil dari penelitian pada gambar. 3 dapat dilihat pada warna biru sentimen negatif dengan presentase 40,9%, warna hijau sentimen positif dengan presentase 27,3%, sedangkan warna orange netral dengan presentase 31,8% dengan semua total 106 sentimen.



Gambar 3. Hasil Visualisasi

Hasil dari visualisasi *bar cart* pada Gambar. 4 memperlihatkan distribusi dari jumlah sentimen pada masing-masing kelas secara jelas. Berdasarkan hasil visualisasi sentimen negatif memiliki jumlah paling tinggi dibandingkan sentimen positif dan netral. Dengan ini menunjukkan bahwa sebagian besar siswa masih merasa kurang terhadap fasilitas pembelajaran TKJ yang belum sepenuhnya memenuhi kebutuhan pembelajaran praktik.



Gambar 4. Visualisasi Bar cart

TF-IDF

Hasil *TF-IDF* pada Gambar. 5 ini digunakan untuk mengetahui frekuensi dari istilah tertentu yang relatif terhadap sebuah kata dalam kumpulan dokumen yang melihat seberapa umum atau tidak umum sebuah kata yang ada diantara sekumpulan teks yang terstruktur atau memberikan bobot pada setiap kata yang ada menggunakan *TF-IDF* (Harahap et al., 2024).

praktek	alat	praktik	aman	bagus	baik	baik	fasilitas
0.000	0.000	0.13	0.00	0.000	0.000	0.000	0.00
0.000	0.00	0.00	0.00	0.000	0.193	0.000	0.26
0.000	0.00	0.00	0.00	0.000	0.000	0.000	0.00
0.000	0.00	0.00	0.00	0.000	0.000	0.000	0.00
0.000	0.00	0.00	0.00	0.000	0.000	0.000	0.00
0.000	0.00	0.00	0.42	0.202	0.000	0.000	0.00
0.000	0.00	0.00	0.00	0.000	0.000	0.000	0.00
0.000	0.00	0.00	0.00	0.000	0.000	0.000	0.00
0.394	0.00	0.00	0.00	0.000	0.000	0.000	0.00
0.000	0.00	0.00	0.00	0.000	0.236	0.000	0.00

Gambar 5. Hasil TF-IDF

Pembagian Dataset

Pembagian data opini sebanyak 106 dibagi menjadi 70% data trainig dan 30% data testing Pembagian data training 74 data dan data testing 32 data, dapat dilihat pada Tabel 8 berikut.

<i>Training</i>	74
<i>Testing</i>	32

Pelatihan *Naïve Bayes*

Pada tahapan pemodelan *Naïve Bayes* dengan menggunakan data training untuk menghitung probabilitas prior menghasilkan nilai setiap kelas dapat dilihat pada gambar. 6 berikut.

Processing (NLP) pada data opini siswa SMK yang masih jarang dilakukan, yang fokus pada evaluasi fasilitas pembelajaran dilingkungan pendidikan vokasi.

Hasil dari penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* mampu mengklasifikasikan sentimen positif, negatif, dan netral dengan nilai akurasi cukup baik sebesar 77,27%, dengan dataset yang berukuran kecil sebanyak 106 data. Sentimen negatif menjadi kategori yang dominan, yang mengindikasikan bahwa siswa merasa tidak puas terhadap fasilitas pembelajaran TKJ. Selain itu, hasil evaluasi menunjukkan bahwa kelas netral memiliki performa klasifikasi paling rendah akibat ketidakseimbangan distribusi data dan adanya kata yang tidak dipamami oleh model pada komentar siswa.

Saran

Untuk penelitian selanjutnya, disarankan melakukan pengembangan model menggunakan algoritma lain seperti *Support Vektor Machine* (SVM), *Random Forest*, maupun pendekatan *Deep Learning* berbasis NLP. Selain itu, penggunaan dataset yang lebih besar, tekning *balancing* data, dan optimasi *preprocessing* yang diharapkan mampu meningkatkan performa klasifikasi serta dapat menghasilkan model analisis sentimen yang lebih akurat dan stabil.

