JEMSI: Jurnal Ekonomi Manajemen Sistem Informasi
E-ISSN: 2686-5238
P-ISSN: 2686-4916
https://dinastirev.org/JEMSI dinasti.info@gmail.com +62 811 7404 455
DOI: https://doi.org/10.38035/jemsi.v7i2
https://creativecommons.org/licenses/by/4.0/
Received: 29 November 2025 | **Revised:** 10 Desember 2025 | **Accepted:** 20 Desember 2025

Menyusun Sistem Informasi untuk Analisis Pertimbangan Preferensi dan Efisiensi Sistem Pembayaran QRIS, Debit, dan Tunai di Masyarakat

Rafael Christoper Suhandi[1], Dedi Trisnawarman[2], Agus Budiyantara[3]
[1]Universitas Tarumanagara, Jakarta, Indonesia, rafaelchrs123@gmail.com
[2]Universitas Tarumanagara, Jakarta, Indonesia, dedit@fti.untar.ac.id
[3]Universitas Tarumanagara, Jakarta, Indonesia, agusb@fti.untar.ac.id

Corresponding Author: rafaelchrs123@gmail.com[1]

Abstract: *This study aims to analyze user perceptions of digital payment systems in Indonesia, particularly QRIS, by combining sentiment analysis and clustering methods. The hybrid model used combines the Naïve Bayes algorithm to analyze public opinion from the Reddit platform and the K-Means Clustering algorithm to group survey respondents based on similar perceptions. Research data was obtained from two sources: secondary data in the form of social media comments and primary data from user surveys. The results of the sentiment analysis indicate that most public opinion is neutral to positive towards the use of QRIS. The integration of sentiment results with survey data produces a combined data set that is more comprehensive in describing user behavior and views. The clustering process produces three main user groups: new users who are still hesitant, adaptive users, and active users with a high level of satisfaction. The results show that the hybrid Naïve Bayes and K-Means approach is effective in revealing user perception patterns and indicates that public acceptance of the QRIS digital payment system is relatively high.*

Keywords: *Sentiment Analysis, Naïve Bayes, K-Means Clustering, QRIS, Digital Payment System.*

Abstrak: Penelitian ini bertujuan untuk menganalisis persepsi pengguna terhadap sistem pembayaran digital di Indonesia, khususnya QRIS, dengan menggabungkan metode analisis sentimen dan klasterisasi. Model hybrid yang digunakan mengombinasikan algoritma Naïve Bayes untuk menganalisis opini publik dari platform Reddit dan algoritma K-Means Clustering untuk mengelompokkan responden survei berdasarkan kesamaan persepsi. Data penelitian diperoleh dari dua sumber, yaitu data sekunder berupa komentar media sosial dan data primer dari survei pengguna. Hasil analisis sentimen menunjukkan mayoritas opini publik bersifat netral hingga positif terhadap penggunaan QRIS. Integrasi hasil sentimen dengan data survei menghasilkan dataset gabungan yang lebih komprehensif dalam menggambarkan perilaku dan pandangan pengguna. Proses klasterisasi menghasilkan tiga kelompok utama pengguna, yaitu pengguna baru yang masih ragu, pengguna adaptif, dan pengguna aktif dengan tingkat kepuasan tinggi. Hasil penelitian menunjukkan bahwa pendekatan hybrid Naïve Bayes dan K-Means

efektif untuk mengungkap pola persepsi pengguna serta menunjukkan bahwa penerimaan masyarakat terhadap sistem pembayaran digital QRIS tergolong tinggi.

Kata Kunci: Analisis Sentimen, Naïve Bayes, K-Means Clustering, QRIS, Sistem Pembayaran Digital.

PENDAHULUAN

Perkembangan teknologi finansial di Indonesia mengalami percepatan pesat dalam beberapa tahun terakhir, terutama dengan hadirnya Quick Response Code Indonesian Standard (QRIS) yang dikeluarkan oleh Bank Indonesia sebagai standar nasional untuk pembayaran berbasis QR. QRIS bertujuan menciptakan transaksi yang cepat, murah, aman, dan andal (CEMUMUAH), serta menjadi solusi interoperabilitas bagi berbagai penyelenggara jasa sistem pembayaran. Sejak diluncurkan pada tahun 2019, penggunaan QRIS terus meningkat dan telah menjadi salah satu inovasi utama dalam mendukung perluasan ekosistem pembayaran digital di Indonesia.

Meskipun tingkat adopsinya tinggi, persepsi dan pengalaman pengguna terhadap QRIS masih menunjukkan variasi. Beberapa pengguna merasakan kemudahan, efisiensi, dan kepraktisan, sementara sebagian lainnya tetap memiliki kekhawatiran terkait biaya transaksi, keamanan data, atau stabilitas layanan. Hal ini menunjukkan perlunya analisis mendalam untuk memahami faktor-faktor yang memengaruhi penerimaan pengguna terhadap sistem pembayaran digital tersebut.

Untuk memperoleh gambaran yang lebih komprehensif, penelitian ini memanfaatkan pendekatan hybrid data mining dengan menggabungkan dua sumber data utama: opini publik di media sosial dan respon survei pengguna. Analisis opini publik dilakukan menggunakan komentar dari platform Reddit, yang kemudian diproses menggunakan algoritma Naïve Bayes untuk mengklasifikasikan sentimen ke dalam kategori positif, netral, dan negatif. Metode Naïve Bayes dipilih karena terbukti efektif dalam analisis teks dan sentimen berbasis probabilistik.

Selain itu, data persepsi pengguna dari survei Google Form dianalisis menggunakan algoritma K-Means Clustering, yaitu metode pengelompokan yang banyak digunakan untuk menemukan pola berdasarkan kemiripan karakteristik responden. Hasil analisis sentimen Reddit kemudian diintegrasikan sebagai variabel tambahan untuk memperkuat segmentasi pengguna, sehingga menghasilkan pemahaman yang lebih menyeluruh mengenai persepsi masyarakat terhadap QRIS.

