

Tingkat Kepuasan "Lazy Night" Menggunakan Metode Naive Bayes

Satisfaction Level of "Lazy Night" Using Naive Bayes Method

Kusdiansyah¹, Muhamad Afif², M. Syafiuddin Usman³

^{1,2,3} S1 Informatika, Universitas Teknologi Digital, Bandung, Indonesia,
Email Correspondency: kusdiansyah20122018@digitechuniversity.ac.id

Info Artikel

Riwayat Artikel:
Diajukan: 07/07/2025
Diterima: 08/07/2025
Diterbitkan: 27/08/2025

Kata Kunci:
Naive Bayes, Machine Learning,
Umpulan Balik Acara, Tingkat
Kepuasan, Klasifikasi Teks

Keyword:
Naive Bayes, Machine Learning, Event
Feedback, Satisfaction Level, Text
Classification



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

<https://doi.org/>
© 2024 iTech

A B S T R A K

Dalam era informasi saat ini, memahami dan menafsirkan feedback dari publik menjadi semakin penting, terutama dalam konteks penyelenggaraan acara. Proses menggali pola atau informasi berharga dari kumpulan data telah menjadi topik yang menarik perhatian banyak pihak. Penelitian ini hadir untuk menjawab kebutuhan tersebut, berfokus pada upaya mengklasifikasikan sentimen para peserta acara "Lazy Night" secara otomatis. Kami memanfaatkan algoritma Naive Bayes, sebuah metode yang dikenal efisien dalam analisis teks. Data penelitian kami dikumpulkan langsung dari formulir umpan balik yang diisi oleh peserta, yang berisi beragam opini dan komentar dalam bentuk tekstual. Setiap masukan kemudian kami telaah dan labeli secara manual ke dalam dua kategori sentimen yang jelas: positif atau negatif. Setelah melalui serangkaian tahapan pra-pemrosesan data yang cermat—mulai dari mengubah format teks hingga mengekstraksi fitur-fitur penting—model Naive Bayes kemudian kami bangun untuk melakukan tugas klasifikasi ini. Hasil evaluasi menunjukkan bahwa pendekatan ini sangat efektif dalam menangkap nuansa sentimen dalam data teks yang relatif pendek, seperti umpan balik acara. Model kami berhasil mencapai tingkat akurasi yang memuaskan, yakni 97,4%. Ini membuktikan bahwa Naive Bayes adalah alat yang andal untuk memahami persepsi peserta, memberikan wawasan berharga bagi penyelenggara acara untuk terus meningkatkan kualitas dan relevansi di masa mendatang.

A B S T R A C T

In today's information age, understanding and interpreting public feedback has become increasingly important, especially in the context of event management. The process of extracting patterns or valuable insights from data sets has become a topic of considerable interest. This research addresses this need, focusing on automatically classifying the sentiments of "Lazy Night" event attendees. We utilize the Naive Bayes algorithm, a method known for its efficiency in text analysis. Our research data was collected directly from feedback forms completed by participants, which contained a variety of opinions and comments in text form. We then manually analyzed each piece of feedback and labeled it into two distinct sentiment categories: positive or negative. After a series of meticulous data preprocessing steps—from transforming the text to extracting key features—we then built a Naive Bayes model to perform this classification task. Evaluation results show that this approach is highly effective in capturing sentiment nuances in relatively short text data, such as event feedback. Our model achieved a satisfactory accuracy rate of 97.4%. This demonstrates that Naive Bayes is a reliable tool for understanding attendee perceptions, providing valuable insights for event organizers to continuously improve quality and relevance in the future.

