

Klasifikasi User Review pada Aplikasi Online Travel Booking Menggunakan Multinomial Naïve Bayes

Mohammad Yoga Pratama¹, Cahyo Crysdiyan²

200605110042@student.uin-malang.ac.id, cahyo@ti.uin-malang.ac.id

^{1,2}Universitas Islam Negeri Maulana Malik Ibrahim

Informasi Artikel

Diterima : 26 Jun 2024

Direview : 30 Apr 2025

Disetujui : 9 Jun 2025

Abstrak

Perkembangan teknologi yang pesat telah membawa perubahan dalam berbagai aspek kehidupan, termasuk dalam sektor pariwisata. Aplikasi tiket travel online seperti Traveloka, pegipegi dan Tiket.com merupakan aplikasi travel yang sangat populer di Indonesia. penelitian ini bertujuan untuk mengukur performa dari metode Multinomial Naïve Bayes dalam mengklasifikasikan ulasan pengguna aplikasi tersebut menjadi kelas "satisfied" dan "unhappy". Dataset berjumlah 1339 ulasan pengguna yang diambil dari Google Play Store. Uji coba dilakukan dengan tiga skenario rasio pembagian dataset (7:3, 8:2, 9:1) dan dievaluasi menggunakan confusion matrix dan K-Fold Cross Validation. Hasil uji coba menunjukkan skenario pembagian data 9:1 menghasilkan akurasi model tertinggi sebesar 81.34% dengan precision 81.47%, recall 81.44% dan F1-Score 81.34%. Analisa kata menggunakan TF-IDF menunjukkan bahwa kata-kata seperti "good", "nice" dan "nice" mendominasi pada kelas "satisfied", sedangkan kata seperti "price", "cant", dan "app" merupakan 3 kata yang paling mendominasi pada kelas "unhappy". Dapat disimpulkan bahwa metode Multinomial Naïve Bayes memiliki performa yang baik untuk klasifikasi ulasan pengguna aplikasi travel online, dan semakin banyak dataset yang digunakan makan semakin bagus pula model yang dihasilkan.

Kata Kunci

Klasifikasi Ulasan,
Multinomial Naïve
Bayes, Online Travel
Agent

A. Pendahuluan

Perkembangan teknologi yang pesat telah membawa perubahan dalam berbagai aspek kehidupan, termasuk dalam sektor pariwisata [1]. Saat ini, pemesanan tiket travel semakin mudah diakses dan praktis dengan adanya aplikasi pemesanan tiket travel online. Aplikasi ini memberikan kemudahan dan benefit bagi penggunanya, seperti akses mencari informasi tentang destinasi wisata, membandingkan pilihan tiket perjalanan mulai dari, maskapai penerbangan, promo dan diskon, waktu keberangkatan hingga fasilitas yang diberikan. Selain itu, proses pemesanan tiket yang mudah dan cepat, serta berbagai variasi harga yang kompetitif.

Bisnis di sektor pariwisata telah berkembang dengan pesat. Menurut data Badan Statistik (BPS) menunjukkan wisatawan nusantara (wisnus) meningkat 12,02 persen pada November 2023 dibandingkan dengan bulan yang sama pada tahun 2022, hal ini menjadi yang tertinggi sejak 5 tahun terakhir sejak 2019. Sementara untuk jumlah wisatawan mancanegara (wisman) pada November 2023 secara kumulatif meningkat 110,86 persen dibandingkan dengan dengan bulan yang sama di tahun 2022 yang dimana ini merupakan peningkatan wisatawan yang sangat drastis. Berdasarkan survei yang dilakukan oleh Avara Research didapatkan Traveloka digunakan oleh 79 persen responden untuk memesan tiket pesawat karena memiliki banyak promo. Sedangkan aplikasi travel lainnya seperti Tiket.com digunakan oleh 8,9 persen responden [2].

