

JMO APP SENTIMENT ANALYSIS FROM GOOGLE PLAY STORE



Presented by
Roy Firman Sihombing

Contents

- 01 Business Introduction
- 02 Scraping Data
- 03 Exploratory Data Analysis
- 04 Data Labelling
- 05 Comment Analysis
- 06 Text Representation
- 07 Prediction Model
- 08 Model Performance Comparison
- 09 Model Evaluation, Conclusion, Business Solution

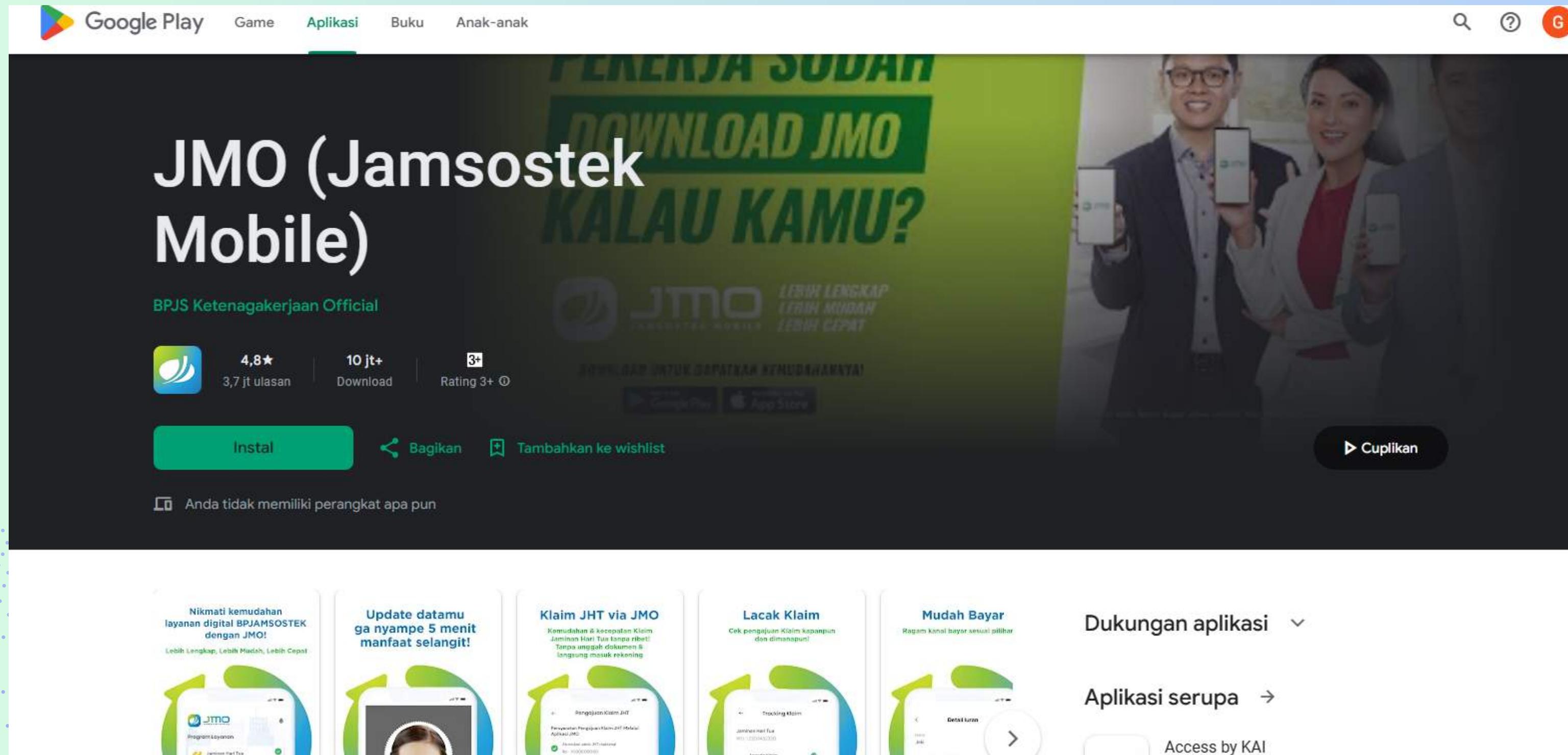
Business Introduction

JMO (Jaminan Sosial Mobile) adalah inovasi digital BPJS Ketenagakerjaan yang mempermudah peserta mengakses layanan jaminan sosial tenaga kerja, seperti klaim, cek saldo, dan informasi kepesertaan, tanpa harus datang ke kantor cabang. Aplikasi ini dirancang untuk memberikan pengalaman layanan yang lebih cepat, praktis, dan efisien.

Analisis sentimen terhadap ulasan pengguna di Google Play Store membantu BPJS memahami tingkat kepuasan serta mendeteksi masalah aplikasi. Temuan ini diharapkan mendukung perbaikan berkelanjutan dan pengambilan keputusan berbasis data di masa depan.