1. PENDAHULUAN

Dalam era informasi saat ini, memahami dan menafsirkan feedback dari publik menjadi semakin penting (Anggraini dkk., 2021; Anggraini & Suroyo, 2019). Dalam konteks sebuah acara, mengumpulkan umpan balik dari para peserta telah menjadi salah Dalam setiap acara yang kita selenggarakan, entah itu konser musik yang meriah, lokakarya edukatif yang mendalam, atau pertemuan komunitas yang akrab, keberhasilan sejati seringkali diukur bukan hanya dari jumlah kehadiran, tetapi juga dari resonansi yang ditinggalkannya di hati para peserta. Bagaimana perasaan mereka? Apakah pengalaman yang kami sajikan sesuai harapan? Jawaban atas pertanyaan-pertanyaan ini terkandung dalam umpan balik yang mereka berikan. Mengumpulkan dan memahami umpan balik ini menjadi elemen krusial untuk mengevaluasi dampak dan menyempurnakan acara di masa mendatang. Salah satu cara yang ampuh untuk menyelami lautan opini ini adalah melalui analisis sentimen. Ini adalah proses cerdas yang memungkinkan kita untuk mengidentifikasi dan mengkategorikan setiap komentar atau ulasan — apakah itu bernada positif, negatif, atau netral. Bayangkan betapa berharganya jika kita bisa dengan cepat mengetahui bagian mana dari acara "Lazy Night" yang paling disukai, atau di area mana ada ruang untuk perbaikan, hanya dari ribuan komentar yang masuk.

Untuk mewujudkan pemahaman ini, kami membutuhkan alat yang tepat. Dalam dunia klasifikasi teks, Algoritma Naive Bayes muncul sebagai pilihan yang menarik. Algoritma ini terkenal karena kesederhanaan implementasinya dan efektivitasnya yang telah terbukti, terutama untuk kumpulan data berukuran kecil hingga menengah. Naive Bayes bekerja dengan menganalisis kemungkinan setiap kata muncul dalam kategori sentimen tertentu, memberikan panduan yang jelas dalam memprediksi apakah suatu umpan balik secara keseluruhan bersifat positif atau negatif. Oleh karena itu, penelitian ini secara khusus memanfaatkan algoritma Naive Bayes untuk menggali sentimen yang terkandung dalam umpan balik dari para peserta acara "Lazy Night". Kami percaya, dengan pendekatan ini, kita dapat memperoleh wawasan berharga yang tidak hanya membantu menilai kesuksesan acara ini, tetapi juga menjadi fondasi untuk penyelenggaraan acara yang lebih baik di masa depan.

2. METODE PELAKSANAAN



Gambar 1. Use Case Diagram

2.1 Dataset

Bayangkan kita ingin tahu apa yang sebenarnya dirasakan oleh para peserta acara "Lazy Night". Untuk itu, kami meminta mereka mengisi sebuah formulir digital (Google Form) yang berisi kolom komentar. Nah, semua komentar atau "umpan balik" dari para peserta inilah yang menjadi "bahan mentah" penelitian kami, atau yang kami sebut sebagai *dataset*. Contoh data:

Dataset ini pada dasarnya adalah kumpulan informasi yang sangat sederhana, terdiri dari dua jenis data utama:

Tim	Departemen	Nama Lengkap	Bagaimana penilaian anda tentang (Saya Mau Bicara mata pelajaran anda terhadap (Pengajar), disiplin dan penilaian anda terhadap (Soal) dan penilaian anda terhadap (Buku kerja)	5 YA	4 YA	3 YA	2 YA	1 YA
1	Tim Manajemen	DEPARTMENT	Unggah	4	3	5	2	1
2	08/12/2024 20:05	Garment	Diksi Indriyani	3	3	3	2	1
3	08/12/2024 20:13	Garment	Lia Wanyani	3	4	3	2	1
4	08/12/2024 20:13	Procurement	VINA PERMATA SETIAWAKA	3	3	3	2	1
5	08/12/2024 20:39	TFU						
6	***	***	***	***	***	***	***	***
701	08/13/2024 13:28:04	Factory	Teti Sofiani	3	3	3	2	1
702	08/13/2024 13:38:51	Product	Ravina Taryana	3	3	4	2	1
703	05/15/2024 14:18:51	procurement	Ujang Setiawan	5	5	5	2	1
704	06/15/2024 14:19:21	(re)procurement	Abdul Amali	5	5	5	2	1
705	06/15/2024 14:19:21	Factory	Kusdiansyah	5	5	5	2	1

Gambar 2. Pengumpulan Data (Dataset)

Teks Umpan Balik : Ini adalah bagian yang paling menarik! Di sinilah peserta menuliskan semua kesan, saran, atau kritik mereka, persis seperti yang mereka ketikkan. Tidak ada perubahan atau penyaringan di sini.