Dengan pendekatan hybrid ini, penelitian diharapkan mampu memberikan insight yang lebih mendalam mengenai bagaimana opini publik dan perilaku pengguna terbentuk, serta menjadi masukan bagi regulator dan penyedia layanan dalam meningkatkan edukasi dan adopsi QRIS di Indonesia.

METODE

Penelitian ini direncanakan untuk dimulai pada bulan Juli 2025 dan diperkirakan akan berlangsung hingga seluruh tahapan penelitian, mulai dari analisis kebutuhan, perancangan, implementasi, hingga evaluasi sistem, selesai dilaksanakan. Rincian jadwal untuk setiap kegiatan akan dijabarkan lebih lanjut dalam diagram waktu (Gantt Chart) penelitian.

Pelaksanaan penelitian ini akan dilakukan secara fleksibel di beberapa lokasi untuk menunjang setiap tahapan agar berjalan dengan efektif. Lokasi utama pelaksanaan penelitian adalah sebagai berikut: Rumah Peneliti: Digunakan untuk kegiatan yang bersifat mandiri seperti studi literatur, perancangan kuesioner, analisis data, dan penyusunan laporan penelitian., dan Kampus Universitas Tarumanagara: Digunakan untuk kegiatan yang memerlukan fasilitas kampus, bimbingan dengan dosen, diskusi, serta proses pengumpulan data dari responden di lingkungan universitas.

Adapun tahapan metodologi penelitian yang digunakan adalah sebagai berikut:

1. **Pengumpulan Data**

Penelitian ini menggunakan dua jenis data, yaitu:

a) **Data Sekunder**

Data diperoleh dari platform Reddit melalui proses scraping menggunakan kata kunci terkait sistem pembayaran digital. Komentar tersebut digunakan sebagai sumber opini publik untuk analisis sentimen terkait QRIS.

b) **Data Primer**

Data diperoleh melalui survei Google Form yang diisi oleh 102 responden, terdiri dari pengguna QRIS dan pelaku UMKM. Survei ini mengukur persepsi pengguna terkait aspek kepuasan, kemudahan, keamanan, kecepatan, dan manfaat penggunaan QRIS.

2. **Preprocessing Data**

Tahap preprocessing dilakukan untuk menyiapkan data sebelum dianalisis.

a. **Data Reddit:**

a) pembersihan teks (cleaning)
b) case folding
c) tokenizing
d) stopword removal
e) stemming

Tahapan ini memastikan teks siap diproses oleh algoritma Naïve Bayes.

b. **Data Google Form:**

a) pengecekan nilai kosong
b) pembersihan data duplikat
c) konversi tipe data
d) normalisasi skala Likert ke rentang 0–1

Normalisasi dilakukan agar data dapat digunakan secara optimal pada proses clustering.

3. **Analisis Sentimen Menggunakan Naïve Bayes**

Komentar Reddit dianalisis menggunakan algoritma Naïve Bayes untuk mengklasifikasikan sentimen ke dalam tiga kategori:

A. positif
B. netral
C. negatif

Hasil sentimen kemudian dikonversi menjadi skor numerik: positif = 1, netral = 0.5, negatif = 0.Skor ini disimpan sebagai variabel tambahan untuk memperkaya analisis pada tahap integrasi data.

4. **Integrasi Data**

Hasil analisis sentimen dari data komentar Reddit kemudian digabungkan dengan data survei Google Form untuk membentuk satu dataset terpadu yang lebih informatif. Integrasi dilakukan dengan menambahkan skor sentimen—yang sebelumnya dikonversi ke nilai numerik (positif = 1, netral = 0.5, negatif = 0)—ke dalam setiap record responden pada data survei. Dengan cara ini, setiap entri Google Form tidak hanya merepresentasikan persepsi pengguna berdasarkan jawaban kuesioner, tetapi juga diperkaya dengan informasi tambahan berupa kecenderungan opini publik yang bersumber dari media sosial.

Proses integrasi ini menghasilkan dataset hybrid yang mampu menangkap dua perspektif sekaligus, yaitu pandangan individual dari responden survei dan dinamika opini publik secara lebih luas. Pendekatan ini memberikan gambaran yang lebih komprehensif mengenai persepsi dan preferensi pengguna terhadap QRIS, serta meningkatkan kualitas analisis pada tahap klasterisasi berikutnya.

5. **Clustering Menggunakan Algoritma K-Means**

Dataset hasil integrasi dianalisis menggunakan algoritma K-Means Clustering untuk mengelompokkan pengguna berdasarkan kemiripan persepsi.Penentuan jumlah cluster optimal dilakukan menggunakan Elbow Method, yaitu metode yang melihat titik penurunan nilai Within-Cluster Sum of Squares (WCSS) untuk menentukan jumlah cluster yang paling ideal.

6. Perancangan Sistem Informasi Berbasis Web

Tahap akhir penelitian adalah membangun sebuah sistem informasi berbasis web yang berfungsi untuk menyajikan hasil analisis, termasuk:

a) Visualisasi distribusi sentimen
b) Visualisasi hasil clustering (distribusi cluster, nilai rata-rata setiap variabel)
c) Tabel data dari Reddit, Google Form, dan hasil integrasi
d) Ringkasan insight berdasarkan hasil analisis

Sistem web ini dikembangkan menggunakan HTML, CSS, dan JavaScript, serta dirancang sebagai media interaktif untuk memudahkan pengguna dalam memahami pola persepsi dan sentimen terhadap QRIS. Seluruh proses analisis seperti preprocessing, sentiment analysis Naïve Bayes, integrasi data, dan K-Means clustering dilakukan terlebih dahulu menggunakan Python. Website hanya digunakan sebagai media penyajian (presentation layer), bukan untuk menghitung machine learning secara langsung.

HASIL DAN PEMBAHASAN

Hasil Analisis dan Integrasi Data

Bagian ini menyajikan hasil analisis dari setiap tahapan proses pengolahan data, mulai dari analisis sentimen terhadap komentar Reddit menggunakan algoritma Multinomial Naïve Bayes, integrasi dengan data survei Google Form, hingga proses klasterisasi menggunakan algoritma K-Means. Hasil analisis ini digunakan sebagai dasar dalam membangun sistem informasi berbasis web yang menampilkan visualisasi persepsi dan preferensi pengguna terhadap sistem pembayaran QRIS.