DAFTAR PUSTAKA

- Adelia, S., Kurniawan, F., Milanda, E., Santari, J., Kesuma, D. T., & Silvia, E. (2023). Analisis Sentimen Belajar Programming Pada Media Sosial Youtube Menggunakan Algoritma Klasifikasi Naive Bayes. *Journal of Information Technology Ampera*, 4(3), 254–264.
- Alvionita, T. M., Achmad Hindasyah, A. H., & Abu Khalid, A. K. (2025). *Analisis Sentimen Persepsi Pengguna Kurikulum Merdeka Belajar Menggunakan Metode Naive Bayes Classifier (Nbc) Dan Support Vector Machine (Svm) Studi Kasus : Smk Teknikom Cikarang*. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, 9(1), 124–133. <https://doi.org/10.47970/siskom-kb.v9i1.878>
- Alzaid, M., & Fkih, F. (2023). *applied sciences Sentiment Analysis of Students ' Feedback on E-Learning Using a Hybrid Fuzzy Model*.
- Febriyani, E., & Februariyanti, H. (2022). *Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter*. 17(1), 25–38.
- Harahap, Y. N., Nisa, R. K., Mulianingias, R., Tumanggor, S. H., & Syahrani, A. (2024). *Analisis Sentimen Pada Review Buku Ajar Siswa Sma Kelas Xi Pada Materi Vektor Abstract : Learning mathematics is a process that someone goes through through three nodes : mental action , thinking process , and understanding process . Because mathematics i. 10(2021), 582–589*.
- Kaka, O., Ananda Putra Tanggu Mara, A., & Wulla Rato, K. (2023). *Analisis Sentimen Dampak Perkembangan Teknologi Informasi dan Komunikasi Terhadap Kemajuan Belajar Siswa SMK Rada Pamba dengan Metode Naive Baiyes*. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 6(3), 191–199.
- Kastrati, Z., Dalipi, F., Imran, A. S., & Nuci, K. P. (2021). *applied sciences Sentiment Analysis of Students ' Feedback with NLP and Deep Learning : A Systematic Mapping Study*.
- Lende, S., Nahak Tetik, Y., Ayu, M. P., Informatika, T., Stella, S., Sumba, M., Tambolaka, I., & Korespondensi, E. P. (2023). *Analisis Sentimen Siswa Terhadap Pelajaran Informatika di*

- SMPK St. Yohanes Kalembu Lona Dengan Metode Naive Bayes Classifier*. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 6, 217–225.
- Nugraha, S. N., Pebrianto, R., Latif, A., & Firdaus, M. R. (2022). *Analisis Sentimen Twitter Terhadap Menteri Indonesia Dengan Algoritma Support Vector Machine Dan Naive Bayes*. *E-Link: Jurnal Teknik Elektro Dan Informatika*, 17(1), 1. <https://doi.org/10.30587/e-link.v17i1.3965>
- Pebdika, A., Herdiana, R., Solihudin, D., Pintar, P. I., Bayes, N., Penentuan, U., & Bantuan, P. (2023). *Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip*. 7(1), 452–458.
- Pratama, R., Perkasa, P., & Rusdi, R. (2022). *Peningkatan Hasil Belajar Mata Pelajaran Komputer Dan Jaringan Dasar Siswa X TKJ Melalui Model Problem Based Learning Di SMK Maestro Banjarmasin*. 5(3), 26–33.
- Rahayu, I. P., Fauzi, A., & Indra, J. (2022). *Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine*. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 296. <https://doi.org/10.30865/json.v4i2.5381>
- Setups, C. N., Subbiah, S., Dheeraj, R., Hidayatullah, A. F., Hakim, A. M., & Gupta, H. (2021). *Sentiment Analysis using Neural Network and Sentiment Analysis using Neural Network and LSTM*. <https://doi.org/10.1088/1757-899X/1074/1/012007>
- Sita, F., Muttaqin, M. S., & Bhinika, P. (2024). *Strategi Pengelolaan Fasilitas dan Sumber Daya di Sekolah Untuk Meningkatkan Kualitas Pembelajaran*. 1, 117–132.
- Syahr, A. U. T., Savitri, A., Widyawati, D., & Ma'Tang, H. (2024). *Analisis Sentimen Fasilitas Belajar dan Alat Laboratorium menggunakan metode Naive Bayes Classifier*. *Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI)*, 7(2), 312–325. <https://doi.org/10.57093/jisti.v7i2.265>
- Umarani, V., Julian, A., & Deepa, J. (2021). *Sciences Sentiment Analysis using various Machine Learning and Deep Learning Techniques*. 3, 385–394.
- Varma, D., Nehansh, A., & Guide, P. S. (2023). *Data Preprocessing Toolkit: An Approach to Automate Data Preprocessing*. 1–5. <https://doi.org/10.55041/IJSREM18270>
- Wiratama Putra, T., Triayudi, A., & Andrianingsih, A. (2022). *Analisis Sentimen Pembelajaran Daring Menggunakan Metode Naive Bayes, KNN, dan Decision Tree*. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 6(1), 20–26. <https://doi.org/10.35870/jtik.v6i1.368>