Scraping The Datasets



Data diambil dari hasil scraping menggunakan library google-play-scraper pada kolom review aplikasi JMO (Jamsostek Mobile) yang tersedia pada platform Google Play Store.

	review	rating
0	mengapa bantuan BSU tidak terlihat diakun saya...	3
1	mudah login dan informatif	5
2	bagus	5
3	sangat membantu	4
4	sangat membantu	5
5	mantap	5
6	bagus pelayanannya	5
7	sangat membantu kami terimakasih atas pelayana...	5
8	sangat membantu	5
9	Terimakasih aplikasih ini sangat membantu jadi...	5

Data hasil scraping disimpan dalam file CSV yang diberi nama ulasan_jmo_playstore.csv. Kumpulan data ini berisi 50.000 ulasan pengguna, di mana setiap baris mewakili satu ulasan dan peringkat yang diberikan oleh pengguna aplikasi JMO (Jamsostek Mobile) di platform Google Play Store.

Exploratory Data Analysis (EDA)

After EDA

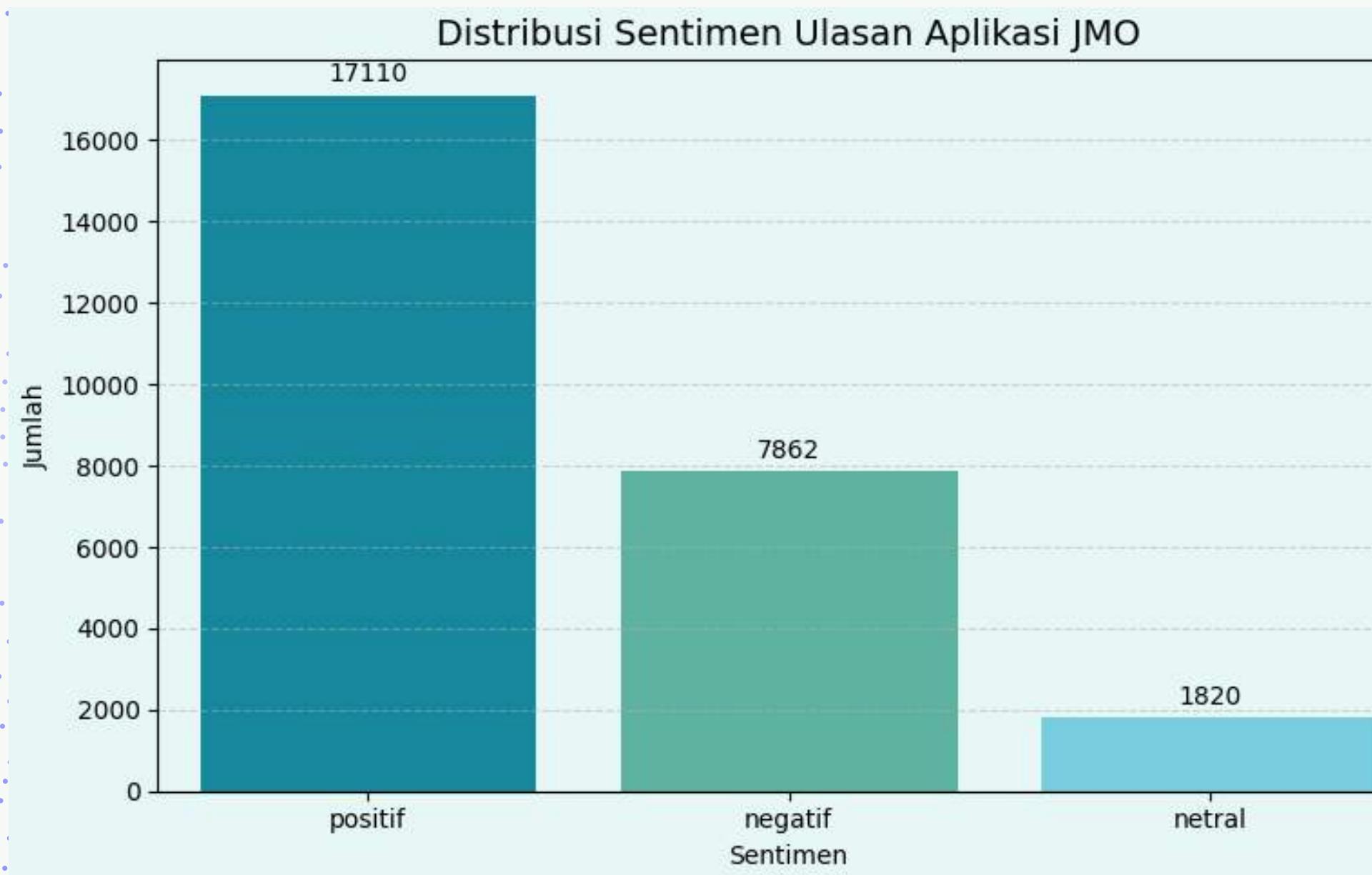
```
Data columns (total 2 columns):  
 #   Column  Non-Null Count Dtype  
 ---  
 0   review   50000 non-null  object  
 1   rating   50000 non-null  int64  
 dtypes: int64(1), object(1)  
 memory usage: 781.4+ KB
```