Label Tingkat Kepuasan : Setelah kami mengumpulkan semua teks umpan balik, tim kami dengan cermat membaca setiap komentar dan menentukan apakah perasaan di baliknya itu 'Puas' (mereka menyukainya) atau 'Tidak' (ada hal yang kurang atau perlu diperbaiki). Proses ini kami lakukan secara manual, memastikan setiap komentar mendapatkan label yang sesuai dengan makna sebenarnya.

2.2 Pra-pemrosesan Data (Preprocessing)

Sebelum data teks ini "diajarkan" ke model komputer, kami perlu sedikit merapikannya, ibaratnya mempersiapkan bahan makanan sebelum dimasak. Proses ini kami sebut pra-pemrosesan data, dan tahapannya meliputi:

- **Case Folding:** Semua teks kami ubah menjadi huruf kecil. Jadi, "LUAR BIASA!" dan "luar biasa!" akan dianggap sama, agar komputer tidak bingung.
- **Punctuation Removal:** Kami menyingkirkan semua tanda baca (koma, titik, tanda seru) dan simbol-simbol yang tidak relevan. Fokus kami adalah pada kata-kata inti.
- **Tokenizing:** Setiap kalimat panjang kami pecah menjadi kata-kata individual, seperti memisahkan biji-bijian dari tangkainya. Ini membantu komputer menganalisis setiap kata secara terpisah.
- **Stopword Removal:** Ada beberapa kata yang sangat umum (seperti "dan", "di", "yang") yang sebenarnya tidak membawa makna sentimen berarti. Kata-kata ini kami hapus agar model lebih fokus pada kata-kata kunci yang penting.
- **Label Encoding:** Label 'Positif' dan 'Negatif' yang tadi kami berikan, kami ubah menjadi angka (misalnya 'Negatif' menjadi 0 dan 'Positif' menjadi 1). Ini karena komputer lebih mudah memproses angka daripada teks.

2.3 Algoritma Naive Bayes

Naive Bayes adalah algoritma klasifikasi probabilistik yang didasarkan pada *Teorema Bayes* dengan asumsi independensi antar fitur (kata). Intinya, algoritma ini menghitung seberapa besar kemungkinan sebuah komentar termasuk dalam kategori 'Positif' atau 'Negatif', berdasarkan kemunculan kata-kata di dalamnya. Rumus inti dari *teorema* ini adalah:

$$P(C | X) = \frac{P(X | C) \cdot P(C)}{P(X)}$$

Dimana:

P(C|X): Ini adalah yang ingin kita cari tahu: seberapa mungkin sebuah komentar (X) masuk ke dalam kategori sentimen tertentu (C).

P(X|C): Ini adalah seberapa sering kata-kata dalam komentar (X) muncul dalam kategori sentimen tertentu (C).

P(C): Ini adalah seberapa sering kategori sentimen (C) itu sendiri muncul di seluruh data.

P(X): Ini adalah seberapa sering kata-kata dalam komentar (X) muncul secara keseluruhan di seluruh data.

Sederhananya, Naive Bayes mencari pola: jika komentar "Acaranya seru sekali!" sering muncul di kategori 'Positif', maka komentar baru yang berisi kata "seru" kemungkinan besar juga 'Positif'.

2.4 Evaluasi Model

Setelah model kita "belajar" dari data, tentu kita ingin tahu seberapa pintar ia dalam memprediksi sentimen baru. Untuk itu, kami mengevaluasi kinerjanya menggunakan beberapa tolok ukur standar seperti Akurasi, Presisi, Recall, dan F1-Score.

Agar pengujian ini adil, kami membagi *dataset* menjadi dua bagian:

- **Data Pelatihan (80%):** Bagian terbesar ini digunakan untuk "mengajari" model kami mengenali pola sentimen. Ini adalah tahap di mana model belajar dari contoh-contoh yang sudah kami labeli sebelumnya.
- **Data Pengujian (20%):** Bagian yang lebih kecil ini adalah "ujian" sesungguhnya bagi model. Ini adalah data rahasia yang belum pernah dilihat model sebelumnya. Kami menggunakan untuk menguji seberapa baik model memprediksi sentimen pada komentar yang benar-benar baru, seolah-olah ini adalah umpan balik segar yang baru masuk.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Penelitian

