

IMPLEMENTASI ALGORITMA RANDOM FOREST DALAM ANALISIS SENTIMEN ULASAN KONSUMEN APLIKASI KLIK INDOMARET

M. Satrio Prasajo¹⁾, Aslam Fatkhudin²⁾, Ahmad Khambali³⁾, Ichwan Kurniawan⁴⁾
Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pekajangan Pekalongan¹²³⁾,
Fakultas Teknologi Informasi, Institut Widya Pratama⁴⁾
satrioprasajo09@gmail.com¹⁾, ichwan,ana10@gmail.com⁴⁾

Abstrak

Dalam kegiatan bisnis, analisis sentimen pengguna merupakan hal yang penting untuk memahami pengalaman pengguna sebagai sarana untuk meningkatkan pelayanan. Analisis sentimen dapat dikaji melalui komentar pengguna pada sebuah aplikasi bisnis, salah satunya adalah aplikasi Klik Indomaret yang menyajikan layanan belanja online. Dengan menggunakan algoritma Random Forest, tujuan dari penelitian ini adalah untuk melihat perspektif pengguna dalam menilai aplikasi Klik Indomaret. Data diperoleh dari ulasan di Google Play Store yang memiliki sentimen positif dan negatif. Hasil memberikan informasi akurasi klasifikasi sebesar 88,56% dengan sentimen positif mendominasi sebesar 64,7%, diikuti oleh sentimen negatif sebesar 35,3%. Meskipun terdapat beberapa fitur yang perlu diperbaiki, penelitian ini menunjukkan kepuasan pengguna terhadap aplikasi. Studi ini memberi pengembang ide untuk meningkatkan layanan aplikasi.

Kata kunci: analisis sentimen, Klik Indomaret, Random Forest.

1. Pendahuluan

Kemajuan dan perkembangan Teknologi Informasi (TI) pada saat ini semakin pesat, dalam hal ini dapat memberikan kemudahan bagi masyarakat dalam pemenuhan kebutuhan dan menyelesaikan masalah dengan cepat. Salah satu contoh fenomena pada masa sekarang ini, terdapat pada industri retail yang telah mengalami transformasi yang pesat dengan tersedianya teknologi digital yang semakin canggih. Teknologi informasi sangat berperan dalam membantu proses bisnis suatu perusahaan dalam meningkatkan produktivitas, daya saing bisnis, dan efisiensi. Dengan memanfaatkan teknologi informasi perusahaan cenderung dapat bekerja secara cermat, tepat, dan efisien dari segi biaya ataupun waktu (Rosyadi et al., 2022).

Perkembangan teknologi menjadi salah satu hal yang dipertimbangkan masyarakat dalam memenuhi kebutuhan. Kemudahan dan kualitas yang diberikan oleh kemajuan teknologi merupakan nilai tambah dan keunggulan kompetitif (Farhatun Nisaul Ahadiyah, 2023). Teknologi mempengaruhi gaya hidup masyarakat yang menjadi serba instan dan menginginkan layanan yang dapat menunjang aktivitasnya seperti berbelanja kebutuhan sehari-hari.

Aplikasi mobile Klik Indomaret termasuk salah satu inovasi digital oleh Indomaret yang merupakan toko retail terbesar di Indonesia yang diluncurkan pada bulan 10 tahun 2015 dengan konsep Pusat Perbelanjaan Terpadu (*One Stop Online Shop*) yang berarti konsumen dapat berbelanja berbagai kebutuhan hanya dengan satu aplikasi (Yuwono & Anggraeni, 2023). Aplikasi ini dapat diunduh secara gratis melalui aplikasi Google Play Store. Pada Google play memiliki fitur penilaian atau ulasan dari pengguna, yang sangat memberikan pengaruh terhadap opini sebelum mengunduh aplikasi. Ulasan dan feedback dari pengguna yang termuat tentunya akan berpengaruh terhadap citra perusahaan (Manik et al., 2021). Dengan demikian, analisis sentimen menjadi salah satu solusi yang tepat untuk mengkaji secara mendalam mengenai opini pengguna terhadap aplikasi Klik Indomaret.

