

## Sentiment Analysis and Identification of Trust and Perceived Risk Aspects in User Reviews of the Dana Application Using Text Mining Methods and Support Vector Machine

Wardatul Jannah<sup>1\*)</sup>, Ulfah Kusnia<sup>2)</sup>, Arif Faizin<sup>3)</sup>

<sup>1)2)3)</sup>Teknik Informatika, Fakultas Teknik, Universitas Yudharta Pasuruan

<sup>\*)</sup>Correspondence author: [wardajannah766@gmail.com](mailto:wardajannah766@gmail.com), Kab. Pasuruan, Indonesia

DOI: <https://doi.org/10.37012/jtik.v12i2.3907>

### Abstract

*This research is motivated by the importance of understanding the trust and risk felt by users in using the DANA application based on user reviews. This research aims to analyze sentiment and identify the aspects of Trust and Perceived Risk in DANA application reviews using the Text Mining method and the Support Vector Machine (SVM) algorithm. The research data consists of 300 user reviews obtained from Kaggle. The research stages include pre-processing, sentiment labeling, label validation, word weighting using TF-IDF, and SVM classification. After eliminating neutral data, 280 data consisting of positive and negative sentiments were obtained, with 224 training data and 56 test data. The test results showed that SVM obtained accuracy, precision, recall, and F1-Score of 98% each. Identification of aspects using keyword dictionaries that have been validated by experts shows that Assurance is the most commonly found aspect of Trust, while Transaction Risk is the most dominant aspect of Perceived Risk. The results of the study recommend that developers maintain the Assurance aspect and improve the security and reliability of transactions to reduce the risks felt by users.*

**Keywords:** DANA, Trust, Perceived Risk, Text Mining, SVM.

### Abstrak

Penelitian ini dilatarbelakangi oleh pentingnya memahami kepercayaan dan risiko yang dirasakan pengguna dalam menggunakan aplikasi DANA berdasarkan ulasan pengguna. Penelitian ini bertujuan untuk menganalisis sentimen serta mengidentifikasi aspek Trust dan Perceived Risk pada ulasan aplikasi DANA menggunakan metode Text Mining dan algoritma Support Vector Machine (SVM). Data penelitian terdiri dari 300 ulasan pengguna yang diperoleh dari Kaggle. Tahapan penelitian meliputi pre-processing, pelabelan sentimen, validasi label, pembobotan kata menggunakan TF-IDF, dan klasifikasi SVM. Setelah menghilangkan data netral, diperoleh 280 data yang terdiri atas sentimen positif dan negatif, dengan 224 data latih dan 56 data uji. Hasil pengujian menunjukkan bahwa SVM memperoleh akurasi, presisi, recall, dan F1-Score masing-masing sebesar 98%. Identifikasi aspek menggunakan kamus kata kunci yang telah divalidasi oleh ahli menunjukkan bahwa Assurance merupakan aspek Trust yang paling banyak ditemukan, sedangkan Transaction Risk menjadi aspek Perceived Risk yang paling dominan. Hasil penelitian merekomendasikan kepada pengembang untuk mempertahankan aspek Assurance serta meningkatkan keamanan dan keandalan transaksi guna mengurangi risiko yang dirasakan pengguna.

**Kata Kunci:** DANA, Trust, Perceived Risk, Text Mining, SVM.

---

## PENDAHULUAN

Perkembangan teknologi digital telah mendorong perubahan dalam aktivitas ekonomi dan keuangan masyarakat, termasuk pergeseran dari transaksi tunai menuju pembayaran digital. Secara global, penggunaan layanan pembayaran digital terus meningkat seiring berkembangnya akses terhadap layanan keuangan berbasis teknologi. World Bank melaporkan bahwa sekitar dua pertiga orang dewasa di dunia telah melakukan atau menerima pembayaran digital, menunjukkan bahwa pembayaran digital telah menjadi bagian penting dalam aktivitas keuangan masyarakat (Demirg et al., 2021). Perkembangan tersebut juga didukung oleh kematangan digital wallet yang menawarkan kemudahan, kecepatan, dan fleksibilitas dalam melakukan transaksi. Pada tahun 2023, nilai transaksi melalui digital wallet secara global mencapai sekitar US\$13,9 triliun dan diproyeksikan terus meningkat pada tahun-tahun berikutnya (Worldpay, 2024). Meningkatnya penggunaan layanan tersebut menjadikan aspek keamanan dan kepercayaan semakin penting karena pengguna tidak hanya mempertimbangkan kemudahan, tetapi juga keamanan dana, keandalan sistem, dan risiko ketika melakukan transaksi digital.

Di Indonesia, perkembangan pembayaran digital juga menunjukkan pertumbuhan yang signifikan. Bank Indonesia mencatat bahwa transaksi pembayaran digital pada tahun 2024 mencapai 34,5 miliar transaksi atau tumbuh 36,1% dibandingkan tahun sebelumnya. Pada periode yang sama, transaksi melalui aplikasi mobile tumbuh sebesar 39,1%, sedangkan transaksi pembayaran melalui QRIS meningkat sebesar 175,2% (Indonesia, 2025). Pertumbuhan tersebut menunjukkan semakin luasnya penerimaan masyarakat terhadap layanan pembayaran digital dalam berbagai aktivitas ekonomi. Kondisi ini sekaligus meningkatkan tuntutan terhadap penyedia layanan untuk menjaga keamanan, keandalan sistem, dan kualitas pengalaman pengguna. Ketika layanan digital digunakan untuk menyimpan uang dan melakukan transaksi, gangguan sistem, kegagalan transaksi, maupun kekhawatiran terhadap keamanan dapat memengaruhi tingkat kepercayaan pengguna terhadap layanan yang digunakan.