Before EDA

```
Data columns (total 2 columns):  
 #   Column  Non-Null Count Dtype  
 ---  
 0   review   26792 non-null  object  
 1   rating   26792 non-null  int64  
 dtypes: int64(1), object(1)  
 memory usage: 627.9+ KB
```

Setelah proses EDA pada data duplikat dan data kosong, Dataset dari 50000 berkurang menjadi 26792.

Data Labelling



Untuk menentukan sentimen sebagai nilai sebenarnya, proses pelabelan dilakukan secara otomatis berdasarkan peringkat data yang dikikis.

- Peringkat 4 hingga 5 akan diberi label sebagai komentar/ulasan positif.
- Peringkat 3 akan diberi label sebagai komentar/ulasan netral.
- Peringkat 1 hingga 2 akan diberi label sebagai komentar/ulasan negatif.

Pelabelan ini digunakan sebagai referensi untuk analisis sentimen dan dalam pelatihan model klasifikasi otomatis.

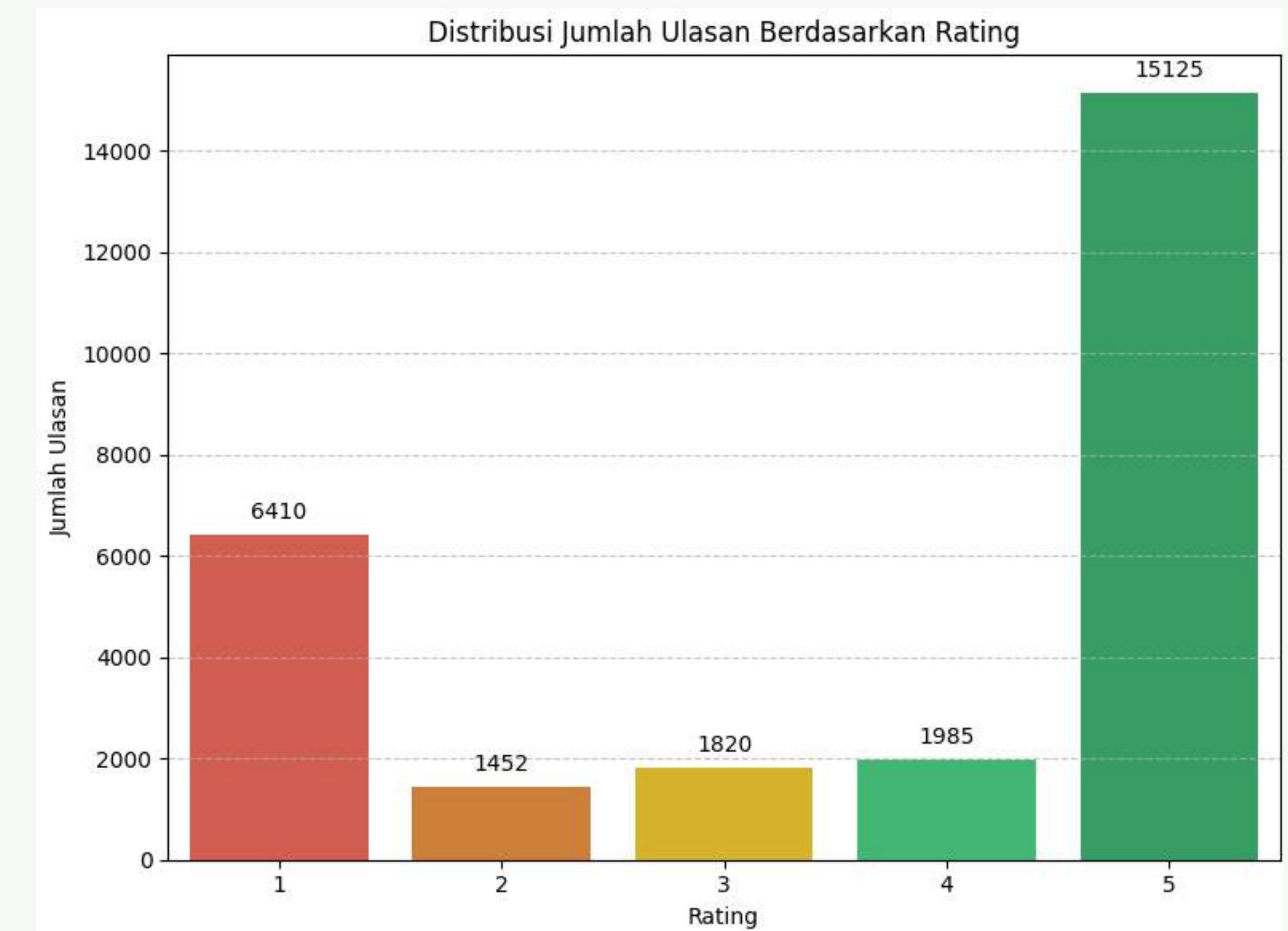


Comment Analysis

Rating Distribution

Berdasarkan distribusi peringkat pengguna aplikasi JMO di Google Play Store, mayoritas memberikan peringkat bintang 5 (15.125 ulasan), yang menunjukkan tingkat kepuasan yang tinggi. Sementara itu, peringkat bintang 1 (6.410 ulasan) merupakan peringkat kedua yang paling umum, yang mencerminkan ketidakpuasan di antara sebagian pengguna.

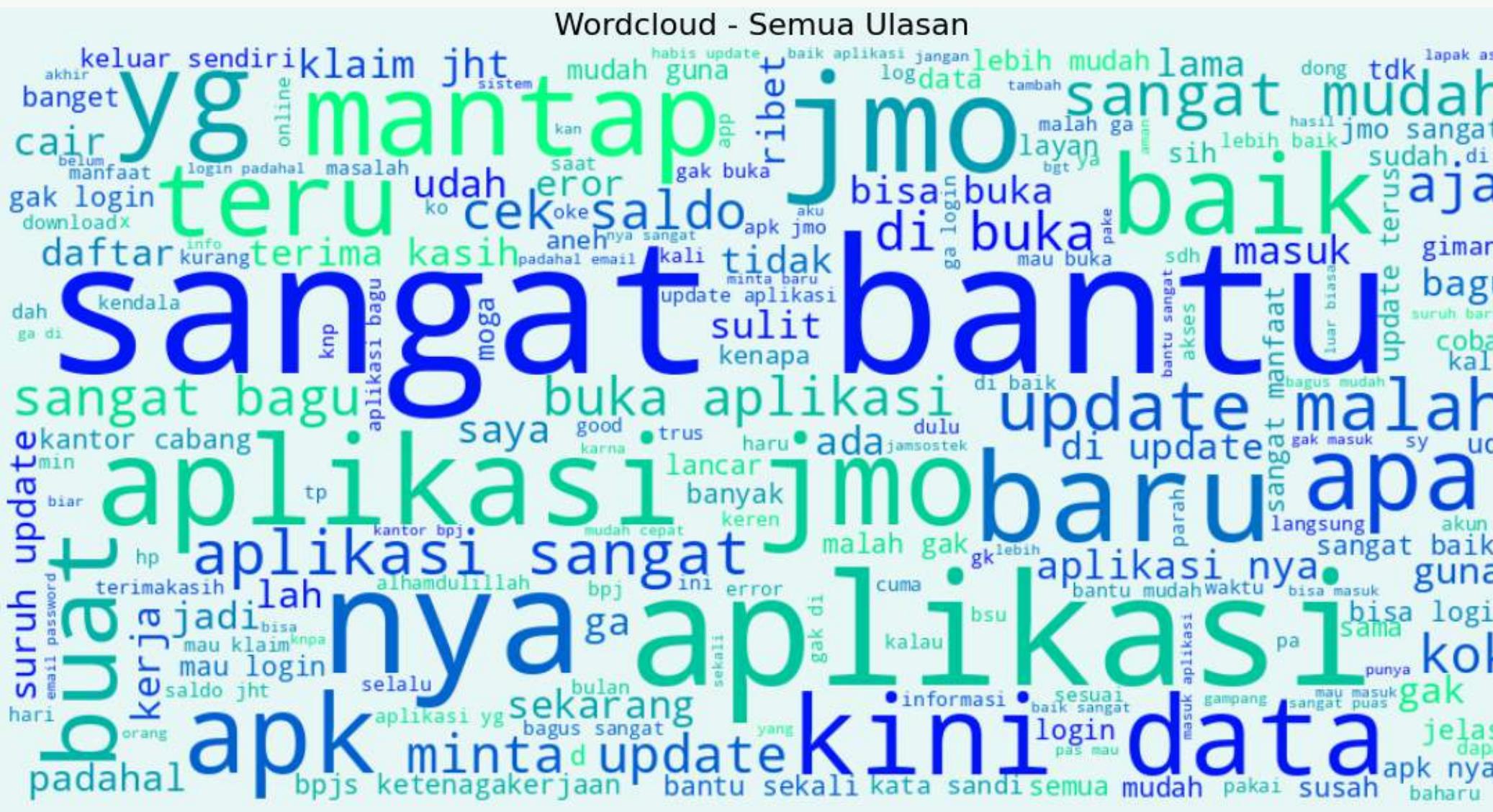
Sebaliknya, peringkat bintang 2, 3, dan 4 menerima ulasan yang jauh lebih sedikit. Pola ini menunjukkan kecenderungan pengguna untuk memberikan peringkat yang ekstrem, baik sangat positif maupun sangat negatif, dengan sedikit evaluasi moderat.



