

# Klasifikasi Sentimen Pengguna Berdasarkan Ulasan Aplikasi JKN Mobile menggunakan Model KNN

Siti Rihastuti<sup>1</sup>, Afnan Rosyidi<sup>2</sup>, Handoko<sup>3</sup>, M Nur Juniadi<sup>4</sup>

<sup>1,2</sup>Program Studi Manajemen Informatika

<sup>3,4</sup>Program Studi Komputersisasi Akuntansi

STMIK AMIKOM Surakarta

Surakarta, Jawa Tengah, Indonesia

e-mail: <sup>1</sup>siti@dosen.amikomsolo.ac.id, <sup>2</sup>afnanrosyidi@gmail.com, <sup>3</sup>handoko@dosen.amikomsolo.ac.id,

<sup>4</sup>nur.juniadi@dosen.amikomsolo.ac.id

Correspondence : e-mail: siti@dosen.amikomsolo.ac.id

Diajukan: 14 Agustus 2024; Direvisi: 24 Agustus 2024; Diterima: 25 Agustus 2024

## Abstrak

Penelitian ini bertujuan untuk mengetahui sentimen pengguna saat mengakses aplikasi JKN mobile. JKN mobile merupakan aplikasi layanan kesehatan nasional yang dapat diakses dan dimanfaatkan secara luas oleh masyarakat. Pengumpulan data menggunakan teknik scraping pada ulasan pengguna JKN mobile dari platform Google Playstore sehingga diperoleh data sejumlah 1000 ulasan. Pengolahan dan pengujian data menggunakan algoritma K-Nearest Neighbors (KNN) dan Google Colab. Dari 1000 dataset yang dikumpulkan, dataset dibagi menjadi 20% data tes dan sisanya menjadi data training. Hasil pengujian menggunakan algoritma KNN menunjukkan sentimen yang positif dari pengguna, dengan nilai  $K=11$  yang menghasilkan tingkat akurasi tertinggi sebesar 89%, presisi 97% dan recall 86%. Hal tersebut menunjukkan bahwa model KNN dapat digunakan untuk mengukur sentimen positif pengguna saat mengakses JKN mobile.

**Kata kunci:** knn, klasifikasi sentimen, ulasan pengguna jkn mobile, google colab.

## Abstract

This study aims to determine user sentiment when accessing the JKN mobile application. JKN mobile is a national health service application that can be accessed and utilized widely by the public. Data collection uses scraping techniques on JKN mobile user reviews from the Google Playstore platform so that 1000 reviews of data are obtained. Data processing and testing use the K-Nearest Neighbors (KNN) algorithm and Google Colab. Of the 1000 datasets collected, the dataset is divided into 20% test data and the rest is training data. The test results using the KNN algorithm show positive sentiment from users, with a value of  $K = 11$  which produces the highest accuracy rate of 89%, precision of 97% and recall of 86%. This shows that the KNN model can be used to measure positive user sentiment when accessing JKN mobile..

**Keywords:** knn, sentiment classification, jkn mobile user reviews, google colab.

## 1. Pendahuluan

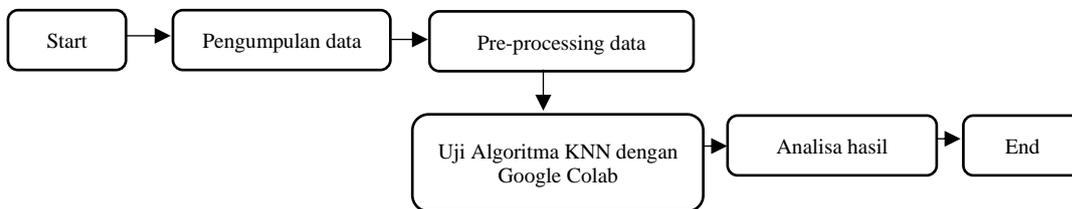
JKN mobile merupakan aplikasi berbasis mobile yang diluncurkan oleh Badan Penyelenggara Jaminan Sosial (BPJS) yang ditujukan untuk memperluas akses dan meningkatkan kualitas layanan kesehatan kepada masyarakat. Berbagai fitur unggulan tersedia di aplikasi JKN mobile, diantaranya peserta bisa mendaftar secara *online* untuk memperoleh layanan kesehatan, pencarian rumah sakit, cek status kepesertaan, riwayat kesehatan pasien serta peserta bisa mengganti lokasi layanan fasilitas kesehatan tingkat 1 sesuai dengan domisili peserta JKN. Pengguna Layanan BPJS Kesehatan di seluruh Indonesia mencapai 241,75 juta jiwa dari total penduduk Indonesia sebesar 275,77 juta jiwa[1]. Penelitian ini akan mengklasifikasi penilaian dan persepsi masyarakat tentang pengaksesan dan penggunaan aplikasi JKN mobile. Persepsi masyarakat diambil dari data ulasan yang ada di platform Google Playstore terkait aplikasi JKN mobile. Algoritma yang akan digunakan adalah KNN (*K-Nearest Neighbors*). Algoritma KNN berfungsi dalam klasifikasi yang dilakukan kepada objek yang didasari terhadap data pembelajaran yang