Sebelum memilih aplikasi travel online, calon pengguna biasanya membaca ulasan yang ada di berbagai platform, seperti media sosial Google Play Store dan App Store. Selain melihat rating aplikasi, ulasan ini menjadi sumber informasi penting bagi mereka untuk memprediksi kualitas layanan dan membuat keputusan mengenai aplikasi mana yang akan digunakan. Ulasan biasanya berisi pengalaman pengguna selama menggunakan aplikasi, bisa berupa ekspresi kepuasan (satisfied) atau keluhan (unhappy) terhadap aplikasi. Dengan membaca ulasan pengguna, dapat membandingkan dan mengidentifikasi aplikasi yang sesuai dengan kebutuhan mereka. Selain itu, ulasan positif dapat meningkatkan kepercayaan pengguna terhadap aplikasi. Bagi pelaku bisnis aplikasi travel, ulasan pengguna merupakan aset berharga yang dapat membantu meningkatkan kualitas mulai dari, layanan teknis hingga layanan yang ada dalam aplikasi (promo, kemudahan akses, keamanan dan lain-lain), sehingga dapat mempertahankan pengguna. Menentukan kepuasan pengguna hanya dengan berpatokan dengan rating saja tidak cukup untuk menggambarkan kredibilitas aplikasi, oleh karena itu ulasan menjadi bukti yang valid [3]. Proses klasifikasi ulasan dengan cara manual akan sangat memakan waktu yang lama, rentan terhadap subjektivitas dan kesalahan. Oleh karena itu, kebutuhan sistem klasifikasi ulasan otomatis menjadi penting agar pengembang dapat memahami lebih mendalam apa yang dirasakan pengguna tentang aplikasi. Dengan adanya sistem klasifikasi dapat mempermudah dan mempercepat proses klasifikasi. Pengembang dapat memprioritaskan perbaikan pada aspek yang paling dikeluhkan pengguna dan meningkatkan kualitas aplikasi secara keseluruhan. Dengan memahami kebutuhan pengguna, pelaku bisnis dapat melakukan pengembangan aplikasi menjadi lebih terarah dan memenuhi ekspektasi pengguna.

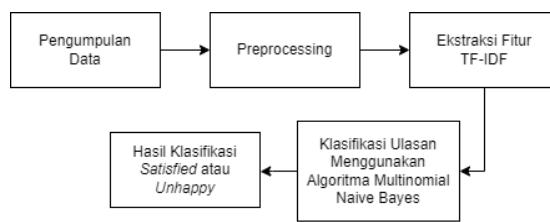
Penelitian sebelumnya membahas tentang analisis sentimen yang dilakukan oleh Diekson [4] dengan menganalisis sentimen ulasan pengguna pada aplikasi Traveloka menggunakan metode Naïve Bayes dengan akurasi 82,91%.

Penelitian sebelumnya oleh Yuyun et al., (2021) [5] dalam penelitiannya tentang klasifikasi sentimen publik terkait penanganan Covid-19 di Indonesia menggunakan metode Multinomial Naïve Bayes. Data yang digunakan pada penelitian ini berasal dari Twitter. Hasil dari penelitian ini tentang klasifikasi sentimen publik terkait penanganan Covid-19 di Indonesia menggunakan metode Multinomial Naïve Bayes nilai akurasi sebesar 74%, presisi 74%, dan recall 74%.

Oktaviani et al., (2021) [6] membahas analisis sentimen ulasan pengguna aplikasi e-commerce Traveloka yang ada di Google Playstore. Penelitian ini menggunakan metode Naïve Bayes. Ulasan pengguna kemudian diklasifikasikan kedalam kategori positif dan negatif. Hasil yang diperoleh dari ulasan yang diberikan pengguna aplikasi Traveloka diperoleh akurasi model Naïve Bayes sebesar 91.20%. Berdasarkan penelitian sebelumnya kami mengajukan metode Multinomial Naïve Bayes dengan ekstraksi fitur TF-IDF untuk melakukan klasifikasi ulasan pengguna aplikasi travel online. Desain sistem berikut dapat memberikan gambaran dari alur sistem klasifikasi mulai dari input data sampai hasil klasifikasi.