General Wordcloud

Kata "aplikasi" muncul paling sering (6.714 kali), menunjukkan fokus utama pengguna pada aplikasi JMO itu sendiri. Kata-kata seperti "sangat", "update", dan "jmo" juga banyak disebut, mengarah pada pengalaman penggunaan dan pembaruan aplikasi.

Kemunculan kata "mudah", "bantu", dan "baik" mencerminkan kesan positif terhadap kemudahan dan manfaat aplikasi. Secara umum, dominasi kata berkonotasi positif mengindikasikan bahwa sentimen pengguna terhadap JMO cenderung positif.



Kata	Frekuensi
aplikasi	6714
sangat	5012
update	4051
jmo	3341
mudah	3337
bantu	3139
nya	3000
buka	2797
login	2528
baik	2404





Specific Wordcloud

Wordcloud - Sentimen Positif



Wordcloud sentimen positif didominasi kata "sangat", "bantu", "mudah", "baik", dan "mantap", mencerminkan pengalaman pengguna yang puas dengan layanan aplikasi JMO. Kata "jmo" dan "aplikasi" juga tampil besar, menunjukkan kepercayaan dan kemudahan dalam penggunaan aplikasi.

Wordcloud - Sentimen Negatif



Wordcloud negatif memperlihatkan kata dominan seperti "update", "gak", "padahal", "eror", dan "login". Hal ini mencerminkan keluhan terkait kesulitan mengakses aplikasi setelah pembaruan atau masalah teknis lainnya seperti klaim gagal.

Wordcloud - Sentimen Netral



Kata-kata seperti "update", "baru", "login", dan "masuk" mendominasi wordcloud netral. Ini menunjukkan ulasan bersifat informatif atau deskriptif, tanpa emosi kuat, misalnya membahas perubahan versi atau proses login.

TEXT REPRESENTATION



TF-IDF (Term Frequency-Inverse Document Frequency)

metode representasi teks yang memberi bobot pada kata berdasarkan seberapa penting kata tersebut dalam satu ulasan dibandingkan seluruh kumpulan data ulasan. Dalam analisis review aplikasi JMO, TF-IDF membantu mengidentifikasi kata-kata yang unik dan bermakna, yaitu kata yang sering muncul di satu ulasan namun jarang muncul di ulasan lainnya. Hal ini memungkinkan model untuk lebih fokus pada informasi yang benar-benar relevan dalam setiap ulasan pengguna.



BoW (Bag of Words)

metode representasi teks yang sederhana namun efektif, di mana setiap ulasan diubah menjadi vektor berdasarkan frekuensi kemunculan kata. Dalam analisis review aplikasi JMO, BoW membantu mengidentifikasi seberapa sering kata-kata seperti "mudah", "eror", atau "bantu" muncul, tanpa memperhatikan urutan kata. Representasi ini memungkinkan sistem mengenali pola kata kunci yang sering digunakan pengguna dalam memberikan ulasan, sehingga memudahkan pengelompokan sentimen berdasarkan isi teks.

Prediction Model

Untuk model prediksi, digunakan 3 jenis model klasifikasi, yaitu Naive Bayes, Logistic Regression, dan Random Forest. Ketiga model ini dipilih karena mewakili pendekatan yang berbeda. Selain itu, untuk mengatasi ketidakseimbangan data, digunakan teknik oversampling dengan RandomOverSampler agar distribusi kelas lebih seimbang.

NAIVE BAYES

LOGISTIC REGRESSION

RANDOM FOREST



Model Performance Comparison



Models with TF-IDF Results

Models with TF-IDF Results					
	Model	Accuracy	Precision	Recall	F1 Score
	Random Forest (TF-IDF)	0.9191	0.9215	0.9191	0.9185
	Tuned Random Forest (TF-IDF)	0.9180	0.9206	0.9180	0.9174
	Tuned Logistic Regression (TF-IDF)	0.8328	0.8351	0.8328	0.8326
	Logistic Regression (TF-IDF)	0.8080	0.8025	0.8080	0.8086
	Tuned Naive Bayes (TF-IDF)	0.7517	0.7581	0.7517	0.7532
	Naive Bayes (TF-IDF)	0.7410	0.7495	0.7410	0.7430

Hasil perbandingan model dengan TF-IDF menunjukkan bahwa Random Forest, baik default maupun yang dituning, memberikan performa terbaik dengan akurasi >91%, F1 Score hingga 91.85%, serta keseimbangan precision dan recall di atas 91%.