Salah satu layanan pembayaran digital yang berkembang di Indonesia adalah DANA. DANA menyediakan berbagai layanan transaksi digital, seperti pembayaran, transfer, pembelian, dan layanan keuangan lainnya. Hingga April 2024, DANA menyatakan telah mencapai 180 juta pengguna, dengan sebagian besar penggunanya berada di wilayah kota

---

tingkat 3 dan 4 di Indonesia (DANA Indonesia, 2024). Besarnya jumlah pengguna tersebut menghasilkan volume interaksi dan pengalaman penggunaan yang beragam. Pengalaman tersebut tercermin dalam ulasan pengguna pada platform distribusi aplikasi yang berisi penilaian, apresiasi, keluhan, maupun pengalaman ketika menggunakan layanan. Ulasan pengguna dapat menjadi sumber informasi yang penting karena memuat tanggapan langsung terhadap kualitas layanan, termasuk pengalaman positif maupun permasalahan yang dialami pengguna.

Namun, jumlah ulasan yang besar dan karakteristik data teks yang tidak terstruktur menyebabkan proses identifikasi opini secara manual menjadi kurang efektif. Oleh karena itu, diperlukan pendekatan yang mampu mengolah data teks secara sistematis. Text Mining dapat digunakan untuk memperoleh informasi dari kumpulan data teks, sedangkan analisis sentimen dapat digunakan untuk mengidentifikasi kecenderungan opini pengguna. Dalam proses klasifikasi, TF-IDF dapat digunakan untuk memberikan bobot berdasarkan tingkat kepentingan suatu kata dalam dokumen, kemudian Support Vector Machine (SVM) dapat digunakan untuk mengklasifikasikan data berdasarkan karakteristik yang telah direpresentasikan (Alpaz et al., 2022). Pendekatan tersebut telah digunakan dalam berbagai penelitian untuk menganalisis ulasan aplikasi digital.

Penelitian terdahulu menunjukkan bahwa SVM memiliki kemampuan yang baik dalam mengklasifikasikan sentimen pada ulasan aplikasi DANA. (Alva & Zuliarso, 2025) Melaporkan akurasi sebesar 88,5% dalam klasifikasi sentimen ulasan aplikasi DANA. Penelitian (Fajar Nawulansih et al., 2025) Menggunakan 1.000 ulasan DANA dan memperoleh akurasi sebesar 90% setelah menerapkan SMOTE dan optimasi parameter, meningkat dari 84% sebelum penerapan teknik tersebut. Hasil tersebut menunjukkan bahwa SVM dapat digunakan untuk mengidentifikasi sentimen pengguna, tetapi penelitian yang berfokus pada klasifikasi sentimen pada umumnya masih menitikberatkan pada pengelompokan opini positif dan negatif serta pengukuran performa model. Dengan demikian, informasi yang lebih spesifik mengenai alasan munculnya sentimen tersebut masih dapat dikembangkan.

Ulasan pengguna sebenarnya tidak hanya menunjukkan apakah pengalaman pengguna bersifat positif atau negatif, tetapi juga dapat mengungkap aspek yang menjadi perhatian pengguna ketika menggunakan layanan digital. Dalam konteks aplikasi keuangan, aspek

---

Trust dan Perceived Risk menjadi penting karena berkaitan dengan keyakinan pengguna terhadap keamanan dan kemampuan layanan serta risiko yang dirasakan selama melakukan transaksi. Ulasan mengenai keamanan akun, perlindungan data, keberhasilan transaksi, kegagalan pembayaran, maupun gangguan sistem dapat memberikan gambaran mengenai tingkat kepercayaan dan risiko yang dirasakan pengguna. Oleh karena itu, analisis ulasan yang hanya menghasilkan klasifikasi sentimen belum sepenuhnya menggambarkan faktor yang mendasari pengalaman pengguna.

