

**KLASIFIKASI ULASAN PENGGUNA APLIKASI PEMESANAN TIKET  
TRAVEL ONLINE MENGGUNAKAN METODE *MULTINOMIAL  
NAÏVE BAYES***

**SKRIPSI**

**Oleh:  
MOHAMMAD YOGA PRATAMA  
NIM. 200605110042**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2024**

**KLASIFIKASI ULASAN PENGGUNA APLIKASI PEMESANAN TIKET  
TRAVEL ONLINE MENGGUNAKAN METODE *MULTINOMIAL  
NAÏVE BAYES***

**SKRIPSI**

Diajukan kepada:  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
Untuk memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:  
**MOHAMMAD YOGA PRATAMA**  
**NIM. 200605110042**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2024**

**HALAMAN PERSETUJUAN**

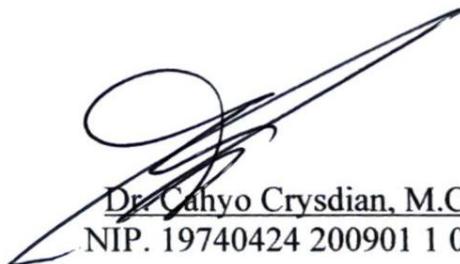
**KLASIKASI ULASAN PENGGUNA APLIKASI PEMESANAN TIKET  
TRAVEL ONLINE MENGGUNAKAN METODE *MULTINOMIAL  
NAÏVE BAYES***

**SKRIPSI**

**Oleh:  
MOHAMMAD YOGA PRATAMA  
NIM. 200605110042**

Telah Diperiksa dan Disetujui untuk Diuji  
Tanggal: 3 Juni 2024

Pembimbing I



Dr. Cahyo Crysdian, M.Cs  
NIP. 19740424 200901 1 008

Pembimbing II



A'la Syauqi, M.Kom  
NIP. 19771201 200801 1 007

Mengetahui,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



M. MT, IPM  
NIP. 19771020 200912 1 001

## HALAMAN PENGESAHAN

# KLASIKASI ULASAN PENGGUNA APLIKASI PEMESANAN TIKET TRAVEL ONLINE MENGGUNAKAN METODE *MULTINOMIAL* *NAÏVE BAYES*

## SKRIPSI

Oleh :  
**MOHAMMAD YOGA PRATAMA**  
NIM. 200605110042

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Sarjana Komputer ( S.Kom )  
Tanggal: 13 Juni 2024

### Susunan Dewan Penguji

Ketua Penguji	: <u>Dr. Zainal Abidin, M.Kom</u> NIP. 19760613 200501 1 004	(  )
Anggota Penguji I	: <u>Tri Mukti Lestari, M.Kom</u> NIP. 19911108202012 2 005	(  )
Anggota Penguji II	: <u>Dr. Cahyo Crysdian, M.Cs</u> NIP. 19740424 200901 1 008	(  )
Anggota Penguji III	: <u>A'la Syauqi, M.Kom</u> NIP. 19771201 200801 1 007	(  )

Mengetahui dan Mengesahkan,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



  
Dr. Firdaus Kurniawan, M.MT, IPM  
NIP. 19771020 200912 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Mohammad Yoga Pratama  
NIM : 200605110042  
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika  
Judul Skripsi : Klasifikasi Ulasan Pengguna Aplikasi Pemesanan  
Tiket *Travel Online* Menggunakan Metode  
*Multinomial Naïve Bayes*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang,  
Yang membuat pernyataan,



Mohammad Yoga Pratama  
NIM. 200605110042

## **MOTTO**

*“Doa mama papa tidak akan berhenti mengalir”*

## **HALAMAN PERSEMBAHAN**

Dengan mengucapkan Alhamdulillah saya dan terima kasih, skripsi ini saya persembahkan kepada :

Allah SWT atas rahmat dan karunia-Nya yang memberikan kekuatan dan kesabaran dalam menyelesaikan skripsi ini.

Mama dan Papa, yang selalu memberikan cinta, dukungan, doa dan semangat yang tiada hentinya. Terima kasih atas segala pengorbanan dan kasih sayang yang telah diberikan selama ini.

Dan diri saya sendiri.

## KATA PENGANTAR

*Assalamualaikum wr wb.*

Puji syukur penulis panjatkan kehadirat Allah Swt. atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian akhir ini yang berjudul “Klasifikasi Ulasan Pengguna Aplikasi Pemesana Tiket *Travel Online* Menggunakan Metode *Multinomial Naïve Bayes*”. Dalam penyusunan karya ini, penulis telah mendapatkan bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Prof. Dr. H.M. Zainuddin, MA., rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Sri Harini, M.Si., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Fachrul Kurniawan, M.MT., IPM, selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Cahyo Crysdiyan, MCS selaku Dosen Pembimbing I yang telah membimbing dan membagikan ilmunya selama bimbingan.
5. A’la Syauqi, M.Kom selaku Dosen Pembimbing II yang telah memberikan banyak koreksi dan saran untuk proses penulisan sehingga karya ini menjadi lebih baik.
6. Dr. Zainal Abidin, M.Kom selaku Ketua Penguji dan Tri Mukti Lestari, M.Kom selaku Dosen Penguji II yang telah menguji isi dari penelitian ini serta memberikan arahan dalam penulisan skripsi ini.

7. Segenap dosen dan jajaran staf program studi Teknik Informatika yang telah memberikan banyak sekali ilmu dan arahan dalam menyelesaikan skripsi ini.
8. Kedua orang tua penulis yang penulis sangat sayangi. Bapak Iwan Yulianto dan Ibu Nurul Hadyante yang senantiasa mendoakan setiap hari dan memberikan dukungan baik moral, motivasi, maupun materi serta kasih sayang sehingga penelitian dapat menyelesaikan skripsi ini dengan baik. Penelitian ini sebagai bentuk tanggung jawab penulis kepada kedua orang tua.
9. Adik-adik Mohammad Rifky Wicaksana dan Mutya Khanza Syaqla yang selalu memberikan semangat kepada penulis.
10. Om koko dan Tante Jum yang selalu memberikan dukungan materi serta motivasi penulis untuk melanjutkan studi ke jenjang selanjutnya.
11. Seluruh keluarga besar penulis yang tidak bisa disebutkan satu-satu.
12. Hikmatul Maulidia Putri, Fahrendra Khoirul Ihtada dan Rizha Alfianita selaku partner seperjuangan penulis, yang selalu memberikan waktu, perhatian dan motivasi hingga penelitian ini selesai
13. Keluarga kecil kos SKD Fahrendra, Faza, Aqza, Fata, Dafa dan Heri yang menemani penulis disaat senang maupun sakit, Sahabat deeptalk Zul Iflah dan Nabila.
14. Squad PUBGM Bagoes, Wahyu, dan Angel yang selalu memberikan support jarak jauh hingga penelitian ini selesai.
15. Teman-teman Angkatan 2020 Teknik Informatika "INTEGER" yang telah memberikan dukungan, semangat, dan kebersamaan selama kuliah dan mengerjakan skripsi ini.

16. Seluruh pihak dan teman-teman penulis lainnya yang tidak dapat penulis sebutkan satu persatu, yang secara tidak langsung juga turut serta memberi dukungan dan membantu penulis dalam pengerjaan skripsi ini,

Sebagai penutup, penulis mengakui bahwa skripsi ini masih terdapat banyak kekurangan. Penulis berharap bahwa skripsi ini dapat membawa manfaat bagi para pembaca dan tentunya penulis sendiri. *Aamiin yaa robbal 'alamin.*

*Wassalamualaikum Warahmatullahi Wabarakatuh.*

Malang, 13 Juni 2024

Penulis

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN</b> .....	<b>iii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iv</b>
<b>PERNYATAAN KEASLIAN TULISAN</b> .....	<b>v</b>
<b>MOTTO</b> .....	<b>vi</b>
<b>HALAMAN PERSEMBAHAN</b> .....	<b>vii</b>
<b>KATA PENGANTAR</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>xi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiii</b>
<b>DAFTAR TABEL</b> .....	<b>xiv</b>
<b>ABSTRAK</b> .....	<b>xv</b>
<b>ABSTRACT</b> .....	<b>xvi</b>
مستخلص البحث .....	<b>xvii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Pernyataan Masalah.....	4
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	5
<b>BAB II STUDI PUSTAKA</b> .....	<b>6</b>
2.1 Kepuasan Pengguna .....	6
2.2 Klasifikasi Ulasan .....	7
<b>BAB III DESAIN DAN IMPLEMENTASI</b> .....	<b>14</b>
3.1 Desain Sistem.....	14
3.2 Akuisisi Data .....	15
3.3 Preprocessing .....	15
3.3.1 Cleaning.....	16
3.3.2 Case Folding.....	16
3.3.3 Stopword .....	17
3.3.4 Lemmatization.....	17
3.4 Ekstraksi Fitur TF-IDF.....	17
3.5 Multinomial Naïve Bayes.....	18
<b>BAB IV UJI COBA DAN PEMBAHASAN</b> .....	<b>21</b>
4.1 Langkah Pengujian .....	21
4.1.1 Pelabelan Data .....	21
4.1.2 Input Dataset .....	21
4.1.3 Pembagian dataset Menjadi 2 Subset .....	22
4.1.4 Pelatihan Model dengan Subset Data Latih .....	23
4.1.5 Pelatihan Model dengan Subset Data Uji .....	23
4.1.6 Evaluasi Performa Model dengan Confuion Matrix.....	24
4.2 Hasil dan Uji Coba .....	25
4.2.1 Hasil Uji Coba Skenario-1 .....	25

4.2.2 Hasil Uji Coba Skenario-2 .....	27
4.2.3 Hasil Uji Coba Skenario-3 .....	28
4.3 Pembahasan.....	30
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>37</b>
5.1 Kesimpulan.....	37
5.2 Saran.....	37
<b>DAFTAR PUSTAKA</b>	
<b>LAMPIRAN</b>	

## DAFTAR GAMBAR

Gambar 3.1 Desain Sistem.....	14
Gambar 3.2 Diagram alur Tahapan Preprocessing .....	16
Gambar 4.1 Confusion Matrix Uji Coba Skenario-1 .....	26
Gambar 4.2 Metrik Performa Uji Coba Skenario-1 .....	27
Gambar 4.3 Confusion Matrix Uji Coba Skenario-2 .....	27
Gambar 4.4 Metrik Performa Uji Coba 1.....	28
Gambar 4.5 Confusion Matrix Uji Coba 3.....	29
Gambar 4.6 Metrik Uji Coba 3 .....	30
Gambar 4.7 Hasil K-Fold Cross Validation Skenario Uji Coba 3 .....	32
Gambar 4.8 Word Cloud kelas Satisfied.....	34
Gambar 4.9 Wordcloud Kelas Unhappy .....	35

## DAFTAR TABEL

Tabel 2.1 Tabel Penelitian Terdahulu .....	11
Tabel 3.1 Tabel Data .....	15
Tabel 4.1 Sampel Dataset.....	22
Tabel 4.2 Pembagian Dataset.....	23
Tabel 4.3 Tabel Confusion Matrix .....	24
Tabel 4.4 Hasil Uji Coba.....	30
Tabel 4.5 Ranking Nilai TF-IDF Kelas Satisfied.....	33
Tabel 4.6 Ranking Nilai TF-IDF Kelas.....	34

## ABSTRAK

Pratama, Mohammad Yoga. 2024. **Klasifikasi Ulasan Pengguna Aplikasi Pemesanan Tiket Travel Online Menggunakan Metode Multinomial Naïve Bayes**. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Cahyo Crysodian, M.Cs (II) A'la Syauqi, M.Kom.