memiliki jarak yang dekat dengan objek yang telah diuji. Pendekatan yang dilakukan pada pencarian kasus perhitungan pendekatan antara masalah baru dengan masalah sebelumnya dengan melakukan penyetaraan bobot dari penjumlahan fitur yang ada.

Beberapa penelitian telah membahas tentang implementasi model KNN terhadap sentimen kepuasan pengguna. Diantaranya pengukuran persepsi konsumen terhadap penggunaan aplikasi JKN mobile[1]. Penelitian tersebut menggunakan *Technology Acceptance Model* (TAM) yang menghasilkan persepsi bahwa masyarakat belum puas dalam memperoleh kemudahan saat menggunakan aplikasi JKN yang ditunjukkan dengan banyaknya keluhan pengguna mengenai fitur dan performa dari aplikasi, sedangkan pada penelitian ini akan menggunakan model KNN dan Google Colab. Pada penelitian[2] menggunakan metode *servqual* dan kuesioner untuk menganalisis kepuasan masyarakat terhadap aplikasi JKN mobile, dan *Model End-User Computing Satisfication* (EUCS)[3] digunakan untuk mengukur penggunaan aplikasi JKN mobile, sedangkan penelitian ini menggunakan ulasan pengguna di aplikasi JKN mobile pada platform Google Playstore. Pengukuran kepuasan pengguna JKN mobile menggunakan variabel EUCS yaitu *Content, Accuracy, Format, Ease of Use* dan *Timeliness* (EUCS) [4] menunjukkan hasil bahwa pengguna mampu merasakan kemudahan dan peningkatan efektifitas pelayanan.

Selain untuk mengukur kepuasan, KNN juga bisa digunakan untuk mengklasifikasikan tingkat stres individu dengan akurasi sebesar 100%[5]. Penelitian yang lainnya KNN mampu mengklasifikasikan sentimen kesehatan mental di twitter[6][7], mengukur kepuasan pengguna PLN Mobile[8], menganalisa kepuasan pengguna e-commerce[9], mengukur tingkat kepuasan pelanggan dalam memanfaatkan platform e-commerce[10], mendeteksi secara dini gejala depresi dan kesehatan mental[11], KNN digabungkan dengan Naïve Bayes dalam klasifikasi tingkat keparahan korban kecelakaan[12] dan diagnosa diabetes[13], membandingkan model KNN, Decision Tree dan Naïve Bayes dalam mengukur sentimen masyarakat pengguna layanan BPJS [14] juga bisa digunakan untuk prediksi penyakit jantung dengan KNN[15]. Penelitian ini ditujukan untuk mengetahui sentimen masyarakat terhadap penggunaan aplikasi JKN mobile agar tujuan negara dalam meningkatkan kualitas pelayanan kesehatan tercapai sesuai dengan UU Jaminan kesehatan diselenggarakan dengan tujuan menjamin agar peserta memperoleh manfaat pemeliharaan kesehatan dan perlindungan dalam memenuhi kebutuhan dasar kesehatan[16].

