

## **ANALISIS SENTIMEN TERHADAP LAYANAN NASABAH BANK MENGGUNAKAN TEKNIK KLASIFIKASI NAIVE BAYES**

**Putri Puspa Wulan<sup>1</sup>, Hasan Basri\*<sup>2</sup>**

<sup>1,2</sup>Sistem Informasi, Universitas Siber Indonesia, Indonesia

Email: <sup>1</sup>[putripwulan@cyber-univ.ac.id](mailto:putripwulan@cyber-univ.ac.id), <sup>2</sup>[basrihasan@cyber-univ.ac.id](mailto:basrihasan@cyber-univ.ac.id)

(Naskah masuk : 28 Mei 2024, Revisi : 31 Mei 2024, Diterbitkan : 31 Mei 2024)

### **Abstrak**

Pelayanan pelanggan di sektor perbankan berpengaruh besar terhadap kepuasan dan kesetiaan nasabah. Namun, mengukur kepuasan dengan kuesioner tertutup seringkali tidak mencakup semua opini. Untuk mengatasi hal ini, analisis sentimen pada kuesioner terbuka digunakan. Sektor perbankan secara umum mengutamakan pelayanan pelanggan, dengan fokus pada senyum, salam, dan sapa. Penelitian ini bertujuan untuk menganalisis sentimen terhadap pelayanan pelanggan bank menggunakan metode klasifikasi *Naive Bayes*. Data diperoleh melalui survei kuesioner terbuka, kemudian dianalisis menggunakan metode *Naive Bayes* untuk mengklasifikasikan sentimen menjadi positif dan negatif. Hasil penelitian menunjukkan bahwa metode *Naive Bayes* efektif dengan akurasi 76.32%. Model yang terbentuk dapat digunakan untuk menganalisis komentar baru dari nasabah, mengekstrak sentimen positif dan negatif, serta memberikan wawasan mendalam kepada bank tentang penilaian nasabah terhadap pelayanan pelanggan. Dengan demikian, bank dapat mengambil tindakan yang sesuai untuk meningkatkan kepuasan nasabah dan kualitas pelayanan secara keseluruhan.

**Kata kunci:** *Customer service, Sentiment Analysis, Naive Bayes*

## **SENTIMENT ANALYSIS OF BANKING CUSTOMER SERVICE USING NAIVE BAYES CLASSIFICATION TECHNIQUE**

### **Abstract**

*Customer service in the banking sector significantly influences customer satisfaction and loyalty. However, measuring satisfaction with closed-ended questionnaires often fails to capture all opinions. To address this issue, sentiment analysis on open-ended questionnaires is employed. The banking sector generally prioritizes customer service, focusing on smiles, greetings, and interactions. This research aims to analyze sentiment towards bank customer service using the Naive Bayes classification method. Data is obtained through open-ended questionnaire surveys, then analyzed using the Naive Bayes method to classify sentiments into positive and negative categories. The research findings indicate that the Naive Bayes method is effective with an accuracy of 76.32%. The resulting model can be used to analyze new customer comments, extract positive and negative sentiments, and provide deep insights to the bank about customer perceptions of the service. Consequently, the bank can take appropriate actions to enhance customer satisfaction and overall service quality.*

**Keywords:** *Customer service, Sentiment Analysis, Naive Bayes*

---

### **1. PENDAHULUAN**

Pelayanan pelanggan merupakan aspek yang sangat penting dalam industri jasa, termasuk dalam industri perbankan. Bagaimana bank memberikan pelayanan pelanggan dapat memiliki dampak yang signifikan pada tingkat kepuasan dan loyalitas nasabah [1]. Namun, interaksi dengan *customer service* seringkali kompleks, dan pengalaman yang dialami oleh nasabah dapat bervariasi.

Terkadang, evaluasi terhadap kepuasan nasabah hanya dilakukan melalui kuesioner tertutup, yang tidak selalu mampu mencakup seluruh pendapat nasabah. Oleh karena itu, diperlukan pendekatan baru yang dapat memperhitungkan semua sudut pandang nasabah. Salah satu metode yang digunakan untuk mengatasi hal ini adalah analisis sentimen, yang memungkinkan penilaian yang lebih akurat dan komprehensif terhadap pelayanan pelanggan melalui pengolahan kuesioner terbuka [2].