Hasil Analisis Sentimen Reddit

Analisis sentimen dilakukan menggunakan algoritma Naïve Bayes dengan representasi vektor TF-IDF (Term Frequency-Inverse Document Frequency). Algoritma ini dipilih karena kesederhanaan komputasi, efisiensi pada data teks berukuran besar, dan performa yang baik untuk klasifikasi multi-kelas pada domain analisis sentimen media sosial.

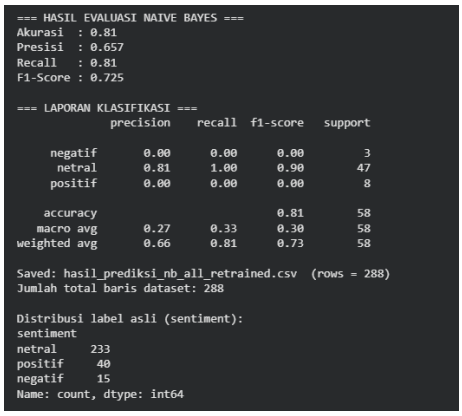
Pembagian Dataset (Train-Test Split)

Dataset Reddit yang telah melalui tahapan preprocessing lengkap (cleaning, case folding, tokenization, stopword removal, stemming, dan transformasi TF-IDF) kemudian dibagi menjadi dua subset:

1) **Data Training (80%)**: Digunakan untuk melatih model Naïve Bayes
   a) Jumlah: ± 238 komentar (80% dari 288 komentar)
   b) Fungsi: Pembelajaran pola bahasa dan probabilitas kemunculan kata untuk setiap kelas sentimen
2) **Data Testing (20%)**: Digunakan untuk evaluasi performa model
   a) Jumlah: ± 60 komentar (20% dari 288 komentar)
   b) Fungsi: Validasi akurasi model pada data yang tidak pernah dilihat sebelumnya (unseen data)

Pembagian dengan rasio 80:20 dipilih karena:

1. Memberikan data training yang cukup untuk pembelajaran model
2. Menyisakan data testing yang cukup untuk evaluasi yang reliable
3. Merupakan praktik standar dalam machine learning untuk dataset ukuran sedang



Gambar 1. Hasil Evaluasi Naïve bayes

Berdasarkan hasil evaluasi pada Gambar 1. model Naïve Bayes menunjukkan akurasi sebesar 0.81. Nilai recall yang lebih tinggi dibandingkan precision menunjukkan bahwa model cenderung mengenali kelas netral secara dominan. Hal ini terjadi karena distribusi data sentimen pada Reddit bersifat imbalanced, dengan 233 komentar netral, 40 positif, dan hanya 15 negatif. Dominasi kelas netral menyebabkan model lebih sensitif terhadap kelas tersebut, yang tercermin dari hasil prediksi model yang seluruhnya menghasilkan label netral.

Sumber Data yang Diintegrasikan

Integrasi data merupakan tahap penggabungan antara data primer yang diperoleh dari Google Form dengan data sekunder yang berasal dari hasil analisis sentimen komentar Reddit. Tujuan dari proses ini adalah menghasilkan satu dataset komprehensif yang memadukan persepsi pengguna berdasarkan survei dengan opini publik dari media sosial sehingga dapat digunakan untuk analisis lanjutan, khususnya proses klasterisasi.

Pada penelitian ini, proses integrasi dilakukan melalui langkah-langkah berikut:

1. **Normalisasi Data Google Form**

   Seluruh variabel persepsi pada Google Form meliputi *kepuasan, kemudahan, keamanan, dan frekuensi penggunaan* dinormalisasi ke dalam rentang 0–1 agar memiliki skala yang seragam. Hal ini penting untuk memastikan bahwa seluruh variabel memiliki kontribusi yang seimbang ketika digunakan dalam K-Means.

2. **Konversi Sentimen Reddit ke Skor Numerik**

   Analisis sentimen menggunakan Naïve Bayes menghasilkan tiga kategori: positif, netral, dan negatif.Agar dapat digabungkan dengan data survei, kategori tersebut dikonversi menjadi skor numerik:

   1. Positif = 1.00
   2. Netral = 0.50
   3. Negatif = 0.00

3. **Penggabungan Dataset**

   Dataset Google Form kemudian digabung dengan skor sentimen Reddit berdasarkan urutan data, menghasilkan satu tabel integrasi yang berisi seluruh fitur yang dibutuhkan untuk proses klasterisasi.

Tabel 6. Hasil integrasi data

No	Usia	Jenis Kelamin	Kepuasan	Kemudahan	Keamanan	Frekuensi Penggunaan	Skor Sentimen Reddit
1	20–30	Laki-laki	0.75	0.75	1.00	0.75	0.50
2	31–40	Perempuan	0.75	1.00	0.75	0.75	0.50

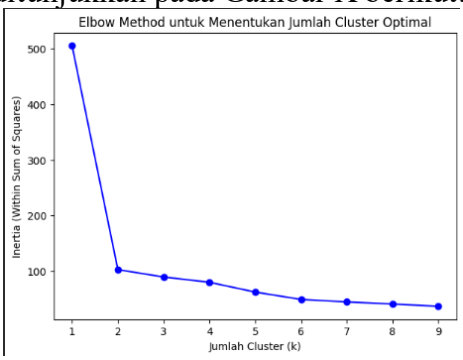
3	20–30	Laki-laki	1.00	0.75	1.00	0.75	0.50
4	31–40	Perempuan	0.75	0.75	0.75	0.75	0.50
5	31–40	Laki-laki	0.75	0.75	0.75	1.00	0.50

Tabel 6 menyajikan hasil integrasi data antara variabel demografis responden dan hasil konversi skala Likert. Tabel ini menggabungkan informasi usia, jenis kelamin, tingkat kepuasan, kemudahan, keamanan, frekuensi penggunaan, serta skor sentimen Reddit untuk setiap responden setelah proses penyelarasan data dilakukan.