Setelah model Naive Bayes kami dilatih dan diuji dengan data umpan balik, inilah hasilnya: Analisis rata-rata penilaian peraspek:

```
# Analisis rata-rata penilaian per aspek
aspek_penilaian = [col for col in df.columns if col.startswith("Bagaimana penilaian anda terhadap")]
print("\n[R] Rata-rata Penilaian Per Aspek:")
print(df_filtered[aspek_penilaian].mean(numeric_only=True).round(2))
```

Gambar 3. Analisis rata-rata penilaian per-aspek

Rata-rata Penilaian Per Aspek:	
Bagaimana penilaian anda terhadap [Bazar Makanan]	4.47
Bagaimana penilaian anda terhadap [Pengisi Acara]	4.48
Bagaimana penilaian anda terhadap [Dekorasi]	4.60
Bagaimana penilaian anda terhadap [Suasana]	4.63
dtype:	float64

Gambar 4. Ini adalah rating 1 sampai 5

```
# Ringkasan persetujuan kegiatan
print("\n[R] Ringkasan Jawaban tentang Acara Kembali:")
print(df_filtered[kolom_kunci].value_counts())
```

Gambar 5. Ringkasan jawaban tentang acara kembali

Ringkasan Jawaban tentang Acara Kembali:	
Menurut teman-teman apakah kegiatan Lazy Night ini perlu diadakan kembali?	
YA	189
TIDAK	5
Name: count, dtype: int64	

Gambar 6. Ini adalah hasil ringkasan jawaban

3.2. Pengaruh Jenis Informasi yang Dibagikan

Melihat tabel di atas, model Naive Bayes menunjukkan performa yang sangat menjanjikan dalam mengklasifikasikan sentimen, bahkan dengan jumlah data yang mungkin tidak terlalu besar. Ini artinya, model kami cukup cerdas dalam membedakan antara komentar positif dan negatif.

Namun, seperti halnya manusia, model pun terkadang bisa "salah paham". Kesalahan klasifikasi ini umumnya terjadi pada umpan balik yang bahasanya sedikit membingungkan atau mengandung sindiran/sarkasme, di mana konteks kalimat lebih kompleks.

Mari kita lihat beberapa contoh:

- **Contoh Klasifikasi Benar:** Teks: "Yah!" → Model kami dengan tepat mengklasifikasikannya sebagai: Positif. (Ini jelas dan mudah dipahami).

- **Contoh Klasifikasi Salah:** Teks: "Tidak" → Model kami mengklasifikasikannya sebagai: Negatif.

Analisis Kesalahan: Dalam kasus ini, kata "Tidak" mutlak menandakan bahwa peserta *Lazy Night* ningthmodel belum sepenuhnya memahami bahwa meskipun ada kata positif, ungkapan keseluruhan menunjukkan ketidakpuasan. Ini adalah tantangan umum dalam analisis sentimen teks yang kompleks, di mana model perlu belajar lebih dalam tentang konteks dan idiom bahasa.

4. KESIMPULAN

Dari penelitian ini, kami dapat menyimpulkan bahwa algoritma Naive Bayes terbukti menjadi metode yang efektif dan efisien untuk analisis sentimen pada dataset umpan balik acara dengan format teks pendek. Meskipun implementasinya relatif sederhana, algoritma ini mampu menghasilkan akurasi yang kompetitif. Artinya, kita bisa mendapatkan gambaran yang cukup akurat tentang bagaimana perasaan peserta acara hanya dengan mengandalkan data teks mereka.

Untuk penelitian di masa mendatang, kami sangat menyarankan untuk menggunakan dataset yang jauh lebih besar dan lebih beragam. Semakin banyak data yang dipelajari model, semakin cerdas dan akurat pula prediksinya. Selain itu, pengembangan metode yang lebih canggih untuk menangani nuansa bahasa yang lebih kompleks, seperti sarkasme atau ambiguitas (kalimat yang bisa diartikan ganda), juga bisa menjadi fokus utama. Ini akan membantu meningkatkan akurasi model lebih lanjut dan membuatnya semakin "memanusiakan" dalam memahami perasaan di balik kata-kata.