Untuk berhasil dalam persaingan, sebuah perusahaan harus mampu menciptakan nilai dan memenuhi kepuasan pelanggan dengan menyediakan produk atau jasa berkualitas dan dengan harga yang bersaing. Oleh karena itu, *customer service* harus mampu menciptakan keunggulan kompetitif tersebut. Pelayanan pelanggan yang memuaskan akan menjadi faktor penting dalam menciptakan kepuasan dan kenyamanan bagi nasabah. Kepuasan nasabah sendiri merupakan kunci utama keberhasilan suatu bank, karena nasabah memiliki interaksi langsung dengan bank dan memiliki penilaian terhadap pelayanan yang diberikan. *Customer service* dan teller, sebagai ujung tombak di dalam dunia perbankan, memegang peran penting dalam melayani nasabah dengan prinsip 3S: Senyum, Salam, dan Sapa.

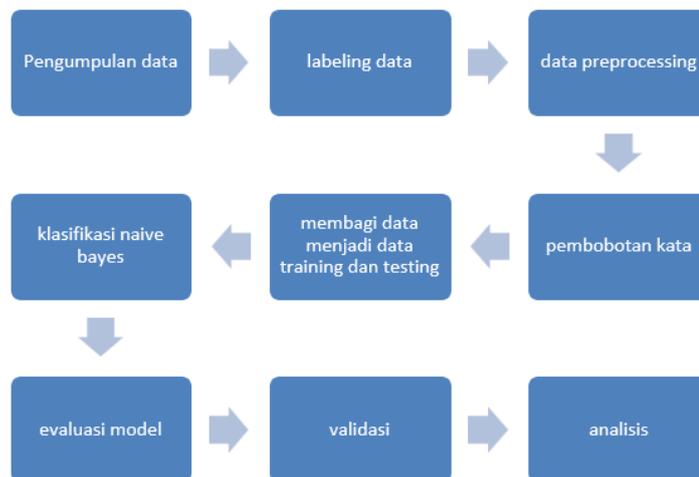
Hasil penelitian yang dilakukan oleh Andrian, dkk menunjukkan bahwa analisis sentimen dapat digunakan untuk menilai kepuasan nasabah terhadap penggunaan bank digital di Indonesia. Dengan tingkat akurasi sebesar 74,89%, penelitian tersebut menunjukkan bahwa analisis sentimen mampu cukup baik dalam menangkap pendapat public [3]. Penelitian lain yang sukses menggunakan analisis sentimen untuk menangkap opini publik adalah penelitian oleh Hossain & Rahman, yang berhasil mengklasifikasikan opini pelanggan terhadap produk asuransi [4].

Dengan demikian, analisis sentimen dapat membantu perusahaan perbankan dalam memahami perasaan dan sikap nasabah terhadap pelayanan pelanggan yang diberikan [5]. Dengan pemahaman ini, perusahaan dapat mengambil tindakan yang tepat untuk meningkatkan kualitas pelayanan dan kepuasan nasabah. Oleh karena itu, penelitian ini akan menjalankan "Analisis Sentimen Terhadap Layanan Nasabah Bank Menggunakan Teknik Klasifikasi *Naive Bayes*".

## 2. METODE PENELITIAN

*Sentiment analysis*, juga dikenal sebagai opinion mining, adalah proses menganalisis teks untuk mengidentifikasi sentimen, seperti positif, negatif, atau netral, yang dapat digunakan untuk memahami opini dan emosi orang terhadap suatu topik atau merek [6]. Berbagai metode termasuk rule-based, machine learning, dan hybrid approach digunakan untuk meningkatkan akurasi dan efektivitas analisis sentimen. Menurut Kadek Ary B, Made Sudarma dan Wayan, analisis sentimen membantu memahami dan mengelompokkan kata-kata dalam beberapa kelas [7].

Machine learning memanfaatkan algoritma untuk memahami pola sentimen dari data latihan, dan kemudian mengaplikasikannya pada data baru. Algoritma seperti *Naive Bayes*, *decision tree*, dan *support vector machine (SVM)* sering digunakan dalam proses ini [8][9]. Penelitian ini dilakukan untuk menggambarkan dan menganalisis sentimen nasabah terhadap pelayanan *customer service* perbankan menggunakan metode klasifikasi *Naive Bayes*. Berikut rancangan yang digunakan untuk melakukan penelitian ini tertera pada Gambar.1.