B. Metode Penelitian

Metode yang digunakan pada penelitian ini untuk melakukan klasifikasi ulasan adalah Multinomial Naïve Bayes. Berikut alur dari penelitian ini.



Gambar 1. Desain Sistem

1. Akuisisi Data

Sumber data yang digunakan berasal dari platform Google Play Store. Untuk melakukan pengumpulan data pada Google Play Store menggunakan library google play scraper. Library ini menyediakan API untuk melakukan scraping pada Google Play Store dengan mudah [7]. Data Ulasan yang diambil dalam bahasa Inggris, berasal tiga aplikasi travel online yaitu Traveloka, pegipegi, dan Tiket.com. Jumlah data yang diambil sebanyak 1339 data dalam rentang waktu 18 April 2017 sampai 12 September 2023.

Tabel 1. Data

No	Aplikasi	Size
1	Traveloka	499
2	tiket.com	432
3	pegipegi	408
Total		1339

2. Preprocessing

Preprocessing bertujuan untuk membersihkan data agar mudah dipahami dan dianalisis oleh berbagai algoritma machine learning, sehingga siap untuk diproses [7]. Data akan melalui empat tahap preprocessing yaitu cleaning, case folding, stopword, dan lemmatization.

2.1 Cleaning

Tahap preprocessing pertama pada penelitian ini adalah cleaning. Pada tahapan ini bertujuan untuk menghapus emoticon, tanda baca (punctuation), URL dan ruang kosong (white space) yang berlebihan. Tanda baca yang dihapus pada data ulasan pengguna seperti `(!\"#$%&()^+-.:/;<=>?@[{}]^_{}~\n)`. Selain itu, untuk emoticon yang dihapus mencakup , , ,  dll.

2.2 Case Folding

Tahap kedua dari preprocessing adalah case folding. Tahapan ini bertujuan untuk mengubah huruf kapital menjadi huruf kecil [8]. Hal ini dilakukan untuk membuat semua huruf menjadi seragam dan juga mengurangi beban komputasi.

2.3 Stopword

Stopword adalah proses menghapus kata pada kalimat yang tidak memiliki makna kontekstual seperti "a", "an", "the", "was", "in" dll [9]. Tahap ini membantu dalam memfokuskan analisis pada kata-kata yang benar-benar relevan.

2.4 Lemmatization

Tahap *preprocessing* terakhir pada penelitian ini adalah *lemmatization*. *Lemmatization* adalah proses mengubah kata menjadi kata dasar dengan mengetahui konteks dari kata tersebut [7]. Berbeda dengan *stemming*, *lemmatization* lebih akurat karena mempertimbangkan konteks dari makna kata tersebut.

3. TF-IDF

Tahap selanjutnya dalam penelitian ini adalah ekstraksi fitur. Term Frequency-Inverse Document Frequency (TF-IDF) adalah algoritma pembobotan kata yang digunakan untuk mengevaluasi pentingnya kata tertentu dalam sebuah ulasan. Besaran dari nilai TF (Term Frequency) bergantung pada seberapa banyak sebuah kata itu muncul dalam ulasan. Semakin sering sebuah kata muncul dalam ulasan, maka kata tersebut dianggap penting untuk dokumen tertentu. Sedangkan IDF (Inverse Document Frequency) adalah seberapa sering sebuah kata tersebut muncul di seluruh ulasan yang ada.