**Kata Kunci:** *Klasifikasi Ulasan, Multinomial Naïve Bayes, Online Travel Agent*

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 maka semakin bagus pula model yang dihasilkan.

## ABSTRACT

Pratama, Mohammad Yoga. 2024. **Classification of Online Travel Ticket Booking Application User Reviews Using the Multinomial Naïve Bayes Method.** Thesis. Informatics Engineering Study Program, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Supervisor: (I) Dr. Cahyo Crysdiyan (II) A'la Syauqi, M.Kom.

The rapid development of technology has brought changes in various aspects of life, including in the tourism sector. Online travel ticket applications such as Traveloka, pegipegi and Tiket.com are very popular travel applications in Indonesia. This research aims to measure the performance of the Multinomial Naïve Bayes method in classifying user reviews of these applications into “satisfied” and “unhappy” classes. The dataset amounted to 1339 user reviews taken from the Google Play Store. Three scenarios of dataset sharing ratio (7:3, 8:2, 9:1) were tested and evaluated using confusion matrix and K-Fold Cross Validation. The test results show that the 9:1 data sharing scenario produces the highest model accuracy of 81.34% with precision 81.47%, recall 81.44% and F1-Score 81.34%. Word analysis using TF-IDF shows that words like “good”, “nice” and “nice” dominate in the “satisfied” class, while words like “price”, “cant”, and “app” are the 3 most dominating words in the “unhappy” class. It can be concluded that the Multinomial Naïve Bayes method has good performance for the classification of online travel application user reviews, and the more datasets used, the better the resulting model.

**Keywords:** *Review Classification, Multinomial Naïve Bayes, Online Travel Agen*

## مستخلص البحث

براتاما، محمد يوغا. 2024. تصنيف مراجعات مستخدمي تطبيق حجز تذاكر السفر عبر الإنترنت بواسطة تطبيق حجز تذاكر السفر عبر الإنترنت باستخدام طريقة بايز الساذج متعدد الحدود. أطروحة. برنامج دراسة هندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: د. كاهيو كريسديان، المشرف الثاني: د. محمد آلا سيوقي، الماجستير. المشرف

الكلمات الرئيسية: تصنيف المراجعات، متعدد الحدود الساذجة متعددة الحدود، وكيل السفر عبر الإنترنت

أحدث التطور السريع للتكنولوجيا تغييرات في مختلف جوانب الحياة، بما في ذلك قطاع السياحة. تُعد تطبيقات حجز تذاكر السفر عبر الإنترنت مثل Traveloka و Pegipegi و Tiket.com من تطبيقات السفر الشائعة جداً في إندونيسيا. تهدف هذه الدراسة إلى قياس أداء طريقة "بايز الساذج متعدد الحدود" في تصنيف مراجعات المستخدمين لهذه التطبيقات إلى فئتين "راضٍ" و"غير راضٍ". بلغ إجمالي مجموعة البيانات 1339 تقييم مستخدم مأخوذة من متجر جوجل بلاي. أُجريت التجارب باستخدام ثلاثة سيناريوهات لنسب مشاركة مجموعة البيانات (7:3، 8:2، 9:1) وتم تقييمها باستخدام مصفوفة الخلط والتحقق المتقاطع *K-Fold Cross Validation*. تُظهر نتائج الاختبار أن سيناريو مشاركة البيانات بنسبة 9:1 ينتج أعلى دقة للنموذج بنسبة 81.34% مع دقة 81.47% واسترجاع 81.44% ودرجة 1 F بنسبة 81.34%. يُظهر تحليل الكلمات باستخدام TF-IDF أن كلمات مثل "جيد" و"لطيف" و"الطيب" تقيمن على فئة "راضٍ"، بينما كلمات مثل "السعر" و"غير قادر" و"التطبيق" هي الكلمات الثلاث الأكثر هيمنة في فئة "غير راضٍ". يمكن استنتاج أن طريقة "ساذج متعدد الحدود (*Multinomial Naive Bayes*)" تتمتع بأداء جيد لتصنيف مراجعات مستخدمي تطبيقات السفر عبر الإنترنت، وكلما زاد عدد مجموعات البيانات المستخدمة، كان النموذج الناتج أفضل

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Perkembangan teknologi yang pesat telah membawa perubahan dalam berbagai aspek kehidupan, termasuk dalam sektor pariwisata (Raja Suhud, 2023). 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 (CNN Indonesia, 2019).

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 (Tama *et al.*, 2019). 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 *et al.*, (2023) dengan menganalisis sentimen ulasan pengguna pada aplikasi Traveloka menggunakan metode *Naïve Bayes* dengan akurasi 82,91%. Yuyun *et al.*, (2021) dalam penelitian analisis sentimen pemerintah terhadap penanganan Covid-19 dengan metode *Multinomial Naïve Bayes*. Hasil *accuracy* model sebesar 74%. Penelitian lainnya yang dilakukan oleh Singh *et al.*, (2019) tentang perbandingan metode *Multinomial Naïve Bayes*. Hasil uji coba menghasilkan akurasi model sebesar 73.4%. Klasifikasi juga dibahas dalam Al-Qur'an Surah Al-Baqarah ayat 31:

وَعَلَّمَ آدَمَ الْأَسْمَاءَ كُلَّهَا ثُمَّ عَرَضَهُمْ عَلَى الْمَلَائِكَةِ فَقَالَ أَنْبِئُونِي بِأَسْمَاءِ هَؤُلَاءِ إِنْ كُنْتُمْ صَادِقِينَ

“Dan Dia ajarkan kepada Adam nama-nama (benda) semuanya, kemudian Dia perlihatkan kepada para malaikat seraya berfirman, “Sebutkan kepada-Ku nama semua (benda) ini, jika kamu yang benar!”(QS. Al-Baqarah :31)

Menurut tafsir yang disampaikan oleh Syekh Jalaludin dalam tafsir Jalalain mengatakan “(Dia mengajarkan Adam semua nama-nama [benda-benda]), Allah memberikan langsung pemahaman nama-nama benda ke dalam hati Adam AS (kemudian menampilkan semuanya) benda-benda itu. Di sini terdapat kemenangan makhluk berakal (di hadapan malaikat, lalu mengatakan) kepada mereka sebagai bentuk kemenangan hujah, (Sebutkanlah) Beritahukanlah (kepada-Ku nama-nama semua benda itu jika kamu memang benar orang-orang yang benar) bahwa Aku

tidak menciptakan makhluk yang lebih pandai dari kamu' atau 'bahwa kamu lebih berhak pada khilafah sebagai [pengganti-Ku]". Hal ini berhubungan dengan klasifikasi ulasan otomatis yang mampu mengenali kepuasan pengguna dengan cara mengelompokkan ulasan pengguna tersebut. Sehingga dapat diketahui kepuasan pengguna berdasarkan ulasannya, pada penelitian ini dikelompokkan menjadi kategori “*satisfied*” dan “*unhappy*”, pengembang aplikasi perjalanan dapat lebih cepat memahami apa yang membuat pengguna merasa puas atau tidak puas dengan layanan yang mereka terima, sehingga dapat meningkatkan layanannya untuk memenuhi kebutuhan pengguna dengan lebih baik.

## **1.2 Pernyataan Masalah**

Seberapa baik performa metode *Multinomial Naïve Bayes* dalam mengklasifikasikan ulasan pengguna aplikasi pemesanan tiket *travel online* untuk membantu pengembang memahami pengalaman pengguna?

## **1.3 Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk mengukur performa dari metode *Multinomial Naïve Bayes* dalam mengklasifikasikan ulasan pengguna aplikasi pemesanan tiket *travel online* untuk membantu pengembang memahami pengalaman pengguna.

## **1.4 Batasan Masalah**

- 1) Sumber data penelitian ini adalah data primer yang diperoleh dari teks komentar pengguna dari empat aplikasi pemesanan tiket *travel online* yaitu Traveloka, pegipegi, dan tiket.com, yang tersedia di *Platform Google Play*

*Store*. Label ditentukan melalui survei yang melibatkan responden umum dengan kategori “*satisfied*” dan “*unhappy*”.

- 2) Objek yang digunakan berupa ulasan pengguna pada *platform Google Play Store* dari aplikasi Traveloka, pegipegi, dan tiket.com dalam bahasa inggris sebanyak 1339 records data.

### **1.5 Manfaat Penelitian**

- 1) Penelitian ini berkontribusi pada literatur akademik mengenai klasifikasi ulasan, khususnya pada konteks klasifikasi ulasan aplikasi *travel online*
- 2) Membantu pelaku bisnis aplikasi *travel* untuk memahami kepuasan pengguna dan meningkatkan pelayanan bagi pengguna berdasarkan klasifikasi ulasan, sehingga pengguna menjadi lebih puas dengan pelayanan yang diberikan.

## BAB II

### STUDI PUSTAKA

#### 2.1 Kepuasan Pengguna

Aplikasi pemesanan tiket *travel online* atau biasa disebut *Online Travel Agent* (OTA) menyediakan kemudahan bagi penggunanya untuk memesan tiket transportasi seperti pesawat, bis, kereta api serta mengakomodasi perencanaan perjalanan dengan lebih efisien. Adanya aplikasi *travel*, pengguna dapat menjelajahi berbagai tempat di dalam negeri maupun mancanegara. Keberadaan aplikasi ini merubah paradigma perjalanan dengan menyederhanakan proses reservasi dan memberikan akses yang cepat bagi pengguna.