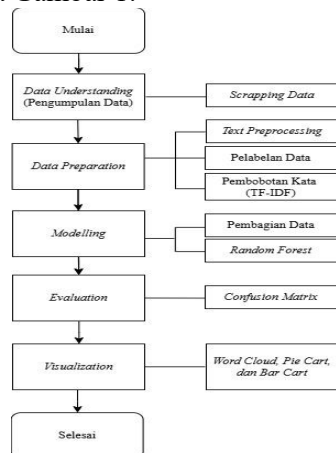
Random Forest merupakan salah satu algoritma machine learning yang sederhana dan fleksibel. Algoritma Random Forest merupakan salah satu algoritma yang sering digunakan, karena kesederhanaan dan keragamannya. Yaitu, dapat digunakan untuk tugas klasifikasi dan regresi (Azhari et al., 2021). Beberapa keunggulan dari algoritma Random Forest yaitu kemampuannya dalam meningkatkan keakuratan

data yang hilang dan data dengan nilai anomali (Kashyap, 2018). Selain itu, Random Forest memiliki proses pemilihan fitur yang memungkinkan pemilihan fitur terbaik yang dapat meningkatkan kinerja model klasifikasi. Dengan proses seleksi fitur ini, Random Forest dipastikan dapat bekerja secara efisien dengan kumpulan data yang besar dan kompleks (Kashyap, 2018).

Penelitian ini diharapkan dapat membantu perusahaan dalam memahami persepsi pengguna dan meningkatkan kualitas layanan pada aplikasi Klik Indomaret.

2. Metode Penelitian

Dalam penelitian ini, penulis membuat desain penelitian yang mencakup berbagai tahapan dalam penelitian, Adapun tahapannya dapat dilihat pada Gambar 1.



Gambar 1 Alur Penelitian

Adapun penjelasan dari setiap tahapan pada gambar 1 adalah sebagai berikut:

2.1 Data Understanding

Tahap ini bertujuan untuk mengumpulkan, mengidentifikasi, memahami, dan memastikan bahwa informasi yang disimpan akurat dan lengkap. Data yang didapat dalam penelitian ini berasal dari ulasan pengguna aplikasi Klik Indomaret di Google Play Store dengan teknik scraping data. Scraping data adalah teknik otomatisasi dengan mengekstrak data dari situs web, basis data, atau sejenisnya yang kemudian menyimpannya dalam bentuk tabel.

2.2 Data Preparation

Untuk memastikan bahwa data yang dihasilkan bersih dan siap untuk dianalisis dalam

penelitian, tahapan ini biasanya melibatkan pengubahan data yang tidak terorganisir menjadi bentuk yang lebih terorganisir. Pada tahap ini akan dilakukan proses text preprocessing, pelabelan data, dan pembobotan kata.

2.3 Modelling

Pada tahap ini akan dilakukan split data (pembagian data) dan klasifikasi Random Forest. Split data dilakukan sebelum proses klasifikasi dengan melakukan pembagian data menjadi data uji (data training) dan data latih (data testing). Data latih merupakan data yang digunakan pada tahap pelatihan untuk mengenali pola dalam data, sedangkan data uji adalah data yang diterapkan pada tahap pengujian untuk mengevaluasi model berdasarkan pola yang telah dipelajari selama proses pelatihan.

2.4 Evaluation

Pada tahap ini, evaluasi dilakukan untuk menilai kinerja model algoritma Random Forest dengan menggunakan confusion matrix.

2.5 Visualization

Setelah evaluasi selesai, tahapan selanjutnya adalah tahap visualisasi data, yaitu dalam word cloud, dan bar chart. Visualisasi bertujuan untuk membantu pembaca dalam memahami hasil penelitian dengan lebih mudah. Dengan menggunakan word cloud, visualisasi menunjukkan kata-kata yang paling sering muncul dalam setiap kelas, baik positif maupun negatif.

3. Hasil dan Pembahasan

Proses analisis sentimen yang dilakukan dengan Google Colab mencakup beberapa tahap sebagai berikut:

3.1 Preprocessing

Data ulasan pengguna aplikasi Klik Indomaret dikumpulkan melalui scraping diambil dari tanggal 30 November 2023 hingga 30 November 2024 menggunakan pustaka Google Play Scraper, menghasilkan sebanyak 7.616 entri.