4. Multinomial Naïve Bayes

Algoritma Multinomial Naïve Bayes merupakan salah satu algoritma pembelajaran probabilistik yang digunakan pada Natural Language Processing (NLP) [10]. Algoritma ini salah satu pengembangan dari algoritma Naïve Bayes yang digunakan untuk melakukan klasifikasi teks, karena dataset pada penelitian ini berupa teks maka metode ini sangat cocok untuk melakukan klasifikasi teks. Algoritma Multinomial Naïve Bayes bekerja dengan menghitung probabilitas kemunculan kata dalam sebuah dokumen [5]. Algoritma ini memanfaatkan konsep probabilistik yang didasarkan pada Teorema Bayes.

C. Hasil dan Pembahasan

Proses pelabelan dataset ulasan pengguna aplikasi tiket *travel online* dilakukan menggunakan metode *crowdsourcing*. Responden diminta untuk

memberikan label “*satisfied*” jika ulasan tersebut menggambarkan kepuasan dan “*unhappy*” jika ulasan menunjukkan ketidakpuasan atau terdapat keluhan. Proses ini mengurangi subjektivitas dan memastikan akurasi pelabelan.

Tabel 2. Sampel Dataset

No	Identitas	Size
1	good, fast respon if any trouble, good solution	Satisfied
2	no verification code when purchase with credit card	Unhappy
3	your website needs attention as it gets stuck at the payment page	Unhappy
4	Traveloka is very useful and user friendly, we can get best flight or deals	Satisfied
5	I have been using this for years, now it's even easier. Of course there's room for improvement.	Satisfied

Dataset yang telah dikumpulkan dan melalui tahap *preprocessing* akan dibagi menjadi dua bagian, yaitu data *train* dan data *test*. Pembagian rasio pengujian data *train* dan data *test* mengacu pada penelitian sebelumnya dilakukan oleh [11] dataset akan dibagi menjadi tiga rasio, yaitu 7:3, 8:2, dan 9:1.

Tabel 3. Pembagian Dataset

Pengujian	Ratio	Jumlah Data	
		Train	Test
1	7:3	937	402
2	8:2	1071	268
3	9:1	1205	134

Berdasarkan hasil ujicoba dengan tiga skenario pembagian data, idapatkan skenario uji coba pada pembagian data 9:1 mendapatkan akurasi tertinggi yaitu sebesar 81.34%, ini menunjukkan bahwa model mampu dengan baik dalam melakukan klasifikasi model. Akurasi terbaik selanjutnya diikuti oleh skenario 7:3 dengan akurasi sebesar 80.84%, kemudian skenario diikuti dengan skenario pembagian data 8:2 dengan akurasi sebesar 79.48%, dapat dilihat pada Tabel. 4 Hasil Uji Coba berikut.

Tabel 4. Hasil Ujicoba

Uji Coba	Ratio Pembagian Data	Accuracy	Precision	Recall	F1-Score
1	7:3	80.84%	80.97%	80.94%	80.84%
2	8:2	79.48%	79.71%	79.60%	79.47%
3	9:1	81.34%	81.47%	81.43%	81.34%

Hasil evaluasi performa yang dilakukan menggunakan *confusion matrix* pada tiap uji skenario pembagian data yang telah ditentukan, dapat diketahui bahwa skenario pengujian pada rasio pembagian data 9:1 memiliki nilai akurasi yang tinggi diantara rasio pembagian data lainnya. Hal ini menunjukkan konsistensi model yang stabil dan memiliki kualitas data yang cenderung seimbang, dimana terdapat penyebaran data yang maksimal. Persebaran kata pada data *training* yang digunakan lebih banyak dari pada rasio pembagian lainnya, sehingga model lebih mampu untuk melakukan prediksi dengan maksimal karena memiliki kosakata yang lebih banyak. Hal ini membuat model menjadi mudah dalam melakukan klasifikasi

ketika ada data baru yang akan diprediksi. Dengan begitu dapat menunjukkan bahwa proporsi data pelatihan yang lebih besar cenderung dapat menghasilkan model yang lebih baik.