5. SARAN

1. Ajak Model Anda "Ngobrol" Lebih Dalam (Mengatasi Sarkasme):

Saat ini, model Naive Bayes Anda sudah cerdas, mencapai akurasi 97,4%. Namun, ia masih sering "salah paham" pada komentar yang bahasanya membingungkan, mengandung sindiran, atau sarkasme.

Saran: Coba ajari model Anda cara "membaca konteks" lebih baik. Anda bisa menggunakan teknik yang lebih modern, seperti word embeddings atau N-gram, agar model tidak hanya melihat kata per kata, tetapi juga melihat hubungan antar kata. Ini akan membuatnya lebih "manusiawi" dalam memahami nuansa di balik kata-kata.

Platform media sosial harus memberikan lebih banyak transparansi dan kontrol kepada pengguna terkait pengumpulan dan penggunaan data mereka. Dengan mengambil langkah-langkah ini, diharapkan dapat menciptakan lingkungan media sosial yang lebih aman dan menghormati privasi setiap individu sesuai dengan spirit Undang-Undang Teknologi Informasi.

2. Coba "Alat" Klasifikasi Lain:

Naive Bayes sudah terbukti andal, tapi di dunia machine learning ada banyak "alat" lain yang bisa Anda coba, seperti SVM atau model berbasis deep learning.

Saran: Lakukan perbandingan. Uji algoritma lain untuk melihat apakah mereka bisa mengatasi tantangan bahasa kompleks (seperti sindiran) dengan lebih baik, sehingga akurasi Anda bisa meningkat lebih jauh.

3. Gabungkan Angka dan Teks:

Anda sudah punya data rating per aspek (Bazar Makanan, Suasana, dll.) dan juga data teks umpan balik.

Saran: Gabungkan kedua hasil ini! Tautkan sentimen negatif dari teks ke aspek yang mendapat rating terendah. Ini akan memberi tahu penyelenggara secara spesifik: "Aspek 'Bazar Makanan' mendapat nilai terendah (4.47), dan inilah semua komentar negatif yang secara spesifik membahas makanan tersebut." Wawasan ini sangat berharga untuk perbaikan acara di masa depan.

DAFTAR PUSTAKA

- Anggraini, N., Harahap, E. S. N., & Kurniawan, T. B. (2021). Text Mining—Analisis Teks Terkait Isu Vaksinasi COVID-19 (Text Mining—Text Analysis Related to COVID-19 Vaccination Issues). *JURNAL IPTEKKOM (Jurnal Ilmu Pengetahuan & Teknologi Informasi)*, 23(2), 141–153. <https://doi.org/10.17933/iptekkom.23.2.2021.141-153>

- Anggraini, N., & Suroyo, H. (2019). Comparison of Sentiment Analysis against Digital Payment “T-cash and Go-pay” in Social Media Using Orange Data Mining. *Journal of Information Systems and Informatics*, 1(2), 152–163. <https://doi.org/10.33557/journalisi.v1i2.21>
- Atina, V., & Sudibyo, N. A. (2023). Pemodelan Prediksi Kelulusan Mahasiswa Dengan Metode Naïve Bayes Di Uniba. *Jurnal Manajemen Informatika dan Sistem Informasi*, 6(2), 148-158.
- Damanik, A. R., Sumijan, S., & Nurcahyo, G. W. (2021). Prediksi tingkat kepuasan dalam pembelajaran daring menggunakan algoritma Naïve Bayes. *Jurnal Sistim Informasi Dan Teknologi*, 88-94.
- Guntur, M., Santony, J., & Yuhandri, Y. (2018). Prediksi harga emas dengan menggunakan metode Naïve Bayes dalam investasi untuk meminimalisasi resiko. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 2(1), 354-360.
- Setiyani, L., Wahidin, M., Awaludin, D., & Purwani, S. (2020). Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes: Systematic Review. *Faktor Exacta*, 13(1), 35-43.
- Khalim, K. A., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 498-504.
- Saepudin, P. A. D., & Usman, M. S. (2024). Prediksi Kehadiran Mahasiswa Universitas Teknologi Digital Menggunakan Naïve Bayes. *TEKNO: Jurnal Penelitian Teknologi dan Peradilan*, 2(2), 16-40.