Hasil Klasterisasi

A. Penentuan Jumlah Cluster Optimal

Penentuan jumlah cluster dilakukan menggunakan Elbow Method, dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk nilai k = 1 hingga k = 9. Grafik hasil perhitungan ditunjukkan pada Gambar X berikut:



Gambar 2. Grafik Elbow Method untuk Menentukan Jumlah Cluster Optimal

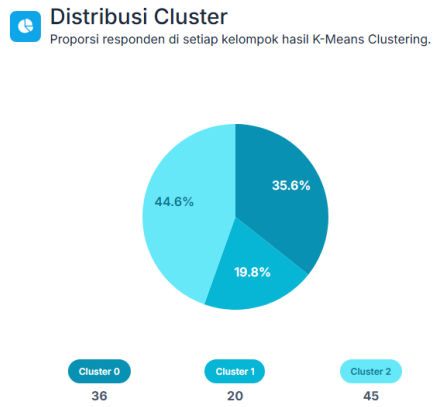
Gambar 2 menunjukan Grafik penurunan WCSS yang signifikan dari k=1 ke k=3, kemudian perubahan menjadi relatif stabil setelahnya. Titik "siku" (elbow) terlihat pada k = 3, sehingga jumlah cluster optimal yang digunakan dalam penelitian ini adalah tiga cluster.

B. Hasil Pembentukan Cluster

Setelah menentukan k=3, proses klasterisasi dilakukan terhadap dataset hasil integrasi (Google Form + Sentimen Reddit). Hasil klasterisasi menghasilkan tiga kelompok dengan ukuran sebagai berikut:

1. Cluster 0: 36 responden
2. Cluster 1: 20 responden
3. Cluster 2: 45 responden

Distribusi cluster dapat dilihat pada Gambar 3. berikut:



Gambar 3. Distribusi Cluster Pengguna Berdasarkan Hasil K-Means

Distribusi di dalam Gambar 3. menunjukkan bahwa masing-masing cluster memiliki jumlah pengguna yang cukup proporsional, sehingga proses klasterisasi berjalan stabil dan tidak terjadi dominasi satu cluster secara berlebihan.

A. Profil dan Karakteristik Tiap Cluster

Analisis dilakukan terhadap nilai rata-rata tiap variabel:

1. Kepuasan
2. Kemudahan
3. Keamanan
4. Frekuesi Penggunaan
5. Skor Sentimen Reddit

Hasil pengamatan profil cluster dapat disajikan sebagai berikut:

Cluster 0 Pengguna Moderat (n = 36)

Cluster ini memiliki skor rata-rata yang cukup seimbang pada semua variabel. Ciri-cirinya:

1. Tingkat kenyamanan dan kemudahan cukup tinggi
2. Penggunaan QRIS cukup sering
3. Skor sentimen Reddit berada pada level netral ke positif

Interpretasi:

Kelompok ini merupakan pengguna yang sudah terbiasa dan merasa cukup puas menggunakan QRIS, namun masih menemukan beberapa aspek yang dapat ditingkatkan.

Cluster 1 Pengguna dengan Persepsi Terendah (n = 20)

Cluster ini memiliki skor rata-rata terendah dalam hampir semua variabel.

Ciri-cirinya:

1. Tingkat kepuasan rendah
2. Keamanan dan kemudahan dinilai paling rendah
3. Penggunaan QRIS jarang
4. Skor sentimen Reddit turut rendah dan cenderung netral-negatif

Interpretasi:

Kelompok ini merupakan pengguna baru atau pengguna yang masih ragu. Persepsi mereka terhadap keamanan dan efisiensi QRIS menjadi hambatan utama dalam penggunaan.

Cluster 2 Pengguna Sangat Puas (n = 45)

Cluster ini memiliki skor rata-rata **paling tinggi** dibanding dua cluster lainnya.

Ciri-cirinya:

1. Kepuasan dan kenyamanan tinggi
2. Penggunaan QRIS lebih sering
3. Persepsi keamanan baik

4. Skor sentimen Reddit yang terintegrasi paling positif

Interpretasi:

Cluster ini menggambarkan kelompok pengguna yang memiliki pengalaman positif dan loyal terhadap penggunaan QRIS.

Implementasi Sistem dan Visualisasi Web

Bagian ini membahas implementasi sistem informasi berbasis web yang dikembangkan sebagai media visualisasi hasil penelitian, mulai dari analisis sentimen, integrasi data, hingga hasil klasterisasi pengguna QRIS. Sistem ini dibangun menggunakan HTML, CSS, dan komponen interface interaktif, sehingga dapat diakses dengan mudah oleh pengguna tanpa instalasi tambahan.

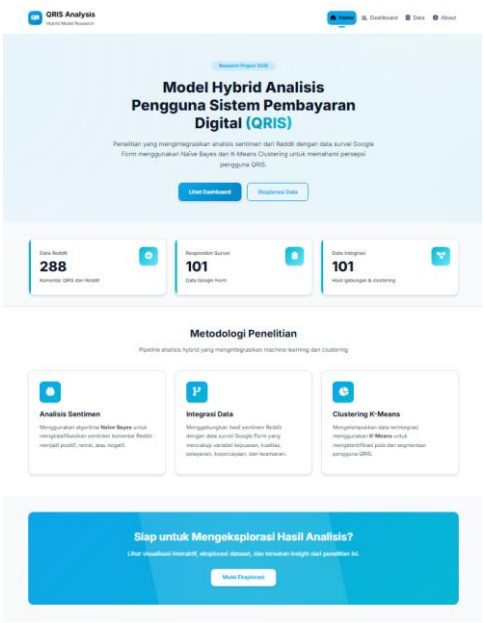
Sistem terdiri dari beberapa halaman utama, yaitu Home, Dashboard, Data Management, dan About. Setiap halaman memiliki peran untuk menyajikan hasil analisis secara terstruktur, informatif, dan mudah dipahami.