## 2. Metode Penelitian



Gambar 1. Alur Penelitian

### a. Pengumpulan data

*Dataset* yang akan digunakan merupakan data yang dikumpulkan dari ulasan atau *review* yang ditulis oleh pengguna JKN mobile di aplikasi Google Playstore sebanyak 1000 data. *Software* Google Colab digunakan untuk proses pengambilan data menggunakan teknik *scraping*.

```
!pip install google-play-scraper
```

```
from google_play_scraper import Sort, reviews_all
import pandas as pd
```

```
# Fetch reviews
try:
    scrapreview = reviews_all(
        'app.bpjs.mobile',
        lang='id',
        country='id',
        sort=Sort.NEWEST,
        filter_score_with=None
    )

    # Convert to DataFrame
    app_reviews_df = pd.DataFrame(scrapreview)

    # Limit to 1000 reviews
    app_reviews_df = app_reviews_df.head(1000)

    # Save to CSV
    app_reviews_df.to_csv('jkn3.csv', index=None, header=True)

    print("Reviews successfully scraped and saved.")
except Exception as e:
    print(f"An error occurred: {e}")
```

Kolom *dataset* yang terkumpul :

```
Column names in the dataset:
Index(['reviewId', 'userName', 'userImage', 'content', 'score',
       'thumbsUpCount', 'reviewCreatedVersion', 'at', 'replyContent',
       'repliedAt', 'appVersion'],
      dtype='object')
```

Menampilkan jumlah *dataset* yang digunakan :

```
[9] df.shape
     #Menampilkan jumlah baris dan kolom dataset
(1000, 11)
```

Gambar 1. Jumlah *dataset*

b. *Pre-processing* data

Selanjutnya adalah melakukan preprocessing data seperti:

- 1) Pembersihan data
- 2) Pemilihan fitur
- 3) Tranformasi
- 4) *Feature engineering*

Ini berfungsi untuk membersihkan nilai/value yang kosong dan bersifat redundansi, dan juga pemilihan atribut karena tidak semua atribut akan dipakai. Pembersihan dan penyaringan data dilakukan terhadap *dataset* supaya data yang digunakan mampu menghasilkan *output* yang akurat dan sesuai yang diharapkan.

	content	score	thumbsUpCount
0	Nmr handphone belum terdaftar malah katanya su...	1	0
1	mau daftar kok susah sekali	1	0
2	Keren aplikasinya 🍌🍌🍌	5	0
3	Sangat membantu	5	0
4	Baru coba	5	0

Gambar 2 . Hasil *Pre-processing* data

c. Pengujian dengan algoritma KNN dan Google Colab

*Dataset* diuji menggunakan algoritma KNN dan Google colab, *dataset* dipisah ke dalam data testing dan data training.

**3. Hasil dan Pembahasan**

Sebanyak 1000 *dataset* diuji menggunakan algoritma KNN dan Google Colab. *Dataset* dipisah kedalam data testing dan data training sebesar 20% dan 80% dengan nilai K dimulai dari 1,3,5,7,9 dan 11 .

```
# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
# Convert text data to numerical features using TF-IDF
vectorizer = TfidfVectorizer()
X_train_vec = vectorizer.fit_transform(X_train)
X_test_vec = vectorizer.transform(X_test)
```

Gambar 3. Memisahkan data testing dan data training  
Selanjutnya dicari akurasi, presisi dan recall.

```
# Initialize and train the KNN classifier
knn = KNeighborsClassifier(n_neighbors=5)
knn.fit(X_train_vec, y_train)

# Make predictions
y_pred = knn.predict(X_test_vec)

# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)

print(f"Accuracy: {accuracy:.2f}")
print(f"Precision: {precision:.2f}")
print(f"Recall: {recall:.2f}")

# Detailed classification report
print("\nClassification Report:")
print(classification_report(y_test, y_pred))
```

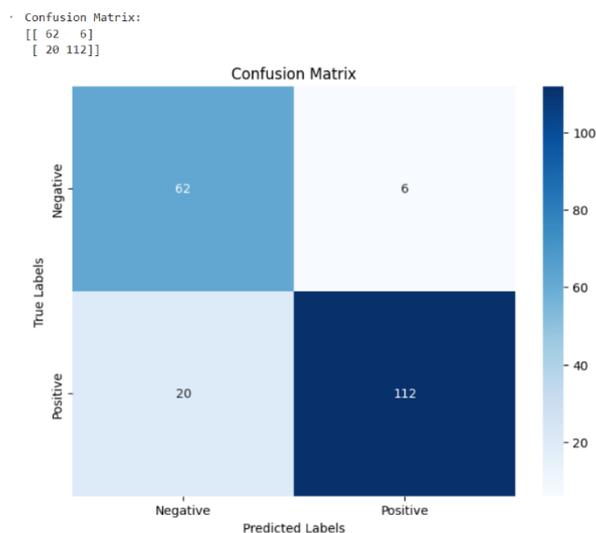
Gambar 4. Mencari akurasi, presisi dan recall  
Dari hasil pengujian diperoleh hasil seperti pada Tabel 1.