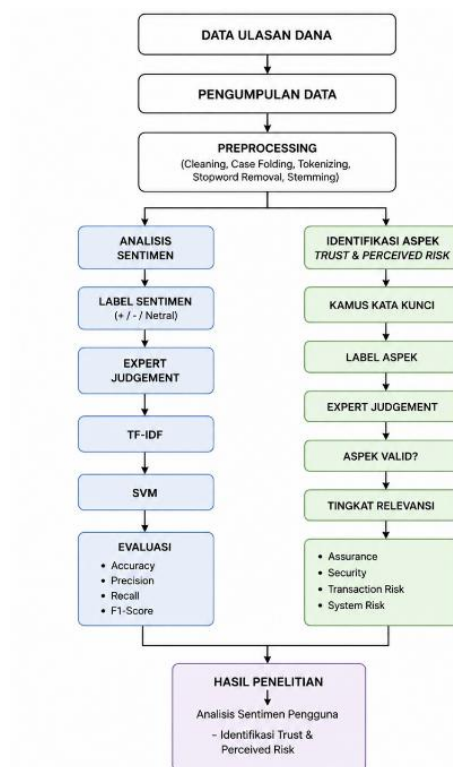
Berdasarkan kesenjangan tersebut, penelitian ini mengembangkan analisis ulasan aplikasi DANA dengan menggabungkan klasifikasi sentimen dan identifikasi aspek Trust serta Perceived Risk. Aspek Trust dalam penelitian ini terdiri atas Assurance dan Security, sedangkan Perceived Risk terdiri atas Transaction Risk dan System Risk. Identifikasi aspek dilakukan menggunakan kamus kata kunci yang disusun berdasarkan kategori penelitian dan divalidasi oleh ahli agar kata kunci yang digunakan sesuai dengan konteks masing-masing aspek. Pendekatan ini diharapkan dapat memberikan informasi yang lebih terarah mengenai aspek kepercayaan dan risiko yang muncul dalam pengalaman pengguna, sekaligus menunjukkan arah sentimen yang menyertai aspek tersebut.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi DANA menggunakan Text Mining, pembobotan TF-IDF, dan algoritma SVM, serta mengevaluasi kinerja model berdasarkan Accuracy, Precision, Recall, dan F1-Score. Selain itu, penelitian ini bertujuan mengidentifikasi aspek Trust yang meliputi Assurance dan Security, serta aspek Perceived Risk yang meliputi Transaction Risk dan System Risk pada ulasan pengguna aplikasi DANA. Hasil penelitian diharapkan dapat memberikan pemahaman yang lebih spesifik mengenai kepercayaan dan risiko yang dirasakan pengguna serta menjadi masukan bagi pengembangan kualitas, keamanan, dan keandalan layanan aplikasi DANA.

## **METODE PENELITIAN**

Penelitian ini menggunakan pendekatan kuantitatif dengan metode Text Mining untuk menganalisis sentimen pengguna serta mengidentifikasi aspek Trust dan Perceived Risk pada ulasan aplikasi DANA. Data yang digunakan merupakan data sekunder berupa ulasan pengguna aplikasi DANA yang diperoleh dari dataset Kaggle. Analisis dilakukan melalui

beberapa tahapan, yaitu pengumpulan dan pelabelan data, pre-processing, pembagian dataset, pembobotan TF-IDF, klasifikasi menggunakan Support Vector Machine (SVM), evaluasi model, serta identifikasi aspek Trust dan Perceived Risk. Alur penelitian ditunjukkan pada Gambar 1.



**Gambar 1.** Alur Penelitian

### Pengumpulan dan Pelabelan Data

Data penelitian terdiri atas 300 ulasan pengguna aplikasi DANA yang diperoleh dari dataset Kaggle. Ulasan tersebut digunakan sebagai populasi data penelitian. Sampel penelitian ditentukan menggunakan teknik purposive sampling, yaitu pemilihan data berdasarkan kriteria yang sesuai dengan tujuan penelitian. Kriteria data yang digunakan adalah ulasan yang berkaitan dengan pengalaman pengguna terhadap aplikasi DANA dan dapat dikategorikan ke dalam sentimen positif, negatif, atau netral.

Setiap ulasan diberi label sentimen positif, negatif, atau netral berdasarkan isi dan makna ulasan. Proses pelabelan dilakukan dengan bantuan expert judgment untuk memastikan kesesuaian antara isi ulasan dan label sentimen. Selanjutnya, ulasan dengan sentimen netral

tidak digunakan dalam proses klasifikasi biner. Dari 300 ulasan awal, diperoleh 280 ulasan yang terdiri atas sentimen positif dan negatif untuk digunakan dalam proses klasifikasi.

### Pre-processing Data

Tahap pre-processing dilakukan untuk membersihkan dan menyeragamkan data teks sebelum dilakukan pembobotan dan klasifikasi. Tahapan ini bertujuan mengurangi unsur yang tidak diperlukan sehingga data menjadi lebih terstruktur dan dapat diproses oleh algoritma klasifikasi. Tahapan pre-processing yang digunakan meliputi cleaning, case folding, tokenizing, stopwords removal, dan stemming (Hakim, 2021).

**Tabel 1.**

**Tahapan Pre-processing**

| Tahapan          | Fungsi                                                           |
|------------------|------------------------------------------------------------------|
| Cleaning         | Menghapus karakter atau bagian teks yang tidak perlu             |
| Case Folding     | Proses mengubah semua huruf menjadi huruf kecil                  |
| Tokenizing       | Memecah teks menjadi kata/token                                  |
| Stopword Removal | Menghilangkan kata-kata yang tidak memberikan informasi tambahan |
| Stemming         | Mengubah kata berimbuhan menjadi bentuk dasar                    |

### Pembagian Dataset

Sebanyak 280 data yang telah melalui proses pelabelan digunakan dalam klasifikasi sentimen. Dataset dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Sebanyak 224 data digunakan sebagai data latih untuk membangun model SVM, sedangkan 56 data digunakan sebagai data uji untuk mengukur kemampuan model dalam mengklasifikasikan data yang belum digunakan pada proses pelatihan.