Dari banyak kemudahan akses dan penggunaan yang disediakan oleh aplikasi pemesanan tiket *travel online*, muncul permasalahan tentang memahami kepuasan pengguna secara menyeluruh. Ketersediaan fitur ulasan dan *rating* pada *platform Google Play Store* dan *App Store* tentang aplikasi *travel* dapat memberikan gambaran, tetapi analisis lebih mendalam sangat diperlukan untuk memahami kepuasan pengguna. Ini sangat berguna bagi pelaku bisnis aplikasi untuk mempertahankan pengguna mereka. Memberikan layanan pengguna yang baik adalah kunci keberhasilan dari aplikasi pemesanan tiket *travel online*. *Rating* aplikasi yang diberikan oleh pengguna biasanya menjadi acuan dalam menentukan kualitas pelayanan teknis maupun pelayan secara keseluruhan. Skala *rating* yang digunakan adalah bintang dari 1 sampai 5 (Reddy *et al.*, 2017).

Namun, mengandalkan *rating* sebagai acuan dalam menentukan kepuasan pelayanan bagi pengguna tidak cukup. Alternatif lain untuk mengetahui kepuasan

pengguna yaitu melalui klasifikasi ulasan yang pengguna berikan untuk aplikasi. Ulasan berisi bagaimana pengguna merasakan, mengevaluasi, dan merespon layanan yang diberikan oleh pihak aplikasi Tama *et al.*, (2019). Ulasan yang pengguna berikan merupakan sumber informasi penting bagi pelaku bisnis aplikasi dan pengembang untuk memahami pengalaman pengguna serta mengevaluasi kualitas layanan yang mereka terima.

Disamping itu, ulasan yang masuk sangat banyak sehingga sulit untuk dibaca dan dianalisis secara manual. Hal ini menyulitkan pengembang untuk mengkategorikan apakah pengguna puas atau tidak dengan pelayanan yang diberikan. Dengan demikian, perlu dilakukan klasifikasi ulasan untuk menentukan kepuasan pengguna aplikasi yang pada penelitian ini menggunakan metode *Multinomial Naïve Bayes* (Oktaviani *et al.*, 2021).

## 2.2 Klasifikasi Ulasan

Proses pemberian label atau kategorisasi pada teks berdasarkan kontennya biasa disebut dengan klasifikasi teks Rahayu & Andi, (2020). Dalam penelitian ini, klasifikasi teks diperlukan untuk mengkategorikan ulasan pengguna dari aplikasi pemesanan tiket *travel online*. Tujuan dari klasifikasi ini adalah untuk mengekstrak sentimen yang ada pada ulasan pengguna untuk memahami ulasan yang mereka berikan terhadap layanan aplikasi.

Penelitian sebelumnya oleh Diekson *et al.*, (2023) tentang klasifikasi ulasan pelanggan aplikasi Traveloka menggunakan metode SVM. Ulasan diklasifikasikan ke dalam dua kategori yaitu “*Satisfied*” dan “*unhappy*”. Dataset berupa *tweet* pengguna Traveloka yang diambil dari Twitter sebanyak 1200 *tweet*. Peneliti hanya

menggunakan 1 preprocessing yaitu *stopword* dan menggunakan TF-IDF sebagai ekstraksi fiturnya. Dari penelitian ini menghasilkan akurasi model SVM sebesar 84.4%.

Penelitian sebelumnya yang dilakukan oleh Wiratama & Rusli (2019) telah menunjukkan performa dari metode *Multinomial Naïve Bayes* dalam mengklasifikasikan *feedback user e-learning* menjadi dua kelas yaitu positif dan negatif. Tahapan *preprocessing* yang digunakan dalam penelitian ini meliputi *stemming, tokenization, stop word*, penghapusan URL, angka dan normalisasi kata-kata gaul sehari-hari. Dataset yang digunakan dalam penelitian ini berjumlah 627 entri, yang terdiri dari 269 *feedback* positif dan 358 *feedback* negatif. Dari model analisis sentimen yang diuji dan dievaluasi menghasilkan akurasi terbaik 71.6% dengan rasio pengujian 70:30.

Poernomo & Suharjito (2019) melakukan penelitian analisis sentimen menggunakan tiga metode yaitu KNN, *Naïve Bayes* dan *Support Vector Machine*. Tujuan penelitiannya untuk mengukur kepuasan pengguna aplikasi Traveloka, Tiket.com dan Agoda dengan menganalisis komentar Facebook dari halaman *Fanpage* mereka. *Preprocessing* yang digunakan yaitu *case folding, remove link*, dan *emoticon conversion*. Rasio pembagian data sebesar 7:3 menghasilkan akurasi model KNN sebesar 96.32%.

Abbas *et al.*, (2019) juga dalam penelitiannya melakukan analisis sentimen ulasan film menggunakan algoritma klasifikasi *Multinomial Naïve Bayes* menunjukkan kinerja yang bagus dalam mengkategorisasi. Pra-proses pada penelitian ini mencakup data *cleanup*, dan *word tokenization*, untuk ekstraksi fitur

menggunakan *Bag of Word* (BOW) dan *Term Frequency-Inverse Document Frequency* (TF-IDF). Pada penelitian ini, akurasi terbaik diperoleh setelah melakukan ekstraksi fitur TF-IDF dengan akurasi sebesar sekitar 90%.

Selanjutnya penelitian tentang analisis sentimen yang dilakukan Sharif *et al.*, (2019) pada ulasan restoran online menggunakan metode *Multinomial Naïve Bayes*. Menurut penelitian ini penting karena orang-orang banyak yang mengandalkan ulasan pelanggan sebelumnya untuk memilih restoran. Pada penelitian ini dataset berjumlah 1000 ulasan restoran yang dikumpulkan dari berbagai sumber seperti artikel dan grup facebook. bahasa Bangladesh kemudian diterjemahkan ke dalam bahasa inggris. Pra-proses yang dilakukan pada penelitian ini seperti *stopword*, *tokenization* dan normalisasi singkatan. Setelah itu dilakukan ekstraksi fitur dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk menghitung frekuensi kata pada ulasan dan seberapa umum kata tersebut dalam seluruh dataset. Setelah dilakukan *preprocessing* dilakukan pengujian dengan rasio 80:20. Hasil dari penelitian ini menunjukkan bahwa model yang diusulkan dapat melakukan klasifikasi ulasan restoran dengan akurasi 80.48%.

Yuyun *et al.*, (2021) dalam penelitiannya tentang klasifikasi sentimen publik terkait penanganan Covid-19 di Indonesia menggunakan metode *Multinomial Naïve Bayes*, menunjukkan performa yang baik. Data yang digunakan pada penelitian ini berasal dari Twitter API, dengan periode *crawling* data antara 25 Agustus – 24 November 2020. Pada penelitian ini sebanyak 2000 tweet sebagai dataset, *tweets* dikategorikan menjadi positif, negatif dan netral. Untuk tahapan *preprocessing* pada penelitian ini meliputi *case folding*, *tokenizing*, *filtering* dan

*stemming*. Setelah itu dilakukan TF-IDF untuk memberikan bobot pada setiap kata dalam dokumen. 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%.

Penelitian yang dilakukan oleh Oktaviani *et al.*, (2021) 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. *Preprocessing* yang digunakan *case folding*, menghapus angka, tanda baca dan menerjemahkan setiap ulasannya. Ekstraksi fitur yang digunakan adalah TF-IDF. Hasil yang diperoleh dari ulasan yang diberikan pengguna aplikasi Traveloka diperoleh akurasi model *Naïve Bayes* sebesar 91.20%.

Ferdino & Rusli (2019) dalam penelitiannya membahas klasifikasi ulasan pengguna *e-learning* Universitas Multimedia Nusantara. Menurut peneliti mengalisis ulasan pengguna yang sangat banyak dapat memakan waktu yang sangat lama oleh karena itu peneliti membuat sebuah sistem yang mampu melakukan klasifikasi ulasan agar mempermudah dalam melakukan klasifikasi ulasan. Metode yang digunakan pada penelitian ini adalah *Multinomial Naïve Bayes*. Tahap *preprocessing* yang digunakan menghapus angka, *case folding*, *stopword*, dan *Tokenization*, dan *stemming*. Kinerja pengklasifikasi diuji dan dievaluasi dengan hasil terbaik yaitu 92.5% untuk akurasi, 85.6% untuk presisi, 85.1% untuk *recall*, dan 85.4%.

Selanjutnya Ferdino & Rusli (2019) pada penelitiannya melakukan klasifikasi ulasan pengguna aplikasi Tokopedia dengan jumlah dataset sebanyak 5000 ulasan yang tersedia di Google Play Store. Ulasan dikategorikan ke dalam dua kategori yaitu positif dan negatif menggunakan metode Naive Bayes. Pembobotan kata yang digunakan adalah TF-IDF. Tahap *preprocessing* meliputi *case folding*, menghapus URL, tanda baca, *white space* dan *tokenizing*. Berdasarkan hasil pengujian dengan rasio pembagian data 8:2 menghasilkan akurasi model NB sebesar 83,9%, presisi 85.1%, recall 83.9%, dan F1-Score 83.8%.

Selanjutnya penelitian yang terakhir oleh Sari, (2019) dalam penelitiannya melakukan analisis sentimen ulasan pengguna aplikasi Grab dan Gojek menggunakan metode *Naïve Bayes*. Dataset diperoleh dari Twitter sebanyak 1000 data menggunakan kata kunci Gojek dan Grab. Pada penelitiannya bertujuan untuk mengekstrak opini, sentimen, evaluasi dan emosi pengguna mengenai layanan tertentu. Ulasan diklasifikasikan menjadi tiga kategori yaitu positif, negatif dan netral. Pembobotan kata yang digunakan adalah TF-IDF. *Preprocessing* yang digunakan meliputi *cleansing*, *filtering*, *tokenizing* dan *stemming*. Dari hasil evaluasi klasifikasi ulasan menggunakan algoritma *Naïve Bayes* berhasil mendapatkan akurasi model sebesar 72.33% dengan *recall* 73.95% dan presisi 73.24%.