Halaman Home

Halaman Home berfungsi sebagai pengantar yang memberikan gambaran umum mengenai penelitian, jumlah dataset, serta alur metodologi yang digunakan.

Konten utama yang ditampilkan:

1. **Header Section**
   a) Judul penelitian dengan typography yang menonjol
   b) Deskripsi singkat mengenai tujuan dan cakupan penelitian
   c) Informasi peneliti dan institusi
2. **Statistik ringkas dataset:**
   a) Ditampilkan dalam bentuk card yang terorganisir dengan ikon representatif:
   b) Jumlah komentar Reddit: ±298 komentar yang telah diproses
   c) Jumlah responden Google Form: 102 responden valid setelah cleaning
   d) Jumlah data terintegrasi: 102 record hasil penggabungan
3. **Jumlah cluster yang terbentuk: 3 kelompok pengguna**
4. **Diagram alur metodologi:**
   a) Tahap 1: Pengumpulan Data - Web scraping Reddit dan survei Google Form
   b) Tahap 2: Preprocessing - Text cleaning untuk Reddit dan normalisasi untuk survei
   c) Tahap 3: Analisis Sentimen - Klasifikasi menggunakan Naïve Bayes
   d) Tahap 4: Integrasi Data - Penggabungan skor sentimen dengan data survei
   e) Tahap 5: Clustering K-Means - Segmentasi pengguna berdasarkan persepsi
   f) Tahap 6: Visualisasi - Penyajian hasil melalui dashboard interaktif



Gambar 4. Tampilan Halaman Home Sistem Informasi Analisis QRIS

Desain dan User Experience:

Seperti yang di tampilkan di **Gambar 4.** Halaman home dirancang dengan prinsip simplicity dan clarity. Penggunaan white space yang cukup memastikan informasi tidak terlihat padat dan mudah dipahami. Color scheme yang dipilih menggunakan kombinasi biru dan putih untuk memberikan kesan profesional dan modern. Setiap card statistik dilengkapi dengan animasi hover yang memberikan feedback visual kepada pengguna.

A. Halaman Dashboard

Dashboard merupakan bagian inti sistem yang menampilkan rangkuman hasil analisis dalam bentuk angka ringkas, grafik, dan visualisasi interaktif.
Tampilan dashboard pada sistem menampilkan beberapa komponen utama:
Ringkasan Statistik

1. **Total Data Integrasi Card**
   a) Menampilkan jumlah total data yang berhasil diintegrasikan
   b) Indikator persentase keberhasilan integrasi
   c) Icon: Database symbol
2. **Jumlah Cluster Card**
   a) Menampilkan jumlah cluster optimal hasil Elbow Method
   b) Informasi distribusi anggota per cluster
   c) Icon: Pie chart symbol
3. **Rata-rata Skor Sentimen Card**
   a) Nilai rata-rata sentimen dari seluruh komentar Reddit
   b) Range skala: 0 (negatif) hingga 1 (positif)
   c) Color indicator: merah (rendah), kuning (sedang), hijau (tinggi)
   d) Icon: Emoticon symbol
4. **Rata-rata Skor Persepsi Card**
   a) Nilai rata-rata persepsi dari data Google Form
   b) Mencakup kepuasan, kemudahan, keamanan, dan frekuensi
   c) Progress bar visual untuk setiap komponen persepsi
   d) Icon: Star rating symbol

Informasi Cluster Dominan

Section ini menampilkan informasi mengenai cluster dengan karakteristik paling menonjol:

1. Cluster dengan Jumlah Anggota Terbanyak: Cluster 2 dengan 45 responden (44.6%)
2. Karakteristik Dominan: Pengguna dengan persepsi tinggi dan kepuasan baik
3. Rekomendasi: Kelompok ini dapat dijadikan target untuk program loyalitas

Grafik Distribusi Cluster

Visualisasi berbentuk pie chart interaktif yang menunjukkan:

1. Proporsi anggota setiap cluster dalam persentase
2. Color coding yang konsisten untuk setiap cluster:
   a) Cluster 0 (Persepsi Moderat): Warna biru
   b) Cluster 1 (Persepsi Rendah): Warna merah
   c) Cluster 2 (Persepsi Tinggi): Warna hijau
3. Fitur hover yang menampilkan detail jumlah dan persentase
4. Legend yang dapat diklik untuk toggle visibility

Interpretasi Distribusi:

1. Cluster 2 mendominasi dengan 44.6%, menunjukkan mayoritas pengguna memiliki pengalaman positif
2. Cluster 1 hanya 19.8%, mengindikasikan sedikit pengguna dengan persepsi negatif
3. Distribusi yang relatif merata menunjukkan segmentasi yang baik

Profil Skor per Cluster

Bagian ini menampilkan perbandingan rata-rata skor untuk setiap cluster dalam bentuk grouped bar chart:
Variabel yang Dibandingkan:

1. **Skor Sentimen Reddit** (0-1)
   a) Cluster 0: 0.50 (netral)
   b) Cluster 1: 0.48 (netral-negatif)
   c) Cluster 2: 0.52 (netral-positif)
2. **Skor Persepsi Total** (0-1)
   a) Cluster 0: 0.68 (moderat)
   b) Cluster 1: 0.42 (rendah)
   c) Cluster 2: 0.85 (tinggi)
3. **Skor Total Integrasi** (rata-rata dari sentimen + persepsi)
   a) Cluster 0: 0.59
   b) Cluster 1: 0.45
   c) Cluster 2: 0.69

Plot Clustering

Visualisasi scatter plot dua dimensi yang menggambarkan:
Sumbu X: Skor Sentimen Reddit (0-1) Sumbu Y: Skor Total Persepsi (0-1)
Karakteristik Visualisasi:

1. Setiap titik merepresentasikan satu responden
2. Warna titik sesuai dengan cluster assignment
3. Size titik dapat disesuaikan berdasarkan frekuensi penggunaan
4. Centroid cluster ditampilkan dengan marker khusus (bintang)
5. Grid lines untuk memudahkan pembacaan koordinat
6. Tooltip interaktif menampilkan detail responden saat hover