**Tabel 2.**

**Pembagian Dataset**

| Jenis Data      | Persentase  | Jumlah     | Fungsi                  |
|-----------------|-------------|------------|-------------------------|
| <i>Training</i> | 80%         | 224        | Membangun model SVM     |
| <i>Testing</i>  | 20%         | 56         | Menguji performa model  |
| <b>Total</b>    | <b>100%</b> | <b>280</b> | <b>Data klasifikasi</b> |

### Pembobotan TF-IDF

Setelah melalui tahap pre-processing, data teks diubah menjadi representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot terhadap setiap kata berdasarkan tingkat kepentingannya dalam suatu dokumen dibandingkan dengan keseluruhan dokumen. TF

menunjukkan frekuensi kemunculan suatu kata dalam dokumen, sedangkan IDF menunjukkan tingkat kekhasan kata tersebut dalam keseluruhan dokumen.

Nilai TF-IDF diperoleh dengan mengalikan nilai Term Frequency (TF) dan Inverse Document Frequency (IDF). Hasil pembobotan tersebut kemudian digunakan sebagai fitur masukan bagi algoritma SVM dalam proses klasifikasi. Penggunaan TF-IDF dan SVM untuk klasifikasi teks telah diterapkan dalam penelitian sebelumnya (Alpaz et al., 2022; Sri Rahayu et al., 2023).

Rumus Term Frequency (TF) dapat dituliskan sebagai berikut:

$$TF(t,d) = \left( \frac{f(t,d)}{\sum t' f(t',d)} \right)$$

Keterangan:

- $TF(t,d)$  = frekuensi kemunculan kata  $t$  dalam dokumen  $d$
- $f(t,d)$  = jumlah kemunculan kata dalam dokumen
- $\sum t' f(t',d)$  = jumlah seluruh kata dalam dokumen

### Gambar 2. Rumus Term Frequency

Nilai Inverse Document Frequency (IDF) dihitung untuk mengetahui tingkat kepentingan suatu kata berdasarkan jumlah dokumen yang mengandung kata tersebut. Selanjutnya, nilai TF dan IDF dikalikan untuk memperoleh bobot TF-IDF setiap kata.

$$IDF(t,D) = \left( \frac{N}{df(t)} \right)$$

Keterangan:

- $IDF(t,D)$  = nilai Inverse Document Frequency dari kata  $t$
- $N$  = jumlah keseluruhan dokumen dalam korpus
- $df(t)$  = jumlah dokumen yang memuat kata  $t$

$$TF-IDF(t,d,D) = TF(t,d) \times IDF(t,D)$$

Keterangan:

- $TF-IDF(t,d,D)$  = bobot kata pada dokumen
- $TF(t,d)$  = nilai *Term Frequency*
- $IDF(t,D)$  = nilai *Inverse Document Frequency*

### Gambar 3. Rumus Inverse Document Frequency dan TF-IDF

## Klasifikasi Support Vector Machine (SVM)

Data yang telah memperoleh bobot TF-IDF digunakan sebagai fitur masukan dalam algoritma Support Vector Machine (SVM). SVM merupakan algoritma pembelajaran mesin

---

yang digunakan untuk menemukan hyperplane terbaik yang dapat memisahkan data ke dalam kelas yang berbeda. Dalam penelitian ini, SVM digunakan untuk mengklasifikasikan ulasan ke dalam dua kelas sentimen, yaitu positif dan negatif.

Model SVM dibangun menggunakan 224 data latih dan selanjutnya diuji menggunakan 56 data uji. Data uji tidak digunakan dalam proses pembentukan model sehingga dapat digunakan untuk mengetahui kemampuan model dalam melakukan klasifikasi terhadap data yang belum pernah dipelajari sebelumnya. SVM banyak digunakan dalam klasifikasi teks karena mampu bekerja pada data dengan jumlah fitur yang tinggi (Lubis et al., 2024; Natasuwarna et al., 2020).

### **Evaluasi dan Pengecekan Keabsahan Hasil**

Evaluasi dilakukan untuk mengetahui tingkat kemampuan model SVM dalam mengklasifikasikan sentimen. Evaluasi menggunakan confusion matrix yang membandingkan hasil prediksi model dengan label sebenarnya. Berdasarkan confusion matrix, performa model dihitung menggunakan Accuracy, Precision, Recall, dan F1-Score.

Accuracy digunakan untuk mengetahui persentase keseluruhan prediksi yang benar. Precision menunjukkan ketepatan model dalam memprediksi suatu kelas, sedangkan Recall menunjukkan kemampuan model dalam menemukan data yang termasuk dalam suatu kelas. F1-Score digunakan untuk mengukur keseimbangan antara Precision dan Recall.

Pengecekan keabsahan hasil dilakukan melalui dua tahap. Pertama, label sentimen diperiksa menggunakan expert judgment untuk memastikan kesesuaian label dengan makna ulasan. Kedua, hasil klasifikasi model dievaluasi menggunakan data uji dan confusion matrix sehingga performa model dapat diukur secara objektif berdasarkan label yang telah ditentukan.

### **Identifikasi Aspek Trust dan Perceived Risk**

Identifikasi aspek dilakukan untuk mengetahui faktor spesifik yang berkaitan dengan kepercayaan dan risiko yang muncul dalam ulasan pengguna. Trust terdiri atas dua kategori, yaitu Assurance dan Security. Assurance berkaitan dengan keandalan, kualitas, dan kemampuan layanan, sedangkan Security berkaitan dengan keamanan akun, perlindungan data, autentikasi, dan keamanan transaksi.