Logistic Regression menempati posisi kedua dengan akurasi terbaik 83.28% dan F1 Score 83.26%, meskipun recall-nya masih di bawah Random Forest.

Naive Bayes memiliki performa terendah, dengan akurasi hanya 74–75% dan F1 Score di kisaran 74–75%, menunjukkan kurang optimalnya model ini untuk data TF-IDF.

Models with Bag of Words Results

Models with Bag of Words Results

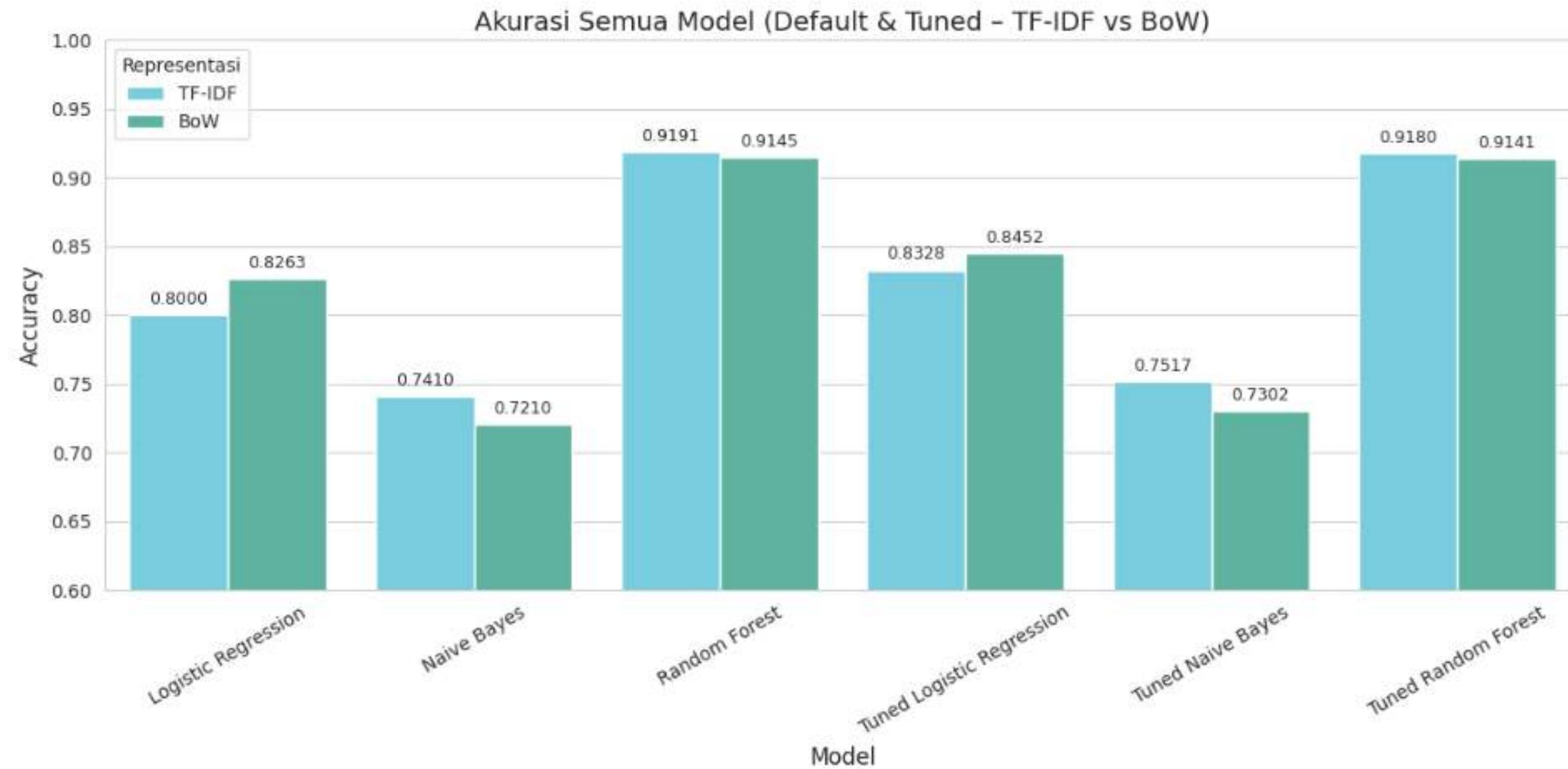
	Model	Accuracy	Precision	Recall	F1 Score
	Random Forest (BoW)	0.9145	0.9161	0.9145	0.9136
	Tuned Random Forest (BoW)	0.9141	0.9161	0.9141	0.9132
	Tuned Logistic Regression (BoW)	0.8452	0.8468	0.8452	0.8447
	Logistic Regression (BoW)	0.8263	0.8272	0.8263	0.8261
	Tuned Naive Bayes (BoW)	0.7302	0.7362	0.7302	0.7316
	Naive Bayes (BoW)	0.7210	0.7295	0.7210	0.7227

Hasil evaluasi dengan representasi Bag of Words menunjukkan bahwa Random Forest, baik standar maupun yang dituning, memberikan performa terbaik dengan akurasi >91% dan F1 Score hingga 91.36%, disertai keseimbangan precision dan recall yang stabil.