Tabel 2.1 Tabel Penelitian Terdahulu

No	Sitasi	Objek	Metode	Hasil	Perbedaan
1	(Diekson et al., 2023)	Klasifikasi ulasan pengguna aplikasi Traveloka	Support Vector Model dengan TF-IDF	Akurasi model SVM sebesar 84.4%	- Metode yang digunakan adalah SVM - Objek yang diteliti fokus pada aplikasi Traveloka

No	Sitasi	Objek	Metode	Hasil	Perbedaan
2	(Wiratama & Rusli, 2019)	Analisis sentimen pengguna <i>website e-learning</i>	Multinomial Naïve Bayes	Menghasilkan akurasi model MNB terbaik dari pembagian data 7:3 sebesar 71.6%	- Penelitian menganalisis sentimen pengguna <i>website e-learning</i>
3	(Poernomo & Suharjito, 2019)	Analisis sentimen komentar pengguna aplikasi Traveloka, Tiket.com, dan Agoda	K-Nearest Neighbors, Naïve Bayes & Support Vector Machine	Metode KNN menghasilkan akurasi terbaik dibandingkan dengan metode SVM & NB dengan akurasi model 96.32%	- Metode yang digunakan adalah KNN, NB dan SVM - Objek penelitian komentar pengguna aplikasi Traveloka, Tiket.com, dan Agoda
4	(Abbas et al., 2019)	Analisis sentimen ulasan film	Multinomial Naïve Bayes dengan pembobotan kata TF-IDF	Akurasi model yang diperoleh sebesar 90%	- Objek penelitian analisis sentimen ulasan film
5	(Sharif et al., 2019)	Analisis sentimen pada ulasan pelanggan restoran online	Multinomial Naïve Bayes dengan TF-IDF	Menghasilkan akurasi model MNB terbaik dari pembagian data 8:2 sebesar 80.48% %	- Objek penelitian analisis sentimen pada ulasan pelanggan restoran online
6	(Yuyun et al., 2021)	Klasifikasi sentimen masyarakat terhadap penanganan <i>Covid-19</i>	Multinomial Naïve Bayes dengan TF-IDF	Menghasilkan akurasi model sebesar 74%, presisi 74%, dan recall 74%	- Objek penelitian klasifikasi sentimen masyarakat terhadap penanganan <i>Covid-19</i>
7	(Oktaviani et al., 2021)	Analisis sentimen ulasan pengguna aplikasi e-commerce Traveloka	Naïve Bayes dengan TF-IDF	Hasil yang diperoleh dari ulasan yang diberikan pengguna aplikasi Traveloka diperoleh akurasi model Naïve Bayes sebesar 91.20%	- Objek penelitian analisis sentimen ulasan pengguna aplikasi e-commerce Traveloka - Metode yang digunakan Naïve Bayes
8	(Ferdino & Rusli, 2019)	Klasifikasi ulasan pengguna <i>e-learning</i> Universitas Multimedia Nusantara.	Multinomial Naïve Bayes	Kinerja pengklasifikasi diuji dan dievaluasi dengan hasil terbaik yaitu 92.5% untuk akurasi, 85.6%, presisi, 85.1%, recall, dan 85.4%.	- Objek penelitian klasifikasi ulasan pengguna <i>e-learning</i> Universitas Multimedia Nusantara
9	(Ferdino & Rusli, 2019)	Klasifikasi ulasan pengguna	Naïve Bayes dengan TF-IDF	Berdasarkan hasil pengujian dengan rasio pembagian data	- Objek penelitian klasifikasi ulasan pengguna aplikasi

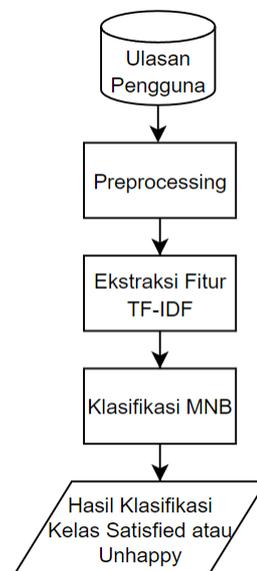
No	Sitasi	Objek	Metode	Hasil	Perbedaan
		aplikasi Tokopedia		8:2 menghasilkan akurasi model NB sebesar 83,9%, presisi 85.1%, recall 83.9%, dan F1-Score 83.8%.	e-commerce Tokopedia - Metode yang digunakan Naïve Bayes
10	(Sari, 2019)	Analisis sentimen ulasan pengguna aplikasi Grab dan Gojek	Naïve Bayes dengan TF-IDF	Dari hasil evaluasi klasifikasi ulasan menggunakan algoritma Naïve Bayes berhasil mendapatkan akurasi model sebesar 72.33% dengan recall 73.95% dan presisi 73.24%	- Objek penelitian analisis sentimen ulasan pengguna aplikasi Grab dan Gojeks - Metode yang digunakan Naïve Bayes

## BAB III

### DESAIN DAN IMPLEMENTASI

#### 3.1 Desain Sistem

Desain sistem menggambarkan alur dari sistem klasifikasi secara menyeluruh. Tahapan penelitian ini dimulai dari input data ulasan pengguna. Selanjutnya data akan masuk tahap *preprocessing* untuk membersihkan data sehingga data siap untuk diproses. Tahap selanjutnya adalah ekstraksi fitur menggunakan algoritma TF-IDF. Setelah dilakukan ekstraksi fitur, data masuk ke proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*. Hasil klasifikasi memberikan informasi tentang apakah ulasan tersebut masuk ke dalam kelas “*satisfied*” atau “*unhappy*”. Dari uraian diatas, berikut alur sistem yang dapat dilihat pada Gambar 3.1.



Gambar 3.1 Desain Sistem

### 3.2 Akuisisi Data

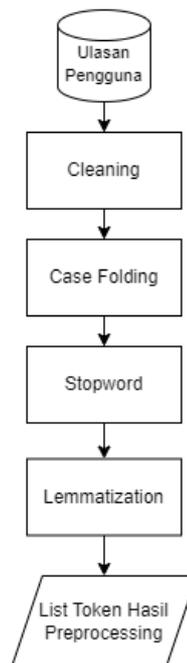
Sumber data yang digunakan berasal dari *platform Google Play Store*. Untuk melakukan pengumpulan data pada *Google Play Store* menggunakan library *google play scraper*. Library ini menyediakan API untuk melakukan *scraping* pada *Google Play Store* dengan mudah (JoMingyu, 2023). Data Ulasan yang diambil dalam bahasa Inggris, berasal dari 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 3.1 Tabel Data

No	Aplikasi	Jumlah Data
1	Traveloka	499
2	tiket.com	432
3	pegipegi	408
<b>Total</b>		<b>1339</b>

### 3.3 Preprocessing

Preprocessing bertujuan untuk membersihkan data agar mudah dipahami dan dianalisis oleh berbagai algoritma machine learning, sehingga siap untuk diproses (Tabassum & Patil, 2020). Data akan melalui empat tahap preprocessing yaitu cleaning, case folding, stopword, dan lemmatization. Berikut tahapan preprocessing pada penelitian ini dapat dilihat pada Gambar 3.2:



Gambar 3. 2 Diagram alur Tahapan Preprocessing

### 3.3.1 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.

### 3.3.2 Case Folding

Tahap kedua dari *preprocessing* adalah *case folding*. Tahapan ini bertujuan untuk mengubah huruf kapital menjadi huruf kecil (Hasan et al., 2020). Hal ini dilakukan untuk membuat semua huruf menjadi seragam dan juga mengurangi beban komputasi.

### 3.3.3 Stopword

Stopword adalah proses menghapus kata pada kalimat yang tidak memiliki makna kontekstual seperti “a”, “an”, “the”, “was”, “in” dll (Ramachandran & Parvathi, 2019). Tahap ini membantu dalam memfokuskan analisis pada kata-kata yang benar-benar relevan.

### 3.3.4 Lemmatization

Tahap *preprocessing* terakhir pada penelitian ini adalah *lemmatization*. *Lemmatization* adalah proses mengubah kata menjadi kata dasar dengan mengetahui konteks dari kata tersebut (Tabassum & Patil, 2020). Berbeda dengan *stemming*, *lemmatization* lebih akurat karena mempertimbangkan konteks dari makna kata tersebut.

## 3.5 Ekstraksi Fitur 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. Apabila sebuah kata tersebut jarang muncul di seluruh ulasan, maka kata tersebut dianggap penting dari seluruh ulasan.

Langkah pertama dalam menghitung besaran nilai TF dan IDF dari setiap kata menggunakan Persamaan 3.1 dan 3.2 berikut :

$$TF_{ij} = \frac{f_{ij}}{\sum_k n_{i,j}} \quad (3.1)$$

Untuk  $f_{ij}$  adalah jumlah kata ( $i$ ) dalam pada ulasan ( $j$ ), kemudian untuk  $\sum_k n_{i,j}$  adalah total semua kata yang ada dalam ulasan ( $j$ ).

$$IDF_i = \log\left(\frac{N}{n_i}\right) \quad (3.2)$$

$N$  merupakan jumlah total dari semua ulasan yang digunakan,  $n_i$  adalah total ulasan yang memiliki  $term_i$  atau kata tersebut. Setelah melakukan perhitungan mulai dari TF kemudian IDF, maka TF-IDF dihitung dengan Persamaan 3.3 sebagai berikut :

$$TFIDF = TF_{ij} \times IDF_i \quad (3.3)$$

### 3.6 Multinomial Naïve Bayes

Algoritma *Multinomial Naïve Bayes* merupakan salah satu algoritma pembelajaran probabilistik yang digunakan pada *Natural Language Processing* (NLP) (Sriram, 2022). 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 (Yuyun *et al.*, 2021). Algoritma ini memanfaatkan konsep probabilistik yang didasarkan pada *Teorema Bayes*. Berikut Persamaan 3.4 merupakan aturan *Theorema bayes*.

$$P(c_i|d_j) = \frac{P(d_j|c_i)P(c_i)}{P(d_j)} \quad (3.4)$$

Pada klasifikasi ulasan,  $P(d_j)$  dapat dianggap sebagai kumpulan kata  $\{w_1, w_2, w_3, \dots, w_n\}$  dimana frekuensi kemunculan kata tersebut diasumsikan sebagai variabel acak dengan distribusi probabilitas tertentu. Pada penelitian ini  $P(d_j)$  merepresentasikan probabilitas ulasan pengguna,  $P(c_i)$  merepresentasikan probabilitas dari kelas “*satisfied*” dan “*unhappy*”. Model klasifikasi MNB dapat dirumuskan pada Persamaan 3.5 berikut ini.

$$P_{(c|w)} = P(c) \prod_{w=1}^n P(w_k|c_i) \quad (3.5)$$

$P(w_k|c_i)$  merupakan probabilitas kata  $w$  ke  $k$  pada kelas  $c$  ke  $i$ . Notasi  $w$  adalah  $\{w_1, w_2, w_3, \dots, w_n\}$  dimana  $w$  merupakan kata dari ulasan  $d$ . Klasifikasi ulasan ini bertujuan untuk mencari nilai *maximum a posteriori* (MAP) kelas  $C_{MAP}$  pada Persamaan 3.6.

$$C_{MAP} = \operatorname{argmax}_{c_i \in C} P(c_i) \prod_{w=1}^n P(w_k|c_i) \quad (3.6)$$

Dimana  $P(c_i)$  adalah probabilitas *prior* yang mewakili kelas “*satisfied*” dan “*unhappy*”. Berikut persamaan dari *prior*.

$$P(c_i) = \frac{d_k}{D} \quad (3.7)$$

Notasi  $d_k$  adalah jumlah ulasan yang masuk dalam kelas “*satisfied*” atau “*unhappy*”, dibagi dengan  $D$  yang merupakan jumlah total ulasan dalam seluruh dataset training.