Sementara itu, Perceived Risk terdiri atas Transaction Risk dan System Risk. Transaction Risk berkaitan dengan risiko yang muncul dalam proses transaksi, seperti kegagalan, keterlambatan, atau permasalahan transaksi. System Risk berkaitan dengan risiko yang disebabkan oleh gangguan sistem, kesalahan aplikasi, atau fitur yang tidak berfungsi sebagaimana mestinya.

Identifikasi aspek dilakukan menggunakan kamus kata kunci yang disusun berdasarkan karakteristik masing-masing kategori. Kata kunci tersebut kemudian divalidasi melalui expert judgment untuk memastikan kesesuaiannya dengan konteks aspek yang diidentifikasi. Ulasan yang mengandung kata kunci tertentu kemudian dipetakan ke dalam kategori Trust atau Perceived Risk. Hasil identifikasi selanjutnya dihitung dan disajikan berdasarkan jumlah kemunculan masing-masing aspek.

**Tabel 3.**

**Kategori Trust dan Perceived Risk**

| <b>Faktor</b>         | <b>Kategori</b>         | <b>Fokus</b>                                    |
|-----------------------|-------------------------|-------------------------------------------------|
| <i>Trust</i>          | <i>Assurance</i>        | Keandalan, kualitas, dan kemampuan layanan      |
| <i>Trust</i>          | <i>Security</i>         | Keamanan akun, data, autentikasi, dan transaksi |
| <i>Perceived Risk</i> | <i>Transaction Risk</i> | Risiko yang berkaitan dengan proses transaksi   |
| <i>Perceived Risk</i> | <i>System Risk</i>      | Risiko akibat gangguan sistem atau fitur        |

## **HASIL DAN PEMBAHASAN**

### **Karakteristik Data Penelitian**

Penelitian ini menggunakan 300 ulasan pengguna aplikasi DANA yang diperoleh dari dataset Kaggle. Data tersebut terlebih dahulu melalui tahap pre-processing, pelabelan sentimen, dan validasi menggunakan expert judgment. Berdasarkan hasil pelabelan, data dikelompokkan menjadi sentimen positif, negatif, dan netral. Data dengan sentimen netral tidak digunakan dalam klasifikasi biner, sehingga diperoleh 280 ulasan yang terdiri atas sentimen positif dan negatif. Selanjutnya, data tersebut dibagi menjadi 224 data latih dan 56 data uji dengan rasio 80:20. Pembobotan TF-IDF diterapkan pada data teks untuk menghasilkan representasi numerik yang kemudian digunakan sebagai fitur dalam proses klasifikasi menggunakan Support Vector Machine (SVM).

## Hasil Klasifikasi Sentimen

Klasifikasi sentimen dilakukan menggunakan algoritma SVM dengan fitur hasil pembobotan TF-IDF. Sebanyak 224 data digunakan untuk proses pelatihan model, sedangkan 56 data digunakan untuk pengujian. Penggunaan TF-IDF sebagai representasi fitur dan SVM sebagai algoritma klasifikasi telah diterapkan pada penelitian sebelumnya dan menunjukkan kemampuan yang baik dalam klasifikasi teks (Alpaz et al., 2022; Sri Rahayu et al., 2023).

Hasil evaluasi model berdasarkan confusion matrix disajikan pada Tabel 4. Tabel tersebut menunjukkan nilai Accuracy, Precision, Recall, dan F1-Score yang diperoleh dari data pengujian.

**Tabel 4.**

**Hasil Evaluasi Klasifikasi SVM**

| Metrik Evaluasi  | Hasil |
|------------------|-------|
| <i>Accuracy</i>  | 98%   |
| <i>Precision</i> | 98%   |
| <i>Recall</i>    | 98%   |
| F1-Score         | 98%   |

Berdasarkan Tabel 4, algoritma SVM memperoleh nilai Accuracy, Precision, Recall, dan F1-Score masing-masing sebesar 98%. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam membedakan sentimen positif dan negatif pada 56 data uji. Nilai Precision sebesar 98% menunjukkan bahwa sebagian besar data yang diprediksi oleh model sebagai suatu kelas sesuai dengan kelas sebenarnya. Sementara itu, nilai Recall sebesar 98% menunjukkan bahwa model mampu mengenali sebagian besar data yang termasuk ke dalam masing-masing kelas sentimen. Nilai F1-Score sebesar 98% menunjukkan adanya keseimbangan yang baik antara kemampuan model dalam menghasilkan prediksi yang tepat dan menemukan data pada masing-masing kelas.

Hasil penelitian ini menunjukkan nilai akurasi yang lebih tinggi dibandingkan dengan beberapa penelitian terdahulu mengenai analisis sentimen ulasan aplikasi DANA. (Alva & Zuliarso, 2025) memperoleh akurasi sebesar 88,5%, sedangkan (Fajar Nawulansih et al., 2025) Memperoleh peningkatan akurasi dari 84% menjadi 90% setelah menerapkan SMOTE dan optimasi parameter. Penelitian (Cahya et al., 2024) Menggunakan 10.000 ulasan aplikasi DANA dan memperoleh akurasi sebesar 94,2%. Perbedaan hasil tersebut dapat dipengaruhi

---

oleh karakteristik dataset, jumlah data, tahapan pre-processing, metode pelabelan, pembagian data, serta parameter model yang digunakan.