Tahap selanjutnya, dilakukan tahap preprocessing untuk mengubah data mentah agar mudah dipahami, sebagai contoh pada gambar dibawah ini.

```
print(content_isi[['content_clean_tanpax', 'tokens']].head())
```

	content_clean_tanpax	tokens
0	mnggu kirim	[mnggu, kirim]
1	kecewa beli pls tanggal mei mei masuk status p...	[kecewa, beli, pls, tanggal, mei, mei, masuk, ...]
2	install kirim outlet alamat isi alamat jangkau...	[install, kirim, outlet, alamat, isi, alamat, ...]
3	barang kirim fiture cancel telepon costumer se...	[barang, kirim, fiture, cancel, telepon, costu...

Gambar 2 Contoh *Preprocessing Tokenizing*

Kemudian tahap pelabelan berdasarkan atribut content, score, dan sentimen. Ulasan dengan sentimen angka 0 dianggap sebagai ulasan dengan sentimen negatif, selanjutnya untuk sentimen dengan angka 1 dianggap sebagai ulasan dengan sentimen positif. Hasil pelabelan menunjukkan 4.927 ulasan positif dan 2.689 ulasan negatif.

```
# Membaca file CSV ('data_ulasan_dengan_label.csv')
df = pd.read_csv('data_ulasan_dengan_label.csv')

if 'sentimen' in df.columns:
    # Menghitung jumlah sentimen positif dan negatif
    jumlah_positif = (df['sentimen'] == 1).sum()
    jumlah_negatif = (df['sentimen'] == 0).sum()

    # Menampilkan hasil
    print(f"Jumlah Sentimen Positif: {jumlah_positif}")
    print(f"Jumlah Sentimen Negatif: {jumlah_negatif}")
```

Jumlah Sentimen Positif: 4927
Jumlah Sentimen Negatif: 2689

Gambar 3 Hasil Pelabelan Data

3.2 Transformation

Tahap transformasi dimulai dengan membagi dataset menjadi dua bagian yaitu data uji dan data latih. Dari 7.616 total data, didapatkan data latih sebanyak 6.854 dan data uji sebanyak 762.

```
print("Total Data: ", X.shape[0])
print("Data Latih: ", X_train.shape[0])
print("Data Uji: ", X_test.shape[0])
```

Total Data: 7616
Data Latih: 6854
Data Uji: 762

Gambar 4 Hasil Pembagian Data

3.3 Clasification

Pada tahap ini, data yang telah dibagi diklasifikasikan berdasarkan sentimen dengan menggunakan algoritma Random Forest. Hasil klasifikasi disajikan pada gambar berikut.

```
from sklearn.metrics import accuracy_score
print("Random Forest Classifier Accuracy Score: ", accuracy_score(y_predict_rf, y_test)*100, "%")
```

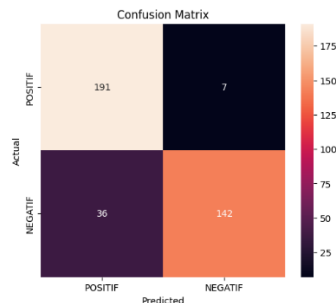
Random Forest Classifier Accuracy Score : 88.56382970234864 %

Gambar 5 Hasil Klasifikasi dengan Random Forest

Pada gambar diatas menunjukkan presentase yang diperoleh dengan metode algoritma Random Forest menghasilkan tingkat akurasi sebesar 88,56%. Hal tersebut menunjukkan metode tersebut mempunyai ketepatan yang cukup baik.

3.4 Evaluation

Evaluasi pada tahap ini dilakukan menggunakan confusion matrix. Hasil evaluasi terhadap data yang telah diklasifikasikan dengan algoritma Random Forest disajikan pada diagram berikut.

Gambar 5 Hasil Evaluasi *Confusion Matrix*

Dari diagram *confusion matrix* di atas, menunjukkan 191 untuk *true positive*, 142 untuk *true negative*, 36 untuk *false positive*, dan 7 untuk *false negative*.