D. Simpulan

Berdasarkan hasil evaluasi performa menggunakan confusion matrix, uji coba skenario-3 dengan pembagian data 9:1 menghasilkan akurasi tertinggi sebesar 81.34% dengan presisi 81.47%, recall 81.43%, dan F1-Score 81.34%, diikuti oleh uji coba skenario-1 dengan akurasi model sebesar 80.84% dan akurasi terendah uji coba skenario-2 dengan akurasi model sebesar 79.48%. Hal ini menunjukkan bahwa model Multinomial Naïve Bayes mampu melakukan klasifikasi ulasan aplikasi travel online dengan baik. Keragaman data dalam dataset berpengaruh signifikan terhadap performa model. Dataset memiliki variasi yang luas memudahkan model dalam melakukan prediksi data baru.

E. Referensi

- [1] Raja Suhud, "Tren Penggunaan Aplikasi Penjualan Tiket bakal Melejit," *Media Indonesia*, 2023. <https://mediaindonesia.com/ekonomi/554815/tren-penggunaan-aplikasi-penjualan-tiket-bakal-melejit> (accessed Feb. 09, 2024).
- [2] CNN Indonesia, "Aplikasi e-Commerce dan Online Travel Pilihan Milenial," *Cnn Indonesia*, 2019. <https://www.cnnindonesia.com/teknologi/20190709195545-185-410676/aplikasi-e-commerce-dan-online-travel-pilihan-milenial> (accessed Jun. 05, 2024).
- [3] V. O. Tama, Y. Sibaroni, and Adiwijaya, "Labeling Analysis in the Classification of Product Review Sentiments by using Multinomial Naïve Bayes Algorithm," in *Journal of Physics: Conference Series*, 2019, vol. 1192, no. 1. doi: 10.1088/1742-6596/1192/1/012036.
- [4] Z. A. Diekson, M. R. B. Prakoso, M. S. Q. Putra, M. S. A. F. Syaputra, S. Achmad, and R. Sutoyo, "Sentiment analysis for customer review: Case study of Traveloka," *Procedia Comput. Sci.*, vol. 216, no. 2022, pp. 682–690, 2023, doi: 10.1016/j.procs.2022.12.184.
- [5] Yuyun, Nurul Hidayah, and Supriadi Sahibu, "Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 820–826, 2021, doi: 10.29207/resti.v5i4.3146.
- [6] V. Oktaviani, B. Warsito, H. Yasin, R. Santoso, and Suparti, "Sentiment analysis of e-commerce application in Traveloka data review on Google Play site using Naïve Bayes classifier and association method," *J. Phys. Conf. Ser.*, vol. 1943, no. 1, 2021, doi: 10.1088/1742-6596/1943/1/012147.
- [7] A. Tabassum and R. R. Patil, "A Survey on Text Pre-Processing & Feature Extraction Techniques in Natural Language Processing," *Int. Res. J. Eng. Technol.*, no. June, pp. 4864–4867, 2020, [Online]. Available: www.irjet.net
- [8] T. Hasan, A. Matin, and M. S. R. Joy, "Machine Learning Based Automatic Classification of Customer Sentiment," *ICCIT 2020 - 23rd Int. Conf. Comput. Inf. Technol. Proc.*, pp. 19–21, 2020, doi: 10.1109/ICCIT51783.2020.9392652.
- [9] D. Ramachandran and R. Parvathi, "Analysis of Twitter Specific Preprocessing

- Technique for Tweets," *Procedia Comput. Sci.*, vol. 165, pp. 245–251, 2019, doi: 10.1016/j.procs.2020.01.083.
- [10] Sriram, "Multinomial Naive Bayes Explained: Function, Advantages & Disadvantages, Applications in 2023," *upgrad*, 2022. <https://www.upgrad.com/blog/multinomial-naive-bayes-explained/> (accessed Oct. 04, 2023).
- [11] Baiq Nurul Azmi, Arief Hermawan, and Donny Avianto, "Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 281–290, 2023, doi: 10.35746/jtim.v4i4.298.