$P(w_i/c_j)$  merupakan probabilitas *likelihood* atau probabilitas suatu kata pada kelas  $c_j$ . Berikut persamaan dari *likelihood*.

$$P(w_k|c_i) = \frac{W_k + 1}{\sum W_k + vocabulary} \quad (3.11)$$

Notasi  $W_i$  merepresentasikan nilai TF-IDF dari kata  $W$  ke  $k$  dalam suatu kelas  $c_j$ . Notasi  $W$  adalah jumlah nilai TF-IDF dari semua kata yang di kelas  $c$  ke  $i$  sementara *vocabulary* merupakan jumlah bobot IDF semua kata yang ada pada dataset *training*.

## **BAB IV**

### **UJI COBA DAN PEMBAHASAN**

#### **4.1 Langkah Pengujian**

##### **4.1.1 Pelabelan Data**

Proses pelabelan dataset ulasan pengguna aplikasi tiket *travel online* dilakukan menggunakan metode *crowdsourcing*. Ulasan dari aplikasi Traveloka, pegipegi dan Tiket.com dikumpulkan dari *Google Play Store* dan disusun dalam sebuah *form* yang kemudian disebarakan kepada responden. 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. Hasil dari tanggapan dari responden dikumpulkan dan disusun dalam format Excel. Ulasan yang sudah terlabel kemudian digunakan untuk analisis lebih lanjut dengan *preprocessing*, ekstraksi fitur TF-IDF dan klasifikasi menggunakan MNB.

##### **4.1.2 Input Dataset**

Penelitian ini menggunakan data ulasan pengguna dari tiga aplikasi pemesanan tiket travel online yaitu Traveloka, pegipegi, dan Tiket.com. Data ulasan ini diperoleh dari platform Google Play Store menggunakan teknik *scraping*. Jumlah data yang diperoleh sebanyak 1339 ulasan pengguna. Proses pelabelan dilakukan dengan menggunakan metode *crowdsourcing* berupa menyebarkan *form* kepada responden untuk memberikan label. Setiap ulasan diberi label “*satisfied*” atau “*unhappy*”. Berikut contoh sampel dataset dapat dilihat pada Tabel 4.1 berikut.

Tabel 4.1 Sampel Dataset

No	Ulasan	Kelas
1	I cannot place an order after updating my phone to Android 12. After each payment, somehow it always leads to a blank screen. Also the CS agent isn't helpful.	Unhappy
2	good, fast respon if any trouble, good solution	Satisfied
3	no verification code when purchase with credit card	Unhappy
4	your website needs attention as it gets stuck at the payment page	Unhappy
5	Traveloka is very useful and user friendly, we can get best flight or deals	Satisfied
6	I have been using this for years, now it's even easier. Of course there's room for improvement.	Satisfied
7	Easy to use, but some copies are too small to notice	Satisfied
8	Latest version is so buggy. Can't open a single page of any hotels. Please fix it asap.	Unhappy
9	This applications quite good and helpful	Satisfied
10	App won't verify users ID when booking flight ticket.	Unhappy

Setelah melalui tahap pelabelan, dataset akan masuk ke tahap *preprocessing* yang bertujuan untuk membersihkan data dari *noise* agar siap masuk ke proses utama. Tahap *preprocessing* yang digunakan mencakup *cleaning*, *case folding*, *stopword* dan *lemmatization*.

#### 4.1.3 Pembagian dataset Menjadi 2 Subset

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 (Baiq Nurul Azmi *et al.*, 2023) dataset akan dibagi menjadi tiga rasio, yaitu 7:3, 8:2, dan 9:1. Berikut Tabel 4.2 yang berisi jumlah detail untuk setiap rasio pembagian data.

Tabel 4.2 Pembagian Dataset

Pengujian	Rasio	Jumlah Data	
		<i>Train</i>	<i>Test</i>
1	7:3	937	402
2	8:2	1071	268
3	9:1	1205	134

#### 4.1.4 Pelatihan Model dengan Subset Data Latih

Pada tahap ini, model klasifikasi akan dilatih menggunakan subset *train* dari dataset. Pelatihan model dimulai dengan *preprocessing* data ulasan untuk dibersihkan. Setelah tahap *preprocessing* selesai, dilakukan ekstraksi fitur menggunakan TF-IDF untuk mengukur kepentingan kata dalam setiap ulasan. Selanjutnya, data ulasan yang telah diekstraksi fitur lalu dilatih menggunakan metode *Multinomial Naïve Bayes*, data akan mempelajari pola kata dalam setiap kelas (*satisfied* dan *unhappy*) dengan menghitung probabilitas probabilitas prior berdasarkan distribusi kelas dalam data latih dan probabilitas likelihood berdasarkan nilai TF-IDF dari setiap kata. Model akan belajar pola-pola tertentu dari data *train* sehingga dapat melakukan prediksi dengan akurat pada data yang belum pernah dilihat sebelumnya.

#### 4.1.5 Pelatihan Model dengan Subset Data Uji

Setelah model dilatih menggunakan subset *train*, tahap selanjutnya adalah menguji performa model menggunakan subset *test*. Data *test* yang telah melalui tahapan preprocessing dan ekstraksi fitur TF-IDF akan diarahkan ke model yang sudah dilatih. Selanjutnya model akan memprediksi kelas untuk setiap ulasan dalam data *train*. Hasil dari tahap ini akan digunakan untuk mengevaluasi performa model sebelum diterapkan pada data real.

#### 4.1.6 Evaluasi Performa Model dengan Confuion Matrix

Selanjutnya mengukur performa dari metode *Multinomial Naïve Bayes* menggunakan *confusion matrix*. Metode *confusion matrix* adalah tabel yang digunakan untuk mengevaluasi performa dan memvisualisasikan kinerja dari model klasifikasi. Matriks ini menunjukkan perbandingan antara prediksi model dan nilai aktual dari data. Berikut contoh tabel dari *confusion matrix*.

Tabel 4.3 Tabel *Confusion Matrix*

<i>Actual Class</i>	<i>Predicted Class</i>	
	<i>Satisfied</i>	<i>Unhappy</i>
<i>Satisfied</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
<i>Unhappy</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Persamaan 4.1 untuk menghitung nilai *accuracy*, Persamaan 4.2 untuk menghitung *precision*, Persamaan 4.3 untuk menghitung nilai *recall* dan Persamaan 4.4 untuk menghitung *F1-Score*.

$$Accuracy (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (4.1)$$

$$Precision (\%) = \frac{TP}{TP + FP} \times 100\% \quad (4.2)$$

$$Recall (\%) = \frac{TP}{TP + FN} \times 100\% \quad (4.3)$$

$$F1 - Score (\%) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \quad (4.4)$$

Keterangan:

TP (*True Positive*) = Jumlah ulasan dengan hasil klasifikasi kelas *i*, diprediksi benar kelas *i* oleh sistem.

TN (*True Negative*) = Jumlah ulasan dengan hasil klasifikasi bukan kelas *i*, diprediksi benar bukan kelas *i* oleh sistem.

FP (*False Positive*) = Jumlah ulasan dengan hasil klasifikasi bukan kelas *i*, namun diprediksi masuk kelas *i* oleh sistem.

FN (*False Negative*) = Jumlah Ulasan dengan hasil klasifikasi kelas *i*, namun diprediksi bukan kelas *i* oleh sistem.

*Accuracy* adalah nilai kedekatan antara nilai *predicted* dengan nilai actual. *Precision* adalah nilai dari semua kelas yang diprediksi positif, berapa banyak yang benar-benar positif. *Recall* merupakan nilai dari seluruh kelas positif, berapa banyak diprediksi benar, dan terakhir *F1-Score* adalah evaluasi hasil kombinasi *Recall* dan *Precision*.

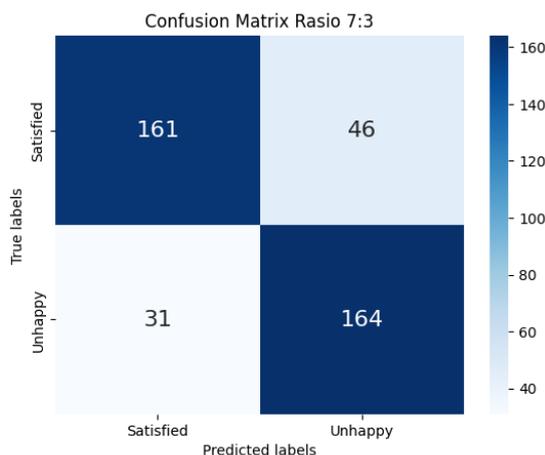
## **4.2 Hasil dan Uji Coba**

Sesuai dengan skenario pengujian, penelitian ini dilakukan 3 kali uji coba dengan membagi dataset menjadi beberapa rasio. Tujuan dari skenario ini adalah untuk menemukan rasio data *train* dan data *test* yang memberikan hasil paling optimal untuk model klasifikasi. Masing-masing model klasifikasi yang dihasilkan dievaluasi performanya dengan menggunakan parameter *accuracy*, *precision*, *recall*, dan *F1-score*. Untuk mendapatkan nilai-nilai tersebut, kelas aktual dibandingkan dengan kelas prediksi sehingga dapat terlihat berapa data pertanyaan yang diprediksi dengan benar dan tidak.