Tabel 1. Hasil akurasi, recall, presisi

Jumlah K	Akurasi	Presisi	Recall
K=1	0,80	0,97	0,71
K=3	0,80	0,97	0,71
K=5	0,87	0,95	0,85
K=7	0,88	0,95	0,86
K=9	0,88	0,96	0,85
K=11	0,89	0,97	0,86

Tabel 1 menunjukkan performa model *K-Nearest Neighbors* (KNN) dalam mengklasifikasikan ulasan pengguna aplikasi JKN *Mobile* dengan nilai K yang berbeda. Pada nilai K=1 dan K=3, hasil akurasinya 80%, presisi 97% dan recall 71%. Pada nilai K=5, hasil akurasi 87%, presisi 95% dan recall 85%. Pada nilai K=7, hasil akurasi 88%, presisi 96% dan recall 85%. nilai K=9, hasil akurasi 88%, presisi 96% dan recall 85%. Dan pada nilai K=11, hasil akurasi 89%, presisi 97% dan recall 86%.

KNN membagi hasil menjadi dua kategori yaitu negatif dan positif dalam visualisasi confusion matrix. Gambar 5 menunjukkan bahwa dari 200 ulasan yang diberikan oleh pengguna, terdapat 62 ulasan yang benar-benar negatif, 20 ulasan negatif namun sebenarnya positif, 112 ulasan diperkirakan positif dan 6 ulasan negatif.



Gambar 5. Confusion Matrix

Model KNN yang diterapkan memiliki kinerja yang baik dengan akurasi sebesar 87% dan keseimbangan yang baik antara *precision* dan *recall*, seperti yang ditunjukkan oleh *weighted F1-Score* sebesar 0.87.

Classification Report:				
	precision	recall	f1-score	support
0	0.76	0.91	0.83	68
1	0.95	0.85	0.90	132
accuracy			0.87	200
macro avg	0.85	0.88	0.86	200
weighted avg	0.88	0.87	0.87	200

Gambar 6. *Classification report*

#### 4. Kesimpulan

Berdasarkan hasil dan pembahasan tingkat kepuasan pengguna dalam memanfaatkan aplikasi JKN mobile menggunakan algoritma KNN dan Google Colab pada 1000 *dataset* ulasan pengguna, pada nilai  $K=11$  diperoleh akurasi tertinggi sebesar 89%, presisi 97 % dan recall 86%. Dari 200 data testing yang diuji, menghasilkan akurasi 87%. Hal tersebut menunjukkan bahwa model KNN dapat digunakan untuk mengukur sentimen positif pengguna saat mengakses JKN mobile. Saran untuk penelitian selanjutnya adalah dapat diterapkannya model KNN yang dibandingkan atau dikombinasikan dengan beberapa model atau algoritma yang lain pada penelitian sejenis agar mendapatkan hasil akurasi yang lebih baik.