3.5 Visualization

Data ulasan divisualisasikan menggunakan word cloud, dalam melihat kata-kata yang paling sering muncul. Gambar 6 menampilkan word cloud untuk ulasan positif, sedangkan Gambar 7 menunjukkan representasi visual untuk ulasan negatif.



Gambar 6 Word Cloud Ulasan Positif



Gambar 7 Word Cloud Ulasan Negatif

Berdasarkan visual word cloud diatas, Kata-kata yang sering muncul dalam sentimen positif, seperti "bantu," "belanja," "mudah," dan "mantap," mencerminkan kepuasan pengguna terhadap layanan aplikasi. Sementara itu, kata-kata dalam sentimen negatif, seperti "barang," "pesan," "bayar," dan "refund," mengindikasikan keluhan terkait proses transaksi dan pengiriman.

4. Kesimpulan dan Saran

Kesimpulan

Hasil Penelitian ini dapat disimpulkan bahwa algoritma Random Forest efektif dalam mengklasifikasikan sentimen ulasan aplikasi Klik Indomaret, dengan tingkat akurasi sebesar 88,56%. Hasil analisis sentimen menunjukkan bahwa 64,7% ulasan bersifat positif dan 35,3% bersifat negatif. Dominasi sentimen positif mencerminkan kepuasan pengguna terhadap layanan aplikasi, sementara sentimen negatif menyoroti aspek yang perlu diperbaiki, seperti sistem pembayaran, pengiriman, dan ketersediaan produk.

Saran

Untuk meningkatkan kualitas layanan, pengembang aplikasi Klik Indomaret disarankan melakukan analisis sentimen secara berkala guna memahami tren opini pengguna. Selain itu, perbaikan terhadap aspek yang sering dikritik dalam ulasan negatif dapat meningkatkan kepuasan pelanggan. Penelitian selanjutnya dapat mengeksplorasi metode lain, seperti SVM, Naïve Bayes atau deep learning, serta menggunakan

dataset yang lebih besar dan beragam untuk meningkatkan akurasi model.

DAFTAR PUSTAKA

- Azhari, M., Situmorang, Z., & Rosnelly, R. (2021). Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes. *Jurnal Media Informatika Budidarma*, 5(2), 640. <https://doi.org/10.30865/mib.v5i2.2937>
- Farhatun Nisaul Ahadiyah. (2023). Perkembangan Teknologi Infomasi Terhadap Peningkatan Bisnis Online. *INTERDISIPLIN: Journal of Qualitative and Quantitative Research*, 1(1), 41–49. <https://doi.org/10.61166/interdisiplin.v1i1.5>
- Kashyap, P. (2018). Machine learning for decision makers: Cognitive computing fundamentals for better decision making. In *Machine Learning for Decision Makers: Cognitive Computing Fundamentals for Better Decision Making*. Apress. <https://doi.org/10.1007/978-1-4842-2988-0>
- Manik, G., Ernawati, I., & Nurlaili, I. (2021). ANALISIS SENTIMEN PADA REVIEW PENGGUNA E-COMMERCE BIDANG PANGAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (Studi Kasus: Review Sayurbox dan Tanihub pada Google Play). *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, 2(2), 64–74. <https://conference.upnvj.ac.id/index.php/senami/ka/article/view/1630>
- Rosyadi, I., Adi, F., & Yusuf, M. (2022). Pengaruh Pengalaman on the Job Training (Ojt), Kemampuan Memasuki Dunia Kerja, Informasi Memasuki Dunia Kerja Terhadap Kesiapan Dunia Kerja (Studi Kasus Siswa Kelas Xii Smk Muhammadiyah Ulujami Pemalang Jurusan Otomotif Tkr). *Neraca*, 18(1), 148–157. <https://doi.org/10.48144/neraca.v18i1.1195>
- Yuwono, A. R., & Anggraeni, N. S. (2023). Persepsi Elemen Visual dan Layout User Interface Aplikasi Alfa Gift dan Klik Indomaret. *Gestalt*, 5(1), 55–72. <https://doi.org/10.33005/gestalt.v5i1.135>