### **4.2.1 Hasil Uji Coba Skenario-1**

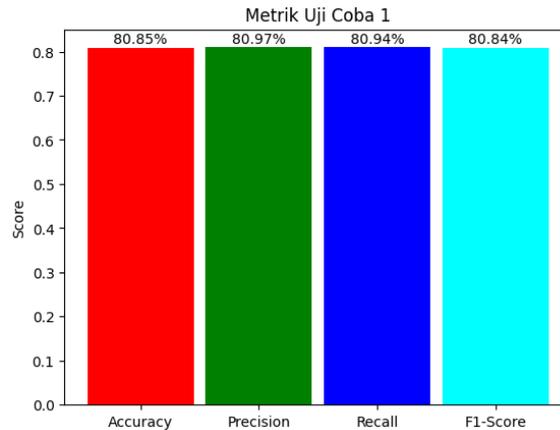
Pada uji coba skenario 1 ini dilakukan rasio pembagian dataset 7:3 dari total keseluruhan yaitu sebanyak 1339 data. Rasio pembagian data ini terdiri dari sebanyak 937 data *train* dan 402 data *test* yang digunakan. Distribusi kelas “*satisfied*” pada data *train* sebesar 51% dari jumlah total data *train*, sedangkan data dengan label “*unhappy*” sebesar 49% dari jumlah total data *train*. Begitu juga pada data *test* label “*satisfied*” berjumlah 207, dan untuk label “*unhappy*” berjumlah 195.

Berikut Gambar 4.1 merupakan hasil evaluasi pengukuran performa model menggunakan confusion matrix pada uji coba skenario 1.



Gambar 4.1 *Confusion Matrix* Uji Coba Skenario-1

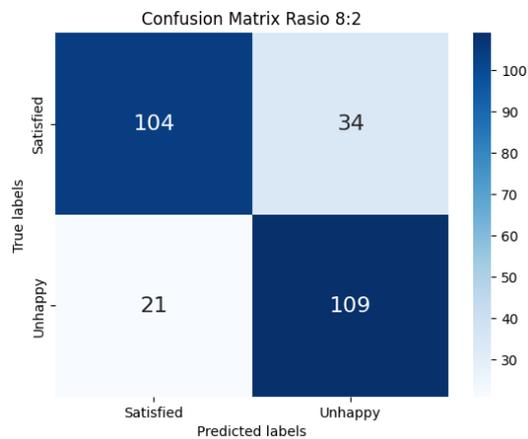
Pada Gambar 4.1 diatas menunjukkan hasil pengukuran dari evaluasi performa model menggunakan *confusion matrix* pada proses klasifikasi uji coba skenario 1 penelitian ini. Dari hasil klasifikasi, dapat diketahui bahwa model berhasil mengklasifikasikan 165 data ulasan sebagai kelas “*satisfied*” dengan tepat (*True Positive*), dan 162 data ulasan sebagai kelas “*unhappy*” dengan tepat (*True Negative*). Namun, terdapat 42 data yang seharusnya diklasifikasikan sebagai kelas “*satisfied*” namun diprediksi sebagai kelas “*unhappy*” oleh model (*False Negative*), serta 33 data yang seharusnya diklasifikasikan sebagai kelas “*unhappy*” namun diprediksi sebagai kelas “*satisfied*” oleh model (*False Positive*). Berdasarkan hasil confusion matrix diatas maka diperoleh hasil perhitungan performa model menggunakan Persamaan 4.1, 4.2, 4.3, 4.4 dapat dilihat pada Gambar 4.2 sebagai berikut.



Gambar 4.2 Metrik Performa Uji Coba Skenario-1

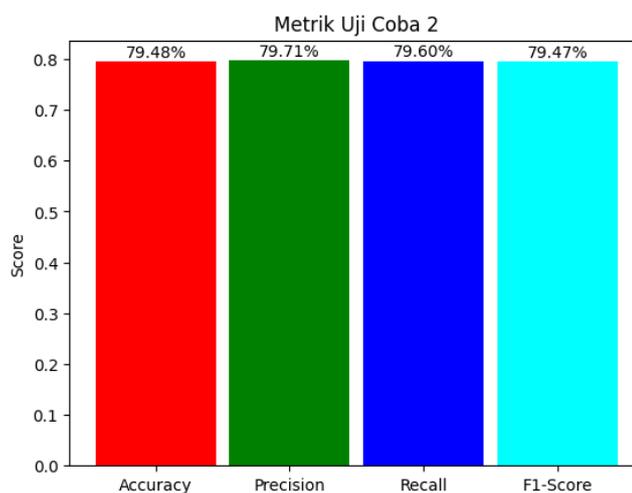
#### 4.2.2 Hasil Uji Coba Skenario-2

Pada uji coba skenario 2 ini dilakukan rasio pembagian dataset 8:2 terdiri dari sebanyak 1071 data train dan 268 data test yang digunakan. Distribusi kelas dengan label “*satisfied*” pada data *train* sebesar 51% dari jumlah total data *train*, sedangkan distribusi kelas dengan label “*unhappy*” sebesar 49% dari jumlah total data *train*. Pada data *test* label “*satisfied*” berjumlah 138, dan untuk label “*unhappy*” berjumlah 130. Berikut Gambar 4.3 merupakan hasil evaluasi pengukuran performa menggunakan confusion matrix pada uji skenario 2.



Gambar 4.3 Confusion Matrix Uji Coba Skenario-2

Pada Gambar 4.3 diatas menunjukkan hasil pengukuran dari evaluasi performa menggunakan confusion matrix pada proses klasifikasi uji skenario-1 penelitian ini. Dari hasil klasifikasi,dapat diketahui bahwa model berhasil mengklasifikasikan 104 data ulasan sebagai kelas “*satisfied*” dengan tepat (*True Positive*), dan 109 data ulasan sebagai kelas “*unhappy*” dengan tepat (*True Negative*). Namun, terdapat 34 data yang seharusnya diklasifikasikan sebagai kelas "*satisfied*" namun diprediksi sebagai kelas "*unhappy*" oleh model (*False Negative*), serta 21 data yang seharusnya diklasifikasikan sebagai kelas "*unhappy*" namun diprediksi sebagai kelas "*satisfied*" oleh model (*False Positive*). Berdasarkan hasil *confusion matrix* diatas maka diperoleh hasil perhitungan performa model menggunakan Persamaan 4.1, 4.2, 4.3, 4.4 dapat dilihat pada Gambar 4.4 sebagai berikut.

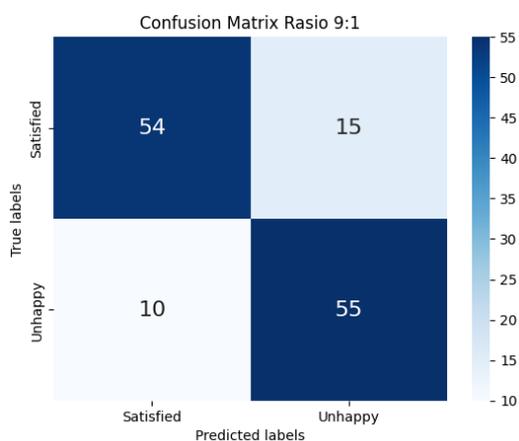


Gambar 4.4 Metrik Performa Uji Coba 1

### 4.2.3 Hasil Uji Coba Skenario-3

Pada uji coba skenario 3 ini dilakukan rasio pembagian dataset 9:1 dari total keseluruhan yaitu sebanyak 1339 data. Rasio pembagian data ini terdiri dari

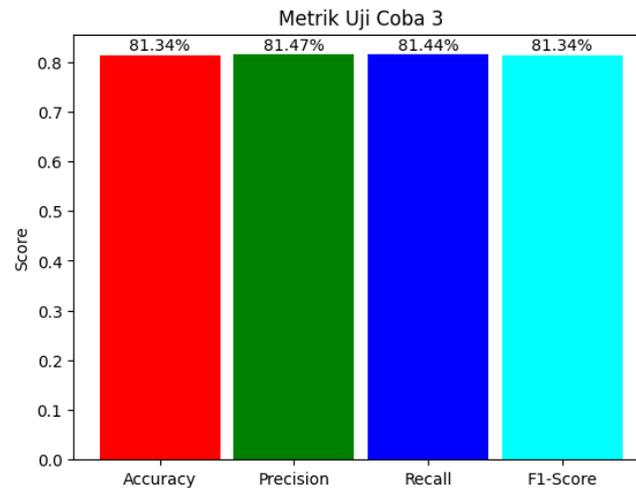
sebanyak 1205 data *train* dan 134 data *test* yang digunakan. Pada data *train*, data dengan label “*satisfied*” berjumlah 619 data, sedangkan data dengan label “*unhappy*” berjumlah 586 data. Pada data test label “*satisfied*” berjumlah 69 data, dan untuk label “*unhappy*” berjumlah 65 data. Berikut Gambar 4.5 merupakan hasil evaluasi pengukuran performa menggunakan *confusion matrix* pada uji skenario 3.



Gambar 4.5 *Confusion Matrix* Uji Coba 3

Pada Gambar 4.5 diatas menunjukkan hasil pengukuran dari evaluasi performa menggunakan *confusion matrix* pada proses klasifikasi uji coba skenario 1 penelitian ini. Dari hasil klasifikasi, dapat diketahui bahwa model berhasil mengklasifikasikan 54 data ulasan sebagai kelas “*satisfied*” dengan tepat (*True Positive*), dan 55 data ulasan sebagai kelas “*unhappy*” dengan tepat (*True Negative*). Namun, terdapat 15 data yang seharusnya diklasifikasikan sebagai kelas "satisfied" namun diprediksi sebagai kelas "unhappy" oleh model (*False Negative*), serta 10 data yang seharusnya diklasifikasikan sebagai kelas "unhappy" namun diprediksi sebagai kelas "satisfied" oleh model (*False Positive*). Berdasarkan hasil *confusion matrix* diatas maka diperoleh hasil perhitungan performa model

menggunakan Persamaan 4.1, 4.2, 4.3, 4.4 dapat dilihat pada Gambar 4.6 sebagai berikut.



Gambar 4.6 Metrik Uji Coba 3

### 4.3 Pembahasan

Berdasarkan uji coba yang telah dilakukan, diperoleh hasil performa dari tiga model klasifikasi dapat dilihat pada Tabel 4.4. Tabel tersebut berisi nilai rata-rata dari semua uji coba yang telah dievaluasi menggunakan confusion matrix menggunakan 4 parameter yaitu *accuracy*, *precision*, *recall*, dan *F1-Score*. Parameter yang digunakan ini bertujuan untuk mengukur seberapa baik model mengklasifikasikan data ulasan pada setiap uji coba.