Gambar 1. Rancangan penelitian

Sesuai dengan Gambar 1, Sentiment analysis dimulai dengan pengumpulan data teks, kemudian dilabeli dengan sentimen tertentu. Data diproses untuk menghilangkan noise dan normalisasi. Kata-kata dinilai bobotnya, kemudian data dibagi menjadi data training dan data testing. Model diklasifikasikan, dievaluasi, dan divalidasi untuk akurasi, sebelum hasil analisis dipelajari dan disajikan.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Pengumpulan Data

Data diperoleh dari cabang Bank XYZ di daerah jakarta, cabang tersebut dipilih karena melayani banyak nasabah. Pengumpulan dilakukan selama satu bulan pada Maret 2023 untuk mendapatkan data yang representatif. Kuesioner terbuka digunakan untuk mengumpulkan pendapat, pengalaman, dan sentimen nasabah terhadap layanan *customer service* Bank XYZ. Sebanyak 191 komentar dari nasabah berhasil dikumpulkan selama periode tersebut.

#### 3.2. Labeling

Setelah pengumpulan data, langkah berikutnya adalah melakukan labeling data secara manual. Setiap komentar dikelompokkan ke dalam kategori sentimen yang sesuai, seperti positif, negatif, atau netral. Sentimen positif mencakup ekspresi kepuasan dan apresiasi, negatif mencakup keluhan, sementara sentimen netral tidak secara jelas menyatakan pandangan positif atau *negative* [10]. Contoh komentar dari nasabah disajikan pada Tabel 1.

Tabel 1. Contoh hasil pengumpulan data kuesioner terbuka

Komentar	Sentiment
untuk pelayanan sudah bagus.. harus di pertahankan untuk pelayanannya.	Positive
Disarankan agar karyawan Bank XYZ tanggap dalam menangani setiap keluhan nasabah.	Negative
Aplikasi XYZ bisa mencetak rekening koran karena internet via web sudah tidak dapat digunakan lagi	Neutral

Setelah melakukan pelabelan, jumlah data untuk masing-masing sentimen terdapat 127 data dengan sentimen positif, 60 data dengan sentimen negatif, dan 4 data dengan sentimen netral. Pada tahap selanjutnya, data sentimen netral tidak dapat digunakan karena jumlahnya sangat terbatas, hanya terdiri dari 4 data. Oleh karena itu, untuk analisis lebih lanjut, fokus akan diberikan pada data sentimen positif dan negatif yang lebih representatif dan memiliki jumlah data yang lebih signifikan.

#### 3.3. Data Processing

##### 1. Case folding

Pada analisis sentimen, *case folding* digunakan untuk merubah semua teks menjadi huruf kecil. Ini dilakukan untuk memastikan konsistensi, di mana kata dengan huruf besar atau kecil dianggap sama. Contohnya, "Baik" dan "baik" memiliki makna yang sama dalam sentimen positif. Tanpa *case folding*, kedua kata tersebut dianggap berbeda dan mempengaruhi hasil analisis. Gambar 2. menjelaskan hasil dari *case folding*.

```

case_folding
0          tetap semangat dalam menjalani hidup
1  untuk pelayanan sudah bagus.. harus di pertaha...
2          pelayanan dengan senang hati
3          dapat melayani setulus hati
4          mohon lebih cepat dalam melayani coustomer
..
186          ditambah fronlinernya biar tidak lama
187          sudah baik
188          cukup
189          ramah semuanya
190          sudah baik dan ramah
    
```

Gambar 2. Contoh *case folding*

##### 2. Stemming

Proses *stemming* adalah salah satu teknik dalam normalisasi teks yang bertujuan untuk mengubah kata-kata ke bentuk dasarnya atau kata dasar. *Stemming* menghilangkan imbuhan atau akhiran dari kata sehingga menyisakan akar kata yang memiliki arti dasar. Gambar 3. menunjukkan contoh hasil dari proses *stemming*.

```

stemmed_text
0          tetap semangat dalam jalan hidup
1  untuk layan sudah bagus harus di tahan untuk l...
2          layan dengan senang hati
3          dapat layan tulus hati
4          mohon lebih cepat dalam layan coustomer
..
186         tambah fronlinernya biar tidak lama
187         sudah baik
188         cukup
189         ramah semua
190        sudah baik dan ramah moga jadi kcp baik