Meskipun memperoleh nilai evaluasi sebesar 98%, hasil tersebut perlu dipahami berdasarkan karakteristik dataset penelitian. Jumlah data yang digunakan relatif terbatas, yaitu 280 data untuk klasifikasi, dengan 56 data sebagai data uji. Oleh karena itu, nilai evaluasi yang tinggi belum dapat secara langsung menunjukkan bahwa model akan menghasilkan performa yang sama ketika diterapkan pada dataset yang lebih besar atau pada ulasan dari periode yang berbeda. Hasil penelitian ini lebih tepat menunjukkan bahwa kombinasi TF-IDF dan SVM bekerja dengan baik pada dataset yang digunakan dalam penelitian.

Hasil tersebut juga sejalan dengan penelitian yang menerapkan SVM untuk klasifikasi teks dan analisis sentimen, seperti (Lubis et al., 2024) dan (Natasuwarna et al., 2020). Penggunaan TF-IDF membantu mengubah data teks menjadi representasi numerik berdasarkan tingkat kepentingan kata, sedangkan SVM menggunakan representasi tersebut untuk menentukan kelas sentimen. Dengan demikian, kombinasi TF-IDF dan SVM dapat digunakan sebagai pendekatan untuk mengklasifikasikan sentimen pada ulasan pengguna aplikasi DANA.

### **Hasil Identifikasi Aspek Trust dan Perceived Risk**

Selain melakukan klasifikasi sentimen, penelitian ini mengidentifikasi aspek Trust dan Perceived Risk untuk memperoleh informasi yang lebih spesifik mengenai pengalaman pengguna. Identifikasi aspek dilakukan terhadap 300 ulasan menggunakan kamus kata kunci yang telah disusun berdasarkan kategori penelitian dan divalidasi melalui expert judgment. Aspek Trust terdiri atas Assurance dan Security, sedangkan Perceived Risk terdiri atas Transaction Risk dan System Risk.

Hasil identifikasi aspek pada 300 ulasan pengguna aplikasi DANA disajikan pada Tabel 5. Penyajian tersebut digunakan untuk melihat distribusi kemunculan masing-masing aspek berdasarkan jumlah ulasan dan persentasenya.

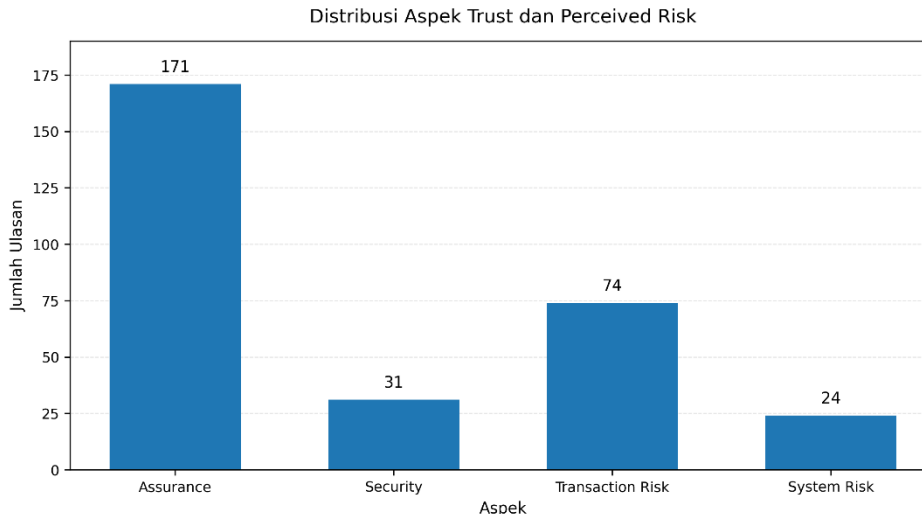
**Tabel 5.**

**Distribusi Aspek Trust dan Perceived Risk**

| Kategori       | Aspek            | Jumlah Ulasan | Persentase  |
|----------------|------------------|---------------|-------------|
| Trust          | Assurance        | 171           | 57%         |
| Trust          | Security         | 31            | 10%         |
| Perceived Risk | Transaction Risk | 74            | 25%         |
| Perceived Risk | System Risk      | 24            | 8%          |
| <b>Total</b>   |                  | <b>300</b>    | <b>100%</b> |

Berdasarkan Tabel 5, Assurance merupakan aspek yang paling banyak ditemukan dalam ulasan pengguna, yaitu sebanyak 171 ulasan atau 57%. Aspek Security ditemukan pada 31 ulasan atau 10%. Pada kategori Perceived Risk, Transaction Risk ditemukan pada 74 ulasan atau 25%, sedangkan System Risk ditemukan pada 24 ulasan atau 8%. Distribusi tersebut menunjukkan bahwa aspek yang berkaitan dengan keandalan dan jaminan layanan lebih sering muncul dalam ulasan dibandingkan aspek keamanan maupun risiko sistem.