Pola yang Teridentifikasi:

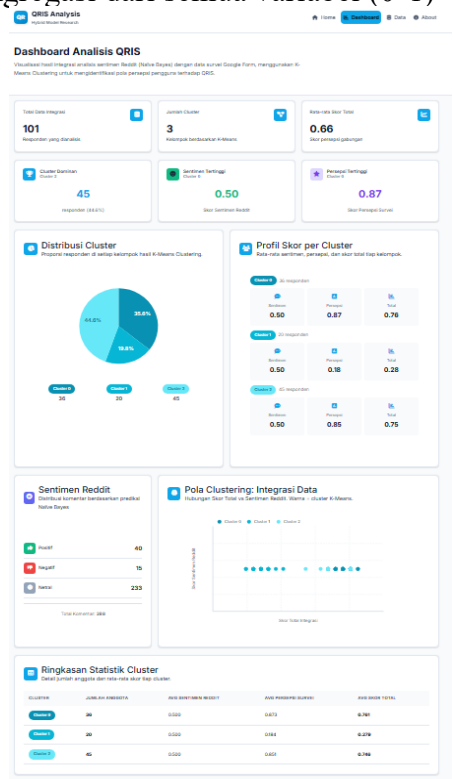
1. Cluster 2 terkonsentrasi di kuadran kanan atas (high sentiment, high perception)
2. Cluster 1 tersebar di kuadran kiri bawah (low sentiment, low perception)
3. Cluster 0 berada di area tengah dengan variasi moderate
4. Terdapat boundary yang relatif jelas antar cluster

Scatter plot yang menggambarkan pemetaan pengguna berdasarkan skor sentimen dan skor total integrasi.

Ringkasan Statistik Cluster

Tabel komprehensif yang menampilkan rata-rata setiap variabel untuk tiap cluster:
Kolom Tabel:

1. **Cluster ID**: Nomor identifikasi cluster (0, 1, 2)
2. **Jumlah Anggota**: Total responden dalam cluster
3. **Kepuasan**: Rata-rata skor kepuasan (0-1)
4. **Kemudahan**: Rata-rata skor kemudahan penggunaan (0-1)
5. **Keamanan**: Rata-rata skor persepsi keamanan (0-1)
6. **Frekuensi Penggunaan**: Rata-rata intensitas penggunaan QRIS (0-1)
7. **Skor Sentimen**: Rata-rata sentimen dari Reddit (0-1)
8. **Skor Total**: Agregasi dari semua variabel (0-1)



Gambar 5.Tampilan Dashboard Analisis QRIS

Gambar 5. menampilkan dashboard analisis QRIS yang digunakan sebagai media visualisasi dalam penelitian ini. Dashboard tersebut memuat beberapa komponen analisis, seperti overview data, persebaran cluster, skor tiap cluster, analisis sentimen Reddit, pola integrasi data, serta ringkasan statistik cluster, sehingga mempermudah proses evaluasi dan interpretasi hasil clustering.

B. Halaman Data Management

Halaman Data Management berfungsi untuk menampilkan dataset penelitian secara terstruktur, yang terdiri atas tiga kategori yaitu:

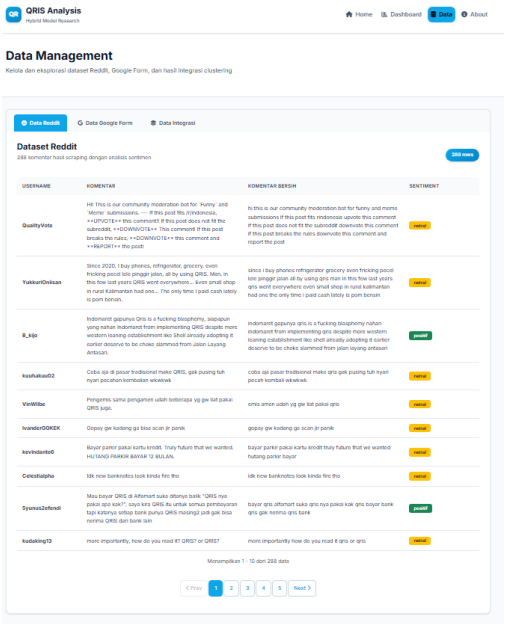
1. Data Reddit

Berisi komentar Reddit yang telah melalui proses:

a) Cleaning
b) Preprocessing
c) Prediksi sentimen dengan Naïve Bayes

Ditampilkan dalam tabel dengan kolom:

a) Username
b) Komentar mentah
c) Komentar bersih
d) Sentimen



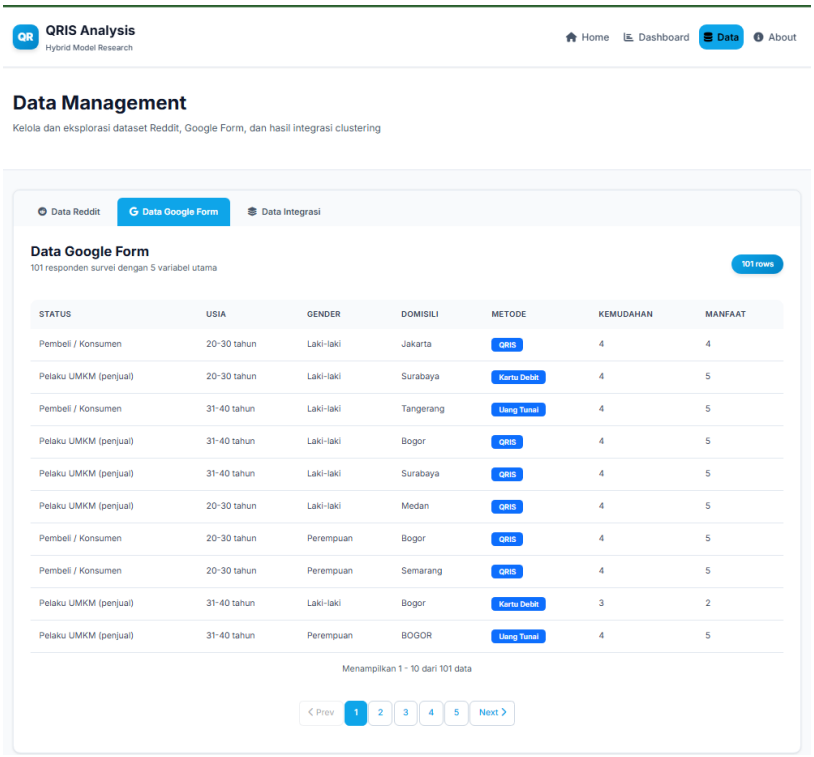
Gambar 6. Tampilan Dataset Reddit

Pada Gambar 6. dapat dilihat tampilan halaman *Dataset Reddit* yang digunakan dalam penelitian ini. Halaman tersebut menampilkan data komentar dari pengguna Reddit secara terstruktur, mulai dari username, isi komentar, hingga label sentimen yang telah dihasilkan oleh sistem. Antarmuka ini berfungsi sebagai pusat pengelolaan data untuk memastikan bahwa setiap entri dapat diperiksa, divalidasi, serta dianalisis dengan mudah sesuai kebutuhan proses pengolahan data selanjutnya.

2. Data Google Form

Berisi data persepsi pengguna yang diperoleh dari 101 responden.
Kolom yang ditampilkan seperti di **Gambar 11.** :

a) Status pengguna (konsumen/UMKM)
b) Usia
c) Jenis kelamin
d) Domisili
e) Metode pembayaran paling sering dipakai
f) Kemudahan
g) Keamanan
h) Kenyamanan

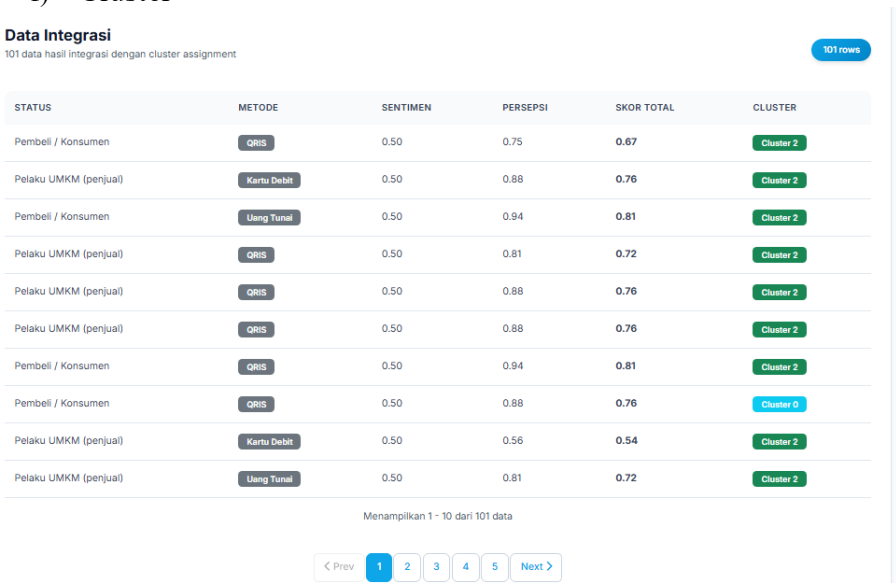


Gambar 7. Tampilan Dataset Google Form

3. Data Terintegrasi

Berisi hasil integrasi antara data Google Form dan skor sentimen Reddit. Ditampilkan dalam tabel yang berisi:

a) Status
b) Metode
c) Sentimen
d) Persepsi
e) Skor total
f) Cluster



Gambar 8. Tampilan Dataset Terintegrasi

Pada Gambar 8. ditunjukkan tampilan *Dataset Terintegrasi* yang merupakan hasil penggabungan data dari berbagai sumber dengan penambahan nilai sentimen, persepsi, skor total, serta label cluster yang dihasilkan dari proses pengelompokan. Setiap baris data

menampilkan informasi mengenai status responden, metode pembayaran yang digunakan, hingga hasil skor dan cluster akhir yang menjadi dasar analisis lebih lanjut. Tampilan ini memudahkan peneliti dalam memahami struktur data yang telah diproses sekaligus melihat distribusi responden pada masing-masing cluster.

KESIMPULAN

Kesimpulan merupakan rangkuman temuan utama yang diperoleh dari seluruh rangkaian penelitian, mulai dari pengumpulan data, preprocessing, analisis sentimen, integrasi data, klasterisasi, hingga implementasi sistem informasi berbasis web. Berdasarkan hasil penelitian yang telah dilakukan, diperoleh beberapa kesimpulan sebagai berikut. Analisis sentimen terhadap opini publik Reddit berhasil dilakukan menggunakan algoritma Naïve Bayes dengan representasi fitur TF-IDF, menghasilkan akurasi sebesar 0,81. Distribusi dataset yang imbalanced membuat model lebih sensitif terhadap kelas netral sehingga mayoritas komentar diklasifikasikan sebagai netral. Meskipun demikian, skor sentimen tetap dapat digunakan sebagai variabel tambahan pada proses integrasi data. Data survei Google Form yang mencakup 102 responden berhasil diproses, dinormalisasi, dan dipetakan ke dalam lima variabel persepsi, yaitu kepuasan, kemudahan, kecepatan, keamanan, dan frekuensi penggunaan. Setelah normalisasi skala 0–1, data menjadi siap untuk diintegrasikan dengan skor sentimen Reddit. Integrasi data antara skor sentimen Reddit dan data persepsi Google Form menghasilkan satu dataset komprehensif yang merepresentasikan pengalaman sekaligus persepsi pengguna. Dataset terintegrasi ini digunakan sebagai input utama dalam proses klasterisasi. Proses klasterisasi K-Means menunjukkan bahwa jumlah cluster optimal adalah tiga (k = 3) berdasarkan metode Elbow. Ketiga cluster merepresentasikan tiga kelompok karakteristik pengguna yang berbeda, yaitu:

Cluster 0: pengguna dengan persepsi sedang dan sentimen netral.
Cluster 1: pengguna dengan persepsi rendah.
Cluster 2: pengguna dengan persepsi tinggi dan tingkat kepuasan lebih baik.