```

Gmabar 3. Contoh hasil process *steaming*

### 3. Tokenisasi

Tokenisasi adalah proses memecah teks menjadi token, seperti kata-kata atau frasa, untuk memudahkan analisis teks. Ini memungkinkan setiap token diperlakukan secara terpisah. Sebagai contoh, diimplementasikan menggunakan *Natural Language Toolkit (NLTK)* dalam kode *Python* setelah *case folding*. Gambar 4. menunjukkan contoh hasil dari proses *steaming*.

```

tokens
0          [tetap, semangat, dalam, jalan, hidup]
1  [untuk, layan, sudah, bagus, harus, di, tahan,...
2          [layan, dengan, senang, hati]
3          [dapat, layan, tulus, hati]
4  [mohon, lebih, cepat, dalam, layan, coustomer]
..
186        [tambah, fronlinernya, biar, tidak, lama]
187        [sudah, baik]
188        [cukup]
189        [ramah, semua]
190        [sudah, baik, dan, ramah]

```

Gambar 4. Contoh hasil tokenisasi

### 4. Filtering

Setelah *case folding* dan tokenisasi, langkah berikutnya dalam *preprocessing* pada analisis sentimen adalah *filtering*. *Filtering* bertujuan menghilangkan kata-kata yang tidak penting dalam analisis, seperti kata-kata umum yang tidak memberi kontribusi signifikan. Jenis *filtering* meliputi penghapusan *stopwords*, kata-kata dengan panjang tertentu, atau frekuensi rendah, atau menggunakan daftar kata-kata yang ditentukan. Contoh penghapusan *stopwords* menggunakan NLTK: *Stemming* dapat dilihat pada Gambar 5.

```

filtered tokens
0          [semangat, jalan, hidup]
1  [layan, bagus, tahan, layan]
2          [layan, senang, hati]
3          [layan, tulus, hati]
4  [mohon, cepat, layan, coustomer]
..
186        [fronlinernya, biar]
187        []
188        []
189        [ramah]
190        [ramah]

```

Gambar 5. Hasil proses *filtering*

### 3.4. Klasifikasi

Dalam tahap klasifikasi, metode *Naive Bayes* dengan pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)* akan digunakan. Proses ini melibatkan persiapan data, pembagian data menjadi data *training* dan data *testing*, serta pembobotan kata menggunakan objek *TfidfVectorizer*. Kemudian, pembobotan *TF-IDF* diterapkan pada kedua jenis data. Model *Naive Bayes* diinisialisasi dan dilatih menggunakan data *training*. Setelah

pelatihan, model digunakan untuk memprediksi data testing dan mengevaluasi akurasi. Dengan pendekatan ini, klasifikasi pada data komentar dalam analisis sentimen dapat dilakukan.

Pada analisis sentimen, proses pembobotan kata menggunakan metode *TF-IDF*. Teknik ini menentukan pentingnya sebuah kata dalam dokumen atau kumpulan dokumen dengan menghitung bobotnya. Bobot ini berguna dalam menilai kesamaan antar dokumen. Proses ini melibatkan perhitungan frekuensi kemunculan kata dalam dokumen (TF), yang kemudian dinormalisasi dengan total kata dalam dokumen. Bobot ini kemudian dikalikan dengan nilai *inverse document frequency (IDF)*, yang merupakan logaritma dari jumlah dokumen dibagi jumlah dokumen yang mengandung kata tersebut. Rumus pembobotan kata dengan metode *TF-IDF* adalah sebagai berikut:

$$TF - IDF(w) = TF(w) * IDF(w)$$

dimana:

- $TF(w)$  = frekuensi kemunculan kata dalam dokumen
- $IDF(w) = \log(N/n)$ , di mana  $N$  adalah total dokumen dalam kumpulan dokumen dan  $n$  adalah jumlah dokumen yang mengandung kata tersebut.

Dengan adanya pembobotan kata ini, maka kata-kata yang muncul lebih sering dan memiliki nilai IDF yang tinggi akan memiliki bobot yang lebih besar dan dianggap lebih penting dalam analisis sentimen.