Distribusi masing-masing aspek juga divisualisasikan pada Gambar 2. Grafik tersebut memberikan gambaran perbandingan jumlah ulasan pada aspek Assurance, Security, Transaction Risk, dan System Risk.



**Gambar 4.** Distribusi Aspek Trust dan Perceived Risk

Gambar 2 menunjukkan bahwa Assurance memiliki jumlah ulasan tertinggi, sedangkan System Risk memiliki jumlah ulasan terendah. Dominasi Assurance menunjukkan bahwa keandalan, kualitas, dan kemampuan layanan merupakan hal yang banyak muncul dalam pengalaman pengguna. Temuan tersebut menunjukkan bahwa keyakinan terhadap kemampuan layanan menjadi salah satu aspek yang menonjol dalam ulasan aplikasi DANA.

---

Aspek Security memiliki jumlah yang lebih rendah dibandingkan dengan Assurance, tetapi tetap menunjukkan bahwa keamanan akun, data, autentikasi, dan transaksi menjadi bagian dari perhatian pengguna. Hal ini menunjukkan bahwa kepercayaan pengguna terhadap layanan pembayaran digital tidak hanya berkaitan dengan kemampuan layanan dalam menjalankan fungsi transaksi, tetapi juga dengan kemampuan layanan dalam memberikan keamanan selama pengguna menggunakan aplikasi.

Pada kategori Perceived Risk, Transaction Risk merupakan aspek yang lebih banyak ditemukan dibandingkan dengan System Risk. Sebanyak 74 ulasan atau 25% berkaitan dengan risiko transaksi, sedangkan 24 ulasan atau 8% berkaitan dengan risiko sistem. Kondisi tersebut menunjukkan bahwa permasalahan yang berhubungan langsung dengan proses transaksi lebih sering muncul dalam ulasan dibandingkan dengan gangguan sistem atau fitur aplikasi. Temuan ini sejalan dengan penelitian (Cahya et al., 2024) yang menunjukkan adanya ulasan pengguna yang berkaitan dengan kegagalan transaksi dan permasalahan sistem pada aplikasi DANA.

### **Pembahasan Hasil Penelitian**

Hasil penelitian menunjukkan bahwa klasifikasi sentimen menggunakan TF-IDF dan SVM mampu menghasilkan performa yang tinggi pada dataset penelitian. Namun, penelitian ini tidak hanya berfokus pada kemampuan model dalam membedakan sentimen positif dan negatif. Identifikasi aspek Trust dan Perceived Risk memberikan informasi tambahan mengenai hal-hal yang menjadi perhatian pengguna ketika menggunakan aplikasi DANA.

Dominasi aspek Assurance menunjukkan bahwa keandalan dan kemampuan layanan menjadi aspek yang paling sering muncul dalam ulasan. Sementara itu, dominasi Transaction Risk pada kategori Perceived Risk menunjukkan bahwa pengalaman yang berkaitan dengan proses transaksi menjadi sumber risiko yang lebih sering dibicarakan oleh pengguna. Temuan ini memperlihatkan bahwa ulasan pengguna dapat memberikan informasi yang lebih rinci dibandingkan sekadar klasifikasi sentimen.

Penelitian terdahulu mengenai aplikasi DANA umumnya berfokus pada klasifikasi sentimen dan pengukuran performa algoritma, seperti penelitian (Alva & Zuliarso, 2025; Cahya et al., 2024; Fajar Nawulansih et al., 2025). Penelitian ini memperluas analisis tersebut dengan mengidentifikasi aspek Trust dan Perceived Risk yang muncul dalam ulasan. Dengan demikian, hasil penelitian tidak hanya menunjukkan bagaimana sentimen pengguna, tetapi

---

juga memberikan gambaran mengenai aspek layanan yang berkaitan dengan kepercayaan dan risiko yang dirasakan.

Secara keseluruhan, kombinasi klasifikasi sentimen dan identifikasi aspek memberikan hasil analisis yang lebih informatif terhadap ulasan pengguna aplikasi DANA. Klasifikasi SVM menunjukkan kecenderungan sentimen pengguna, sedangkan identifikasi aspek menunjukkan bagian layanan yang berkaitan dengan Trust dan Perceived Risk. Hasil tersebut dapat menjadi masukan bagi pengelola layanan untuk mempertahankan keandalan layanan, meningkatkan keamanan transaksi, serta mengurangi permasalahan yang berkaitan dengan risiko transaksi dan gangguan sistem.

## KESIMPULAN DAN REKOMENDASI

Penelitian ini menunjukkan bahwa penerapan Text Mining, pembobotan TF-IDF, dan algoritma Support Vector Machine (SVM) dapat digunakan untuk menganalisis sentimen pada ulasan pengguna aplikasi DANA. Dari 300 ulasan yang dikumpulkan, setelah proses pelabelan dan validasi, diperoleh 280 data bersentimen positif dan negatif yang digunakan dalam proses klasifikasi, dengan 224 data sebagai data latih dan 56 data sebagai data uji. Hasil pengujian menunjukkan nilai Accuracy, Precision, Recall, dan F1-Score masing-masing sebesar 98%. Hasil tersebut menunjukkan bahwa model SVM mampu melakukan klasifikasi sentimen positif dan negatif dengan tingkat kinerja yang tinggi pada data penelitian.