Tabel 4.4 Hasil Uji Coba

Uji Coba	Rasio Pembagian Data	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
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	<b>81.34%</b>	<b>81.47%</b>	<b>81.43%</b>	<b>81.34%</b>

Pada Tabel 4.4 dapat diketahui hasil uji coba yang didapatkan berdasarkan evaluasi performa menggunakan *confusion matrix* pada tiap skenario pembagian

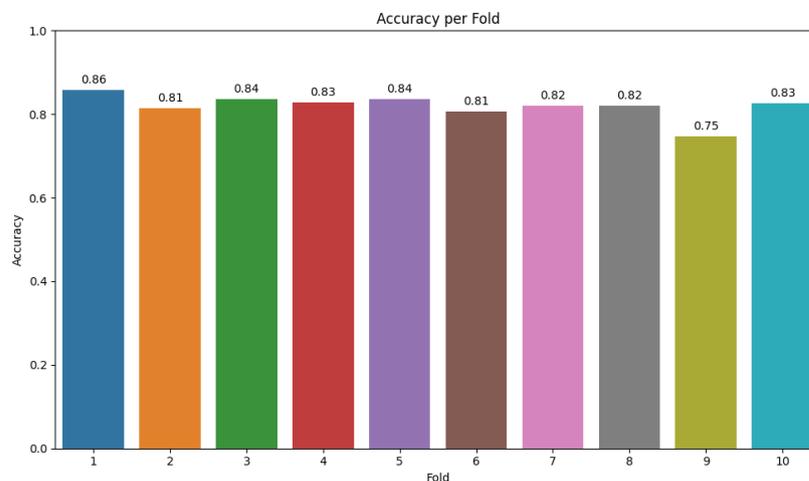
data yang telah ditentukan. Didapatkan skenario uji coba pada pembagian data 9:1 mendapatkan akurasi tertinggi yaitu sebesar 81.34%, ini menunjukkan bahwa model mampu dengan baik dalam melakukan klasifikasi model. Akurasi terbaik selanjutnya diikuti oleh skenario 7:3 dengan akurasi sebesar 80.84%, kemudian skenario diikuti dengan skenario pembagian data 8:2 dengan akurasi sebesar 79.48%.

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.

Perbandingan data *train* dan data *test* memiliki pengaruh yang cukup besar terhadap performa model. Namun, tidak hanya itu, keragaman data juga berpengaruh pada performa model. Jika dataset memiliki kecenderungan pola yang hampir sama atau berfokus pada karakteristik tertentu, maka model akan kesulitan menghadapi data baru. Begitu juga sebaliknya, jika data memiliki variasi yang luas, maka akan semakin mudah model untuk melakukan prediksi dengan benar. Karena

data yang diambil secara *random*, hal ini menyebabkan seberapa bagus kualitas dari dataset yang diperoleh. Oleh karena itu, pada penelitian ini menerapkan *K-Fold Cross Validation* untuk memastikan model diuji dengan berbagai macam kombinasi data *train* dan data *test*.

Penelitian ini menerapkan *K-Fold Cross Validation* untuk mengevaluasi performa model dengan cara membagi dataset menjadi K subset yang sama ukurannya. Hal ini dilakukan untuk melihat apakah model tidak mengalami penurunan performa pada input baru dari distribusi data yang sama. Dataset dibagi menjadi K bagian. model dilatih dan dievaluasi sebanyak K kali menggunakan *fold* yang berbeda setiap iterasi. Kemudian metrik performa dihitung pada setiap iterasi untuk mengetahui rata-rata hasil evaluasi performa model. Penelitian ini menggunakan skenario *K-Fold Cross Validation* yaitu K=10.



Gambar 4.7 Hasil *K-Fold Cross Validation*

Pada Gambar 4.7 menunjukkan hasil *K-Fold Cross Validation* dengan k=10 mendapatkan nilai rata-rata akurasi sebesar 82%. Nilai akurasi terendah didapatkan

pada *fold* ke-9 dengan nilai 75% dan nilai tertinggi didapatkan pada *fold* ke-1 dengan nilai 86%.

Ekstraksi fitur pada penelitian ini menggunakan TF-IDF untuk menganalisis kata. Nilai dari TF-IDF dapat mengidentifikasi kata-kata yang sering muncul serta dapat merepresentasikan sebuah kelas ulasan. Analisis kata digunakan untuk mengidentifikasi ulasan pengguna aplikasi tiket *travel online* apakah termasuk ke dalam kelas “*satisfied*” atau “*unhappy*”.

Tabel 4.5 *Ranking* Nilai TF-IDF Kelas *Satisfied*

No	Kata	Nilai TF-IDF
1	good	0.089195
2	easy	0.064696
3	nice	0.052335
4	app	0.049246
5	simple	0.042454
6	use	0.040678
7	price	0.039662
8	apps	0.037525
9	fast	0.036542
10	helpful	0.036385

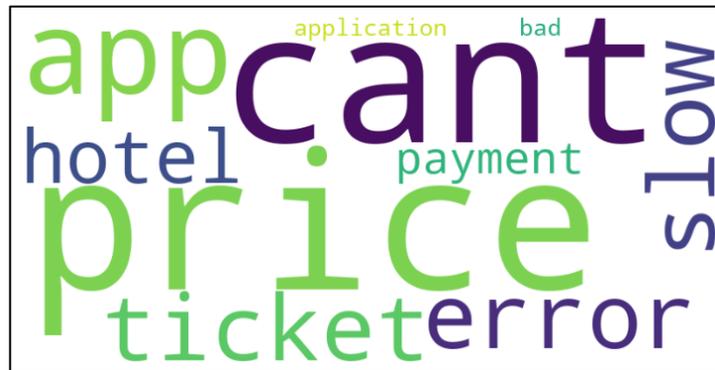
Tabel 4.5 dapat dilihat kata-kata yang mendominasi pada kelas “*satisfied*”. Kata-kata seperti “*good*”, “*easy*”, dan “*nice*” merupakan 3 kata yang paling mendominasi pada kelas “*satisfied*”. Hal ini menunjukkan bahwa pengguna yang memberikan ulasan puas atau “*satisfied*” cenderung menggunakan kata-kata tersebut untuk menggambarkan pengalaman mereka menggunakan aplikasi *travel online*, dengan demikian, pengembang atau pemilik aplikasi dapat menggunakan informasi ini untuk memperbaiki atau mempertahankan fitur-fitur yang dinilai positif oleh pengguna.

Gambar 4. 8 Word Cloud kelas *Satisfied*Tabel 4. 6 *Ranking* Nilai TF-IDF Kelas

No	Kata	Nilai TF-IDF
1	price	0.030453
2	cant	0.028797
3	app	0.027105
4	ticket	0.026971
5	error	0.026372
6	slow	0.022543
7	hotel	0.022321
8	payment	0.022097
9	application	0.021855
10	bad	0.019223

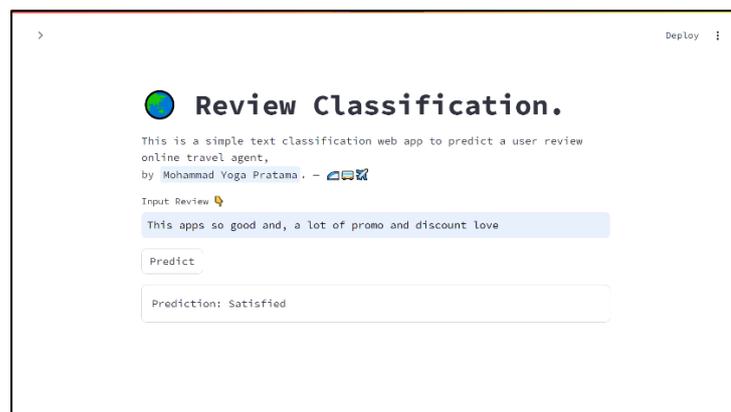
Tabel 4.6 dapat dilihat kata-kata yang mendominasi pada kelas “*satisfied*”. Kata-kata seperti “*price*”, “*cant*”, dan “*app*” merupakan 3 kata yang paling mendominasi pada kelas “*unhappy*”. Hal ini menunjukkan adanya ketidakpuasan terhadap harga produk atau layanan, kendala atau masalah dalam penggunaan aplikasi, serta kualitas atau performa dari aplikasi yang kurang memuaskan. Dari analisis kata ini memberikan gambaran yang lebih jelas terhadap aspek-aspek yang perlu diperbaiki oleh pemilik atau pengembang aplikasi untuk meningkatkan kepuasan pengguna, seperti meninjau kembali strategi harga, memperbaiki *bug* atau masalah teknis, serta meningkatkan kualitas dan fungsionalitas aplikasi secara keseluruhan. Dengan demikian, dari analisis ini dapat menjadi landasan untuk

melakukan perbaikan yang lebih baik di masa yang akan datang guna untuk meningkatkan kepuasan pengguna dan keberhasilan aplikasi tiket *travel online*.



Gambar 4. 9 Wordcloud Kelas *Unhappy*

Pada penelitian membuat implementasi sederhana dari model dengan akurasi terbaik yaitu skenario-3. Model diimplementasikan dalam bentuk *web app* sederhana yang dimana user dapat memasukkan data ulasan baru lalu diprediksi apakah ulasan tersebut masuk ke dalam kelas “*satsified*” atau “*unhappy*” yang dapat dilihat pada Gambar 4.10 berikut.



Gambar 4.10 Implementasi Model

Berdasarkan hasil penelitian yang telah dilakukan, model mampu digunakan untuk melakukan klasifikasi terhadap ulasan pengguna pada aplikasi *travel online*.

Model ini dapat bermanfaat sebagai informasi terkait kepuasan pengguna bagi pemilik atau pengembang aplikasi untuk ditindaklanjuti jika terdapat keterkaitan masalah pada aplikasi. Dari model ini diharapkan dapat meningkatkan kualitas layanan berdasarkan klasifikasi ulasan pengguna, sehingga pengguna menjadi lebih puas. Dalam hal ini berkaitan dengan surah Az-Zumar Ayat 18 yang berbunyi :

الَّذِينَ يَسْتَمِعُونَ الْقَوْلَ فَيَتَّبِعُونَ أَحْسَنَهُ أُولَئِكَ الَّذِينَ هَدَى اللَّهُ وَأُولَئِكَ هُمُ الْأَوْلَىٰ

*“Yang mendengarkan perkataan lalu mengikuti apa yang paling baik diantaranya. Mereka itulah orang-orang yang telah diberi Allah petunjuk dan mereka itulah orang-orang yang mempunyai akal.”(QS. Az-Zumar : 18)*

Menurut tafsir Jalayn terhadap surah Az-Zumar ayat 18, menjelaskan tentang pentingnya orang-orang yang mendengarkan dan mengikuti nasihat yang terbaik di antara mereka, serta menjauhi hal-hal yang merugikan mereka, adalah termasuk orang-orang yang telah diberi petunjuk oleh Allah. Mereka adalah orang-orang yang memiliki kecerdasan dan pemikiran yang baik. cara mengikuti dan mendengarkan nasihat juga sesuai dengan memilah dan memilih jika terdapat suatu informasi, dan dianjurkan untuk memilih informasi yang baik. Hal ini berhubungan dengan adanya sistem klasifikasi ulasan pengguna pada aplikasi *travel online* ini, dengan adanya sistem klasifikasi ini maka pengguna akan dimudahkan dalam mencari informasi terkait suatu aplikasi yang baik, sehingga dapat memberikan kenyamanan, kelancaran, dan keselamatan ketika pengguna sedang bepergian. Begitu juga sistem ini berguna bagi perusahaan dalam mengelompokkan terkait masalah yang sedang terjadi pada suatu aplikasi berdasarkan ulasan pengguna yang telah diberikan.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

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.