*Naive Bayes* adalah teknik klasifikasi probabilistik yang memanfaatkan teorema Bayes. Metode ini umumnya dipakai untuk mengklasifikasikan teks dalam analisis sentimen. Dasar dari *Naive Bayes* adalah mengestimasi probabilitas kemunculan sebuah kata dalam dokumen berdasarkan label. Dari situ, probabilitas setiap kata digunakan untuk menghitung probabilitas dokumen termasuk dalam label tertentu. Meskipun mengasumsikan independensi antar kata, metode ini efektif dalam banyak kasus. Rumus dasar *Naive Bayes* adalah sebagai berikut:

$$P(c|d) = P(c) * \frac{P(d|c)}{P(d)}$$

di mana:

- $P(c|d)$  adalah probabilitas kelas  $c$  berdasarkan dokumen  $d$
- $P(c)$  adalah probabilitas prior dari kelas  $c$
- $P(d|c)$  adalah probabilitas dokumen  $d$  jika diketahui kelas  $c$
- $P(d)$  adalah probabilitas dokumen  $d$

Untuk melakukan klasifikasi menggunakan *Naive Bayes*, data hasil *filtering* yang sudah diproses diubah menjadi kalimat kembali, kemudian data tersebut dibagi menjadi data training dan data split. Dalam kasus ini akan membagi data training menjadi 80% dan data testing menjadi 20%.

Proses pembobotan kata *TF-IDF* pada data training dan data testing melibatkan langkah-langkah berikut:

1. Data training ( $X_{train}$ ) diubah menjadi vektor *TF-IDF* dengan menggunakan metode *fit\_transform*.
2. Data testing ( $X_{test}$ ) juga diubah menjadi vektor *TF-IDF* menggunakan metode *transform*.

Proses `fit_transform` dalam konteks pembobotan kata *TF-IDF* melibatkan dua langkah utama. Pada tahap pertama (*fit*), model (dalam hal ini, *vectorizer*) belajar tentang data yang diberikan ( $X_{train}$ ). Ini melibatkan menghitung bobot *TF-IDF* untuk setiap kata dalam data pelatihan dan membangun kamus yang berisi informasi ini. Model ini menggunakan informasi ini untuk menyesuaikan diri dengan struktur dan karakteristik data pelatihan.

Tahap kedua (*transform*) melibatkan mengubah data pelatihan yang telah dipelajari oleh model menjadi representasi vektor *TF-IDF*. Ini berarti setiap kata dalam set data pelatihan diubah menjadi vektor yang mewakili bobot *TF-IDF*-nya. Proses transformasi ini menghasilkan set data baru yang terdiri dari vektor-vektor *TF-IDF* yang siap digunakan dalam analisis atau klasifikasi selanjutnya. `fit_transform` menggabungkan kedua langkah ini menjadi satu, yaitu melatih model pada data pelatihan dan secara bersamaan mengubah data tersebut menjadi representasi vektor *TF-IDF*.

Data pelatihan ( $X_{train\_tfidf}$ ) adalah data teks yang telah diubah menjadi representasi vektor *TF-IDF* menggunakan proses sebelumnya. Model *Naive Bayes* dilatih menggunakan metode *fit*. Ini berarti model akan mempelajari pola dan hubungan antara fitur-fitur (kata-kata yang telah diubah menjadi vektor *TF-IDF*) dan label-label ( $y_{train}$ ) yang sesuai. Model menggunakan data training ini untuk menyesuaikan parameter-parameter internalnya agar dapat melakukan prediksi yang akurat pada data baru.

Hasil menggunakan metode klasifikasi *Naive Bayes* untuk menganalisis sentimen pada komentar nasabah terkait layanan *customer service* Bank XYZ. Model dilatih dengan data training dan dievaluasi dengan data testing,

mencapai akurasi 76.32%. Meskipun masih ada ruang untuk perbaikan, hasil ini menunjukkan efektivitas model dalam mengidentifikasi sentimen positif dan negatif.

### 3.5. Evaluasi Model

*Confusion matrix* adalah tabel evaluasi untuk membandingkan prediksi model dengan nilai aktual, terdiri dari *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)*. TP menunjukkan jumlah prediksi positif yang benar, TN adalah prediksi negatif yang benar, sedangkan FP adalah prediksi positif yang salah, dan FN adalah prediksi negatif yang salah. Ini membantu mengevaluasi kesalahan dan akurasi model klasifikasi dengan memberikan gambaran yang jelas tentang performa klasifikasi. Hasil *Confusion matrix* dari data yang kita latih dapat dilihat pada Tabel. 2.