Sistem informasi berbasis web yang dikembangkan berhasil menyajikan hasil analisis dalam bentuk dashboard interaktif yang mencakup visualisasi cluster, profil skor tiap cluster, pola persepsi pengguna, ringkasan statistik, serta halaman manajemen data. Sistem ini memberikan media eksplorasi yang efektif untuk memahami persepsi pengguna terhadap QRIS secara menyeluruh. Secara keseluruhan, penelitian ini berhasil mengintegrasikan analisis sentimen dan klasterisasi menggunakan pendekatan hybrid serta mengimplementasikannya dalam bentuk sistem informasi berbasis web yang informatif.

REFERENSI

A. N. Marbun, Y. S. J. Nasution, and I. Harahap, “Analisis Penggunaan Sistem Pembayaran QRIS terhadap Perkembangan Bisnis UMKM melalui Kepuasan Nasabah sebagai Variabel Intervening (Studi Kasus Bank Syariah Indonesia (BSI) Kc Sibolga),” *AHKAM*, vol. 2, no. 2, pp. 361–372, Jun. 2023, doi: 10.58578/AHKAM.V2I2.1198.

A. N. Illah and I. Ilham, “Analisis Risiko, Keamanan, dan Efektivitas Framework pada Layanan Pembayaran Digital Menggunakan Systematic Literature Review,” *J. Teknol. Sist. Inf.*, vol. 6, no. 1, pp. 89–102, Apr. 2025, doi: 10.35957/JTSI.V6I1.9860.

A. R. Maulidah *et al.*, “Perkembangan Sistem Pembayaran Digital : Pada Era Revolusi Industri 4.0 Di Indonesia,” *J. Ekon. dan Bisnis Digit.* , vol. 1, no. 4, pp. 798–803, Jun. 2024, Accessed: Nov. 19, 2025. [Online]. Available: https://jurnal.ittc.web.id/index.php/jebd/article/view/991

A. Z. Rizquina and C. I. Ratnasari, “Implementasi Web Scraping untuk Pengambilan Data Pada Website E-Commerce,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 4, pp. 377–383, Oct. 2023, doi: 10.47233/JTEKSIS.V5I4.913.

E. Sholihah and R. Nurhapsari, “Percepatan Implementasi Digital Payment Pada UMKM:

Intensi Pengguna QRIS Berdasarkan Technology Acceptance Model," vol. 12, no. 1, pp. 1–12, 2023, doi: 10.21831/nominal.v12i1.52480.

H. D. Listyaningrum, "Social Media Digital Footprints for Strengthening Corporate Forensic Auditing Practices," *Kompak J. Ilm. Komputerisasi Akunt.* , vol. 18, no. 2, pp. 507–517, Oct. 2025, doi: 10.51903/KOMPAK.V18I2.3130.

I. Standard Sebagai Alat Transaksi Pembayaran Raja Sakti Putra Harahap, A. Afandi, M. Lubis, L. Indriani, S. H. Tinggi Agama Islam Syekh Abdul Halim Hasan Al-ishlahiyah Binjai, and U. Syekh Ali Hasan Ahmad Addary Padangsidimpuan, "Determinan Preferensi Mahasiswa Dalam Menggunakan Quick Response Code Indonesian Standard (Qris) Sebagai Alat Transaksi Pembayaran," *Innov. J. Soc. Sci. Res.*, vol. 3, no. 1, pp. 312–319, Mar. 2023, Accessed: Nov. 19, 2025. [Online]. Available: http://j-innovative.org/index.php/Innovative/article/view/2396

M. Muslimawati, "Analisis Penggunaan Aplikasi Qris Sebagai Alat Pembayaran Non Tunai Untuk Mempermudah Transaksi Bagi Pelaku Usaha UMKM di Kecamatan Abepura, Kota Jayapura," *J. Manag. Bus.*, vol. 2, no. 1, pp. 185–196, Jun. 2024, doi: 10.61912/LAJUMEN.V2I1.43.

M. U. Albab, Y. K. P., and M. N. Fawaiq, "Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic," *J. Transform.*, vol. 20, no. 2, pp. 1–12, Jan. 2023, doi: 10.26623/TRANSFORMATIKA.V20I2.5374.

M. Z. Haq, C. S. Octiva, A. Ayuliana, U. W. Nuryanto, and D. Suryadi, "Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial," *J. Minfo Polgan*, vol. 13, no. 1, pp. 1079–1084, Jul. 2024, doi: 10.33395/JMP.V13I1.13937.

N. Riyanah, S. Informasi, S. Tinggi, M. Informatika, D. Komputer, and N. Mandiri, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 206–213, Feb. 2021, doi: 10.35746/JTIM.V2I4.117.

P. Baumann, "The OGC web coverage processing service (WCPS) standard," *GeoInformatica 2009 144*, vol. 14, no. 4, pp. 447–479, Jul. 2009, doi: 10.1007/S10707-009-0087-2.

R. Handoyo, R. Mangkudjaja, and S. M. Nasution, "Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K - Means pada Pengelompokan Dokumen," *J. Sifo Mikroskil*, vol. 15, no. 2, pp. 73–82, Oct. 2014, doi: 10.55601/JSM.V15I2.161.

S. K. Dirjen *et al.*, "Terakreditasi SINTA Peringkat 4 Klasterisasi Karakter Konsumen Terhadap Kecenderungan Pemilihan Produk Menggunakan K-Means," vol. 3, no. 1, pp. 195–202, 2018.

R. W. Abiba and R. Indrarini, "Pengaruh Penggunaan Uang Elektronik (E-Money) Berbasis Server Sebagai Alat Transaksi terhadap Penciptaan Gerakan Less Cash Society Pada Generasi Milenial di Surabaya," *J. Ekon. dan Bisnis Islam*, vol. 4, no. 1, pp. 196–206, Apr. 2021, doi: 10.26740/JEKOBI.V4N1.P196-206.