Logistic Regression menempati posisi kedua dengan akurasi tertinggi 84.52% dan F1 Score 84.47% pada versi tuned.

Sementara itu, Naive Bayes memiliki performa terendah, dengan akurasi hanya 73.02% dan F1 Score 73.16%, menunjukkan hasil yang kurang optimal untuk data BoW.

Accuracy of All Models (Default & Tuned – TF-IDF vs BoW)



Berdasarkan grafik "Accuracy Scores of All Models", dapat dilihat bahwa model dengan representasi teks Bag of Words (BoW) cenderung menghasilkan akurasi yang lebih tinggi dibandingkan model dengan TF-IDF, terutama setelah dilakukan tuning. Model dengan akurasi tertinggi adalah Tuned Naive Bayes (BoW), disusul oleh Naive Bayes (BoW) dan Tuned Logistic Regression (BoW), yang semuanya mencapai akurasi mendekati atau di atas 0.88.

Sebaliknya, model dengan akurasi terendah adalah Logistic Regression (TF-IDF), menandakan bahwa representasi TF-IDF tanpa tuning kurang optimal dalam kasus ini. Proses tuning hyperparameter terbukti meningkatkan performa hampir di semua model, baik dengan TF-IDF maupun BoW. Secara keseluruhan, grafik ini menunjukkan bahwa kombinasi representasi teks BoW dan tuning model memberikan hasil prediksi paling akurat pada dataset yang digunakan.