Tabel 2. *Confusion matrix*

Actual \ Predictive	Positive	Negative
Positive	3	9
Negative	0	26

Analisis dari Tabel. 1, *Confusion matrix* menggambarkan kinerja model klasifikasi dengan membandingkan prediksi model terhadap label yang sebenarnya. Dalam kasus ini, kita memiliki dua kelas sentimen, yaitu positif (positive) dan negatif (negative).

Di baris pertama, kolom pertama (3) menunjukkan jumlah data dengan sentimen sebenarnya positif yang diprediksi dengan benar sebagai positif oleh model. Kolom kedua (9) menunjukkan jumlah data dengan sentimen sebenarnya positif yang salah diprediksi sebagai negatif oleh model.

Di baris kedua, kolom pertama (0) menunjukkan jumlah data dengan sentimen sebenarnya negatif yang salah diprediksi sebagai positif oleh model. Kolom kedua (26) menunjukkan jumlah data dengan sentimen sebenarnya negatif yang diprediksi dengan benar sebagai negatif oleh model.

Dalam evaluasi performa model, terdapat beberapa metrik yang digunakan untuk mengukur kualitas klasifikasi model sentiment analysis. Pertama, akurasi sebesar 76.32% mengindikasikan sejauh mana model mampu mengklasifikasikan dengan benar seluruh data testing. Ini berarti model dapat memprediksi dengan tepat 76.32% data testing. Selanjutnya, presisi sebesar 82.41% menggambarkan tingkat ketepatan model dalam mengidentifikasi setiap kelas sentimen. Dengan presisi tersebut, model memberikan prediksi positif yang benar-benar positif sebanyak 82.41% dari semua prediksi positif yang dilakukan. Recall sebesar 76.32% menggambarkan kemampuan model dalam mendeteksi data yang termasuk dalam setiap kelas sentimen. Dalam hal ini, model dapat mengidentifikasi 76.32% data yang sebenarnya memiliki sentimen positif atau negatif. Terakhir, F1-Score sebesar 70.96% digunakan untuk memberikan gambaran keseluruhan tentang performa model dalam mengklasifikasikan sentimen dengan memadukan informasi presisi dan recall. Semakin tinggi F1-Score, semakin baik performa model dalam mengklasifikasikan sentimen secara seimbang. Dengan demikian, metrik evaluasi ini memberikan pemahaman yang lebih mendalam tentang kualitas klasifikasi yang dilakukan oleh model sentiment analysis.

Klasifikasi yang dilakukan dalam model sentiment analysis memiliki kualitas yang cukup baik. Hal ini dapat dilihat dari akurasi, presisi, recall, dan F1-Score yang memperoleh nilai yang relatif tinggi. Model mampu mengklasifikasikan dengan benar sebagian besar data testing (akurasi), memberikan prediksi yang akurat untuk setiap kelas sentimen (presisi), mampu mendeteksi sebagian besar data yang memiliki sentimen positif atau negatif (recall), dan mencapai keseimbangan antara presisi dan recall (F1-Score). Dengan demikian, dapat disimpulkan bahwa model *sentiment analysis* yang digunakan dalam penelitian ini memberikan hasil klasifikasi yang cukup baik.

## 4. KESIMPULAN

Berdasarkan hasil klasifikasi dengan metode *Naive Bayes*, dapat disimpulkan bahwa metode tersebut efektif dalam mengklasifikasikan sentimen terkait pelayanan *customer service* bank. Dengan akurasi 76.32%, model ini berhasil membedakan antara sentimen positif dan negatif. Model ini dapat digunakan untuk mengklasifikasikan komentar baru dari nasabah bank, memberikan wawasan tentang evaluasi mereka terhadap layanan *customer service*. Ini memungkinkan bank untuk memahami aspek positif yang telah dicapai dan masalah yang perlu diperbaiki guna meningkatkan kepuasan pelanggan.

## DAFTAR PUSTAKA

- [1] X. Zhang *et al.*, "Personalized Digital *Customer services* for Consumer Banking Call Centre using Neural Networks," in *2020 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2020, pp. 1–7. doi: 