Selain klasifikasi sentimen, penelitian ini mengidentifikasi aspek Trust dan Perceived Risk pada 300 ulasan. Aspek Assurance merupakan aspek Trust yang paling dominan dengan 171 ulasan (57%), diikuti Security sebanyak 31 ulasan (10%). Pada aspek Perceived Risk, Transaction Risk ditemukan pada 74 ulasan (25%), sedangkan System Risk sebanyak 24 ulasan (8%). Temuan tersebut menunjukkan bahwa keandalan dan kemampuan layanan menjadi aspek kepercayaan yang paling banyak dibahas pengguna, sementara risiko yang berkaitan dengan proses transaksi menjadi aspek risiko yang paling dominan.

Berdasarkan temuan tersebut, pengembang DANA disarankan untuk mempertahankan kualitas dan keandalan layanan sebagai bagian dari Assurance, sekaligus meningkatkan keamanan serta keandalan proses transaksi untuk mengurangi Transaction Risk yang dirasakan pengguna. Untuk penelitian selanjutnya, penggunaan jumlah data yang lebih besar

dan beragam dapat dilakukan agar hasil lebih representatif. Identifikasi aspek juga dapat dikembangkan menggunakan metode Natural Language Processing atau pembelajaran mesin sehingga tidak hanya bergantung pada kamus kata kunci. Selain itu, perbandingan SVM dengan algoritma klasifikasi lainnya dan penerapan cross-validation dapat dilakukan untuk memperoleh evaluasi kinerja model yang lebih komprehensif.

## REFERENSI

1. Alpaz, F., H. Gustiana, M., Rosmala, D., & Fitrianti F, N. (2022). Akurasi dan presisi pengklasifikasian abstrak paper informatika menggunakan TF-IDF dan multiclass Support Vector Machine (SVM). *Prosiding Diseminasi FTI*, 1–12.
2. Alva, I. G. P., & Zuliarso, E. (2025). Analisis sentimen berdasarkan ulasan Google Play Store pada aplikasi DANA menggunakan metode Support Vector Machine (SVM). *Journal of Information Technology and Computer Science (INTECOMS)*, 4(6), 1–9. <https://doi.org/10.55041/isjem04122>
3. Cahya, I. T., Irma, A., Ali, I., Nurdiawan, O., & Martanto. (2024). Analisis sentimen ulasan aplikasi DANA di Play Store menggunakan Support Vector Machine. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, 6(1), 197–207. <https://doi.org/10.31849/zn.v6i1.17994>
4. DANA Indonesia. (2024). *Wujudkan ekosistem ekonomi digital yang positif, Rudiantara bergabung menjadi Komisaris Utama DANA Indonesia*. <https://www.dana.id/corporate/newsroom/wujudkan-ekosistem-ekonomi-digital-yang-positif-rudiantara-bergabung-menjadi-komisaris-utama-dana-indonesia>
5. Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2021). [Judul sumber belum tersedia].
6. Dumas, D., Organisciak, P., Maio, S., & [penulis lainnya]. (2021). Four text-mining methods for measuring elaboration. *The Journal of Creative Behavior*. <https://doi.org/10.1002/jocb.471>
7. Fajar Nawulansih, D., Ceisa Santi, N., & Aristia Sa'ida, I. (2025). Analisis sentimen ulasan aplikasi DANA di Google Play Store: Penerapan Support Vector Machine dan

- 
- Synthetic Minority Over-sampling Technique. *Jurnal Pendidikan dan Teknologi Indonesia*, 5(9), 2660–2671. <https://doi.org/10.52436/1.jpti.1053>
8. Hakim, B. (2021). Analisa sentimen data text preprocessing pada data mining dengan menggunakan machine learning. *Journal of Business and Audit Information Systems*, 4(2), 16–22.
  9. Indonesia, B. (2025). *Laporan kebijakan moneter*.
  10. Kobayashi, V. B., Mol, S. T., Berkers, H. A., & [penulis lainnya]. (2018). Text mining in organizational research. *Organizational Research Methods*. <https://doi.org/10.1177/1094428117722619>
  11. Lubis, S. K., Dar, M. H., & Nasution, F. A. (2024). Analisis sentimen ulasan pengguna aplikasi pada Google Play Store menggunakan algoritma Support Vector Machine. *Informatika*, 11(2), 120–128. <https://doi.org/10.36987/informatika.v11i2.5860>
  12. Natasuwarna, A. P. (2020). Seleksi fitur Support Vector Machine pada analisis sentimen keberlanjutan pembelajaran daring. *[Nama jurnal belum tersedia]*, 19(4), 437–448.
  13. Percha, B. (2021). Modern clinical text mining: A guide and review. *Annual Review of Biomedical Data Science*. <https://doi.org/10.1146/annurev-biodatasci-030421-030931>
  14. Sri Rahayu, J., Jaya Purnama, J., Hamid, A., & Kurnia Hikmawati, N. (2023). Analisis sentimen AicoGPT (Generative Pre-trained Transformer) menggunakan TF-IDF. *Jurnal Buana Informatika*, 14(2), 97–106. <https://doi.org/10.24002/jbi.v14i02.7039>
  15. Worldpay. (2024). *Global payments report 2024*. <https://www.worldpay.com/en/press-releases/worldpay-global-payments-report-2024-digital-wallet-maturity>