#### 5.2 Saran

Dalam penelitian ini tentunya masih jauh dari kata sempurna. Terdapat kekurangan yang diperlukan untuk memperbaiki performa model. Berikut beberapa saran yang perlu dilakukan untuk penelitian selanjutnya:

1. Menggunakan ekstraksi fitur selain TF-IDF seperti *Word Embedding* dan *Word2Vec*.
2. Melakukan percobaan dengan menggunakan algoritma lainnya, seperti *Support Vector Machine (SVM)*, *Random Forest* dan algoritma lainnya. Hal ini dapat memberikan gambaran lebih lanjut tentang performa algoritma yang berbeda dalam meningkatkan kualitas klasifikasi ulasan.

## DAFTAR PUSTAKA

- Abbas, M., Memon, K. A., & Jamali, A. A. (2019). *Multinomial Naive Bayes Classification Model for Sentiment Analysis Multinomial Naive Bayes Classification Model for Sentiment Analysis*. March. <https://doi.org/10.13140/RG.2.2.30021.40169>
- Baiq Nurul Azmi, Arief Hermawan, & Donny Avianto. (2023). Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver. *JTIM : Jurnal Teknologi Informasi Dan Multimedia*, 4(4), 281–290. <https://doi.org/10.35746/jtim.v4i4.298>
- CNN Indonesia. (2019). *Aplikasi e-Commerce dan Online Travel Pilihan Milenial*. Cnn Indonesia. <https://www.cnnindonesia.com/teknologi/20190709195545-185-410676/aplikasi-e-commerce-dan-online-travel-pilihan-milenial>
- Diekson, Z. A., Prakoso, M. R. B., Putra, M. S. Q., Syaputra, M. S. A. F., Achmad, S., & Sutoyo, R. (2023). Sentiment analysis for customer review: Case study of Traveloka. *Procedia Computer Science*, 216(2022), 682–690. <https://doi.org/10.1016/j.procs.2022.12.184>
- Ferdino, I., & Rusli, A. (2019). Using naïve bayes classifier for application feedback classification and management in Bahasa Indonesia. *Proceedings of 2019 5th International Conference on New Media Studies, CONMEDIA 2019*, 217–222. <https://doi.org/10.1109/CONMEDIA46929.2019.8981830>
- Hasan, T., Matin, A., & Joy, M. S. R. (2020). Machine Learning Based Automatic Classification of Customer Sentiment. *ICCIT 2020 - 23rd International Conference on Computer and Information Technology, Proceedings*, 19–21. <https://doi.org/10.1109/ICCIT51783.2020.9392652>
- JoMingyu. (2023). *google-play-scraper 1.2.4*. Pypi. <https://pypi.org/project/google-play-scraper/>
- Oktaviani, V., Warsito, B., Yasin, H., Santoso, R., & Suparti. (2021). Sentiment analysis of e-commerce application in Traveloka data review on Google Play site using Naïve Bayes classifier and association method. *Journal of Physics: Conference Series*, 1943(1). <https://doi.org/10.1088/1742-6596/1943/1/012147>
- Poernomo, A. D., & Suharjito, S. (2019). Indonesian online travel agent sentiment analysis using machine learning methods. *Indonesian Journal of Electrical Engineering and Computer Science*, 14(1), 113. <https://doi.org/10.11591/ijeecs.v14.i1.pp113-117>
- Rahayu, C., & Andi. (2020). *Klasifikasi Teks*. Binus. <https://mti.binus.ac.id/2020/09/03/klasifikasi-teks/>

- Raja Suhud. (2023). *Tren Penggunaan Aplikasi Penjualan Tiket bakal Melejit*. Media Indonesia. <https://mediaindonesia.com/ekonomi/554815/tren-penggunaan-aplikasi-penjualan-tiket-bakal-melejit>
- Ramachandran, D., & Parvathi, R. (2019). Analysis of Twitter Specific Preprocessing Technique for Tweets. *Procedia Computer Science*, 165, 245–251. <https://doi.org/10.1016/j.procs.2020.01.083>
- Reddy, C. S. C., Kumar, K. U., Keshav, J. D., Prasad, B. R., & Agarwal, S. (2017). Prediction of star ratings from online reviews. *IEEE Region 10 Annual International Conference, Proceedings/TENCON, 2017-Decem*, 1857–1861. <https://doi.org/10.1109/TENCON.2017.8228161>
- Sari, E. Y. (2019). *Sentiment Analysis of Customer Satisfaction on Transportation Network Company Using Naive Bayes Classifier*.
- Sharif, O., Hoque, M. M., & Hossain, E. (2019). Sentiment Analysis of Bengali Texts on Online Restaurant Reviews Using Multinomial Naïve Bayes. *1st International Conference on Advances in Science, Engineering and Robotics Technology 2019, ICASERT 2019, 2019(Icasert)*, 1–6. <https://doi.org/10.1109/ICASERT.2019.8934655>
- Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 593–596. <https://doi.org/10.1109/ICACTM.2019.8776800>
- Sriram. (2022). *Multinomial Naive Bayes Explained: Function, Advantages & Disadvantages, Applications in 2023*. Upgrad. <https://www.upgrad.com/blog/multinomial-naive-bayes-explained/>
- Tabassum, A., & Patil, R. R. (2020). A Survey on Text Pre-Processing & Feature Extraction Techniques in Natural Language Processing. *International Research Journal of Engineering and Technology*, June, 4864–4867. [www.irjet.net](http://www.irjet.net)
- Tama, V. O., Sibaroni, Y., & Adiwijaya. (2019). Labeling Analysis in the Classification of Product Review Sentiments by using Multinomial Naive Bayes Algorithm. *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012036>
- Wiratama, G. P., & Rusli, A. (2019). Sentiment analysis of application user feedback in Bahasa Indonesia using multinomial naive bayes. *Proceedings of 2019 5th International Conference on New Media Studies, CONMEDIA 2019*, 223–227. <https://doi.org/10.1109/CONMEDIA46929.2019.8981850>
- Yuyun, Nurul Hidayah, & Supriadi Sahibu. (2021). Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter. *Jurnal RESTI (Rekayasa Sistem Dan*

*Teknologi Informasi*), 5(4), 820–826. <https://doi.org/10.29207/resti.v5i4.3146>

# **LAMPIRAN**

## LAMPIRAN-LAMPIRAN

### Lampiran I

#### Sampel Pelabelan Ulasan

No	Ulasan	Responden			Hasil
		Responden I	Responden II	Responden III	
		Kelas			
1	Disadvantages: Unable to check in flight tickets via this application.	Unhappy	Unhappy	Unhappy	Unhappy
2	The user interface is actually better when compared to competitors for similar applications. However, I cannot do payment settlement for all the payment methods offered.	Unhappy	Unhappy	Unhappy	Unhappy
3	It's easy to book flight from this app.	Satisfied	Satisfied	Satisfied	Satisfied
4	How do I change my email address? There is no option in account settings. Please help. Thanks.	Unhappy	Unhappy	Unhappy	Unhappy
5	A pretty useful apps to have	Satisfied	Satisfied	Satisfied	Satisfied
6	The app froze when I tried to upload photos on the review page. Also maybe some support for line breaks when writing a review? (LG V20 running Oreo 8.0)	Unhappy	Unhappy	Unhappy	Unhappy
7	Sometimes apps couldn't be operated, just reload reload. You guys as program developers need to fix it	Unhappy	Unhappy	Unhappy	Unhappy
8	After updating... the refund info doesn't work	Unhappy	Unhappy	Unhappy	Unhappy
9	good for flaying....	Satisfied	Unhappy	Satisfied	Satisfied
10	How come you can't log in, bro?	Unhappy	Unhappy	Unhappy	Unhappy
11	Easy way to book flight and other ticket. The bad thing is finding transit flight is hard.	Satisfied	Satisfied	Satisfied	Satisfied
12	very easy on payment	Satisfied	Satisfied	Satisfied	Satisfied
13	easy search and book hotel, lots of accomodations, to improve more option for booking payment e.g more bank for money transfer thru mbanking	Satisfied	Satisfied	Satisfied	Satisfied
14	In the filter, also add the hotel rooftop fool facilities	Satisfied	Satisfied	Satisfied	Satisfied
15	ugly product	Unhappy	Unhappy	Unhappy	Unhappy

## Lampiran II

### Pseudocode Preprocessing

```
FUNCTION clean_text(text):
    // Convert text to lowercase
    CONVERT text TO lowercase
    // Remove emoticons
    REMOVE emoticons FROM text USING regex '\\u....'
    // Remove URLs
    REMOVE URLs FROM text USING regex 'http\S+'
    // Remove punctuation
    REMOVE punctuation FROM text USING translate method
    // Remove leading and trailing whitespace
    STRIP whitespace FROM text
    RETURN cleaned text # remove whitespace
    text = text.strip()
    return text

FUNCTION remove_stopword(text):
    // Define stop words set
    SET stop_words TO list OF English stopwords
    // Split text into words
    SPLIT text INTO word_tokens
    // Filter out stop words
    SET filtered_text TO empty list
    FOR each word IN word_tokens:
        IF word NOT IN stop_words:
            ADD word TO filtered_text
    RETURN filtered_text JOINED by spaces

FUNCTION lemmatize(text):
    // Initialize lemmatizer
    SET lemmatizer TO instance OF WordNetLemmatizer
    // Split text into words
    SPLIT text INTO word_tokens
    // Apply lemmatization
    SET lemmatized_text TO empty list
    FOR each word IN word_tokens:
        ADD lemmatized word TO lemmatized_text
    RETURN lemmatized_text JOINED by spaces

FUNCTION case_folding(text):
    // Convert text to lowercase
    CONVERT text TO lowercase
    RETURN text
```