#### Daftar Pustaka

- [1] D. Rahmawati, A. Febisatria Sekolah Tinggi Ilmu Ekonomi Surakarta Aditya Liliyan Sekolah Tinggi Ilmu Ekonomi Surakarta Alamat, J. Slamet Riyadi No, and K. Kabupaten Sukoharjo Jawa Tengah, "Analisis Persepsi Konsumen Terhadap Penggunaan Aplikasi Mobile JKN," vol. 2, no. 2, pp. 384–394, 2024, [Online]. Available: <https://doi.org/10.61722/jiem.v2i2.976>
- [2] I. C. Firmana, S. G. Amalia, T. A. Syahputra, N. Lutfiyana, K. Kunci-Bpjs, and S. Kesehatan, "Analisis Kepuasan Masyarakat Terhadap Aplikasi JKN Mobile Dengan Metode Servqual di Klinik Beringin," *J. Sist. Inf.*, vol. 12, no. 1, pp. 1–9, 2023, [Online]. Available: <https://ejournal.antarbangsa.ac.id/jsi/article/view/507>
- [3] A. Nurul Jannah, I. Susanto, D. Putra Rakhmadani, and P. Korespondensi, "Analisis Penggunaan Aplikasi Mobile JKN dengan Metode EUCS," *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 7, no. 3, pp. 1491–1502, 2023, [Online]. Available: <http://doi.org/10.33395/remik.v7i3.12826>
- [4] R. A. Wigatie and I. Zainafree, "Analisis Kepuasan Pengguna Aplikasi Mobile JKN di BPJS Kesehatan Kantor Cabang Magelang," *Politek. Kesehat. kendari*, vol. 15, no. 2, pp. 1–7, 2023, [Online]. Available: <https://repository.pnb.ac.id/7328/>
- [5] V. Manurung and A. F. Rozi, "TIN: Terapan Informatika Nusantara Analisis Perbandingan Algoritma K-Nearest Neighbor dan Decision Tree Pada Klasifikasi Tingkat Stress Individu TIN: Terapan Informatika Nusantara," *TIN Terap. Inform. Nusant.*, vol. 5, no. 1, pp. 73–80, 2024, doi: 10.47065/tin.v5i1.5286.
- [6] M. L. Wicaksono, R. Rusdah, and D. Apriana, "Sentiment Analysis Of Mental Health Using K-Nearest Neighbors On Social Media Twitter," *Bit (Fakultas Teknol. Inf. Univ. Budi Luhur)*, vol. 19, no. 2, p. 98, 2022, doi: 10.36080/bit.v19i2.2042.
- [7] A. Ilham and W. Pramusinto, "MENTAL PADA TWITTER MENGGUNAKAN ALGORITME K-NEAREST NEIGHBOR ANALYSIS OF PUBLIC SENTIMENT ON MENTAL HEALTH ON TWITTER USING K-NEAREST NEIGHBOR ALGORITHM," *Semin. Nas. Mhs. Fak. Teknol. Inf.*, vol. 2, no. September, pp. 539–547, 2023.
- [8] S. Syafrizal, M. Afdal, and R. Novita, "Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 10–19, 2023, doi: 10.57152/malcom.v4i1.983.
- [9] A. Van Fadhila *et al.*, "Implementasi Metode Machine Learning Untuk Mendeteksi Tingkat Stress Manusia Berdasarkan Kualitas Tidur," *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, vol. 4, no. 1, pp. 130–143, 2023.
- [10] T. N. Halim, R. Martin, and ..., "Klasifikasi Kepuasan Pelanggan Terhadap Platform E-Commerce dengan Metode K-Nearest Neighbor (K-NN)," *Jurasik (Jurnal Ris. ...)*, vol. 8, pp. 512–523, 2023, [Online]. Available:

- <http://ejurnal.tunasbangsa.ac.id/index.php/jurasik/article/view/636><https://ejurnal.tunasbangsa.ac.id/index.php/jurasik/article/download/636/609>
- [11] M. Anastasia, V. S. Maulivia, and S. Suharjo, "Metode Pembelajaran Mesin Untuk Menilai Data Depresi Dan Kesehatan Mental," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 7, no. 3, pp. 606–612, 2024, doi: 10.31539/intecom.v7i3.9584.
- [12] D. S. Wisdayani, "Perbandingan Algoritma K-Nearest Neighbor Dan Naive Bayes Untuk Klasifikasi Tingkat Keparahan Korban Kecelakaan Lalu Lintas Di Kabupaten Pati Jawa Tengah," *J. Dr. Diss. Muhammadiyah Univ. Semarang*, pp. 9–25, 2019, [Online]. Available: <http://repository.unimus.ac.id/id/eprint/3832>[http://repository.unimus.ac.id/3832/8/Jurnal Ilmiah.pdf](http://repository.unimus.ac.id/3832/8/JurnalIlmiah.pdf)
- [13] Q. A. Puteri, T. Sagirani, and J. Lemantara, "Perbandingan Algoritma Naive Bayes dan K-Nearest Neighbor (KNN) untuk Mengetahui Keakuratan Diagnosa Penyakit Diabetes," *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 3, pp. 247–254, 2023, doi: 10.25077/teknosi.v9i3.2023.247-254.
- [14] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 646, 2021, doi: 10.32493/informatika.v5i4.7622.
- [15] H. W. Dhany, "Performa Algoritma K-Nearest Neighbour dalam Memprediksi Penyakit Jantung," *Semin. Nas. Inform.*, pp. 176–179, 2021, [Online]. Available: <https://www.kaggle.com/>
- [16] R. Indonesia, "Undang-Undang Nomor 40 Tahun 2004," 2004.