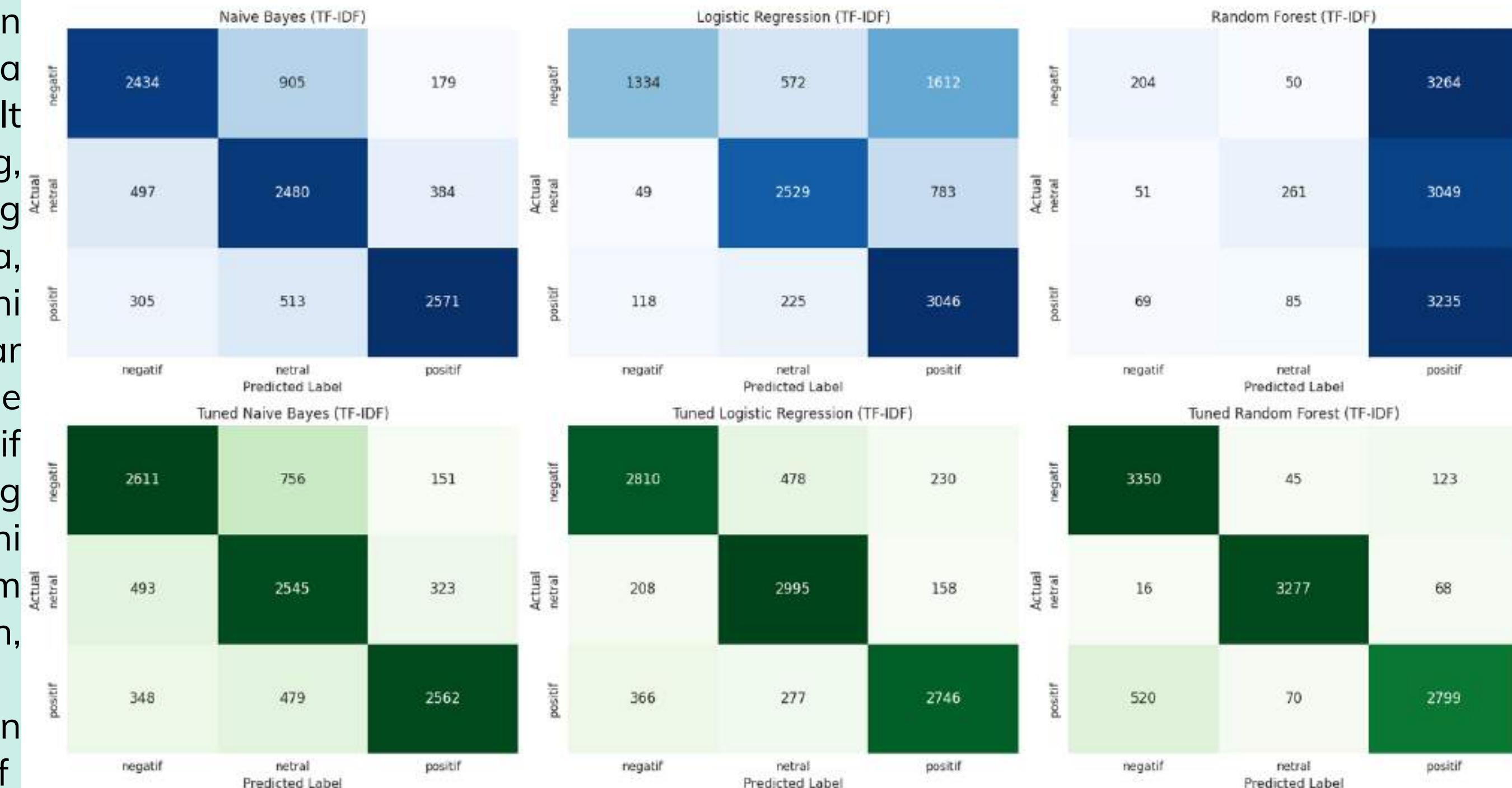
Model Evaluation



Berdasarkan enam confusion matrix pada grafik, terlihat bahwa Random Forest, baik default maupun setelah tuning, menunjukkan performa paling stabil dan seimbang. Misalnya, pada versi tuned, model ini mencatat true positif sebesar 2799, true netral 3277, dan true negatif 3350, dengan false positif dan false negatif masing-masing hanya 123 dan 520. Hal ini menunjukkan akurasi tinggi dalam mengenali ketiga kelas sentimen, terutama netral dan positif.

Sebaliknya, Logistic Regression default menghasilkan false negatif tertinggi (1612), menunjukkan kelemahan dalam mengenali sentimen positif. Setelah tuning, Naive Bayes dan Logistic Regression memang mengalami peningkatan dalam mengenali kelas negatif (true negatif naik, false positif turun), namun perbaikannya masih belum mampu menyaingi konsistensi dan keseimbangan Random Forest, yang tetap menjadi model dengan performa terbaik pada representasi TF-IDF.

Confusion Matrix - TF-IDF Models (Default vs Tuned)



CONCLUSION

Analisis terhadap 50.000 ulasan pengguna aplikasi JMO menunjukkan bahwa mayoritas pengguna memberikan respons positif, terlihat dari dominasi rating bintang 5 dan kata-kata seperti “mudah”, “cepat”, dan “terima kasih” dalam wordcloud.

Namun, sejumlah ulasan juga menunjukkan keluhan, terutama terkait masalah teknis, seperti gagal login dan klaim saldo, yang menjadi catatan penting bagi pengembang.

Meskipun beberapa model BoW memiliki akurasi tinggi, pendekatan TF-IDF, khususnya pada model Random Forest, memberikan performa paling seimbang dalam mengenali ketiga sentimen berdasarkan metrik F1-Score, precision, dan recall. Confusion matrix juga menunjukkan bahwa model mampu membedakan sentimen positif dan negatif dengan baik, meskipun sentimen netral masih cenderung ambigu. Temuan ini menegaskan bahwa strategi analisis sentimen yang tepat dapat memberikan insight penting untuk peningkatan layanan JMO.



Business Solution

Monitoring Feedback dengan Model Sentimen

Dengan model sentimen seperti Random Forest (TF-IDF), BPJS dapat memantau ulasan pengguna secara otomatis. Hasilnya membantu tim mengidentifikasi masalah dan menetapkan prioritas perbaikan fitur atau layanan dengan lebih cepat dan tepat.

-  **Peningkatan Fitur Login dan Aksesibilitas**
Banyak ulasan negatif dan netral menyoroti masalah saat login, dengan kata seperti “login”, “buka”, dan “error” sering muncul. Untuk mengatasinya, sistem perlu dioptimalkan agar lebih cepat saat login, dilengkapi panduan visual, serta mempermudah akses fitur “Lupa PIN/Password”.

-  **Optimasi Proses Klaim JHT dan JP**
Proses klaim JHT dan JP menjadi keluhan utama karena dianggap rumit dan membingungkan. Pengguna membutuhkan alur layanan yang lebih sederhana dan jelas di aplikasi JMO. Penambahan fitur chatbot untuk menjawab pertanyaan umum, serta notifikasi progres klaim secara real-time, dapat meningkatkan transparansi, mempercepat proses, dan memberikan pengalaman yang lebih baik bagi pengguna.

-  **Edukasi Digital tentang Layanan BPJS**
Banyak keluhan muncul karena kurangnya pemahaman pengguna terhadap layanan BPJS di JMO. Edukasi digital melalui media sosial, banner aplikasi, atau video interaktif dapat membantu menjelaskan fitur-fitur penting sekaligus mengurangi beban call center.

Let's Connect!

I'm always open to networking, collaboration, or just a quick chat.



linkedin.com/in/roy-firman-sihombing



roy-firman-sihombing.free.nf/



github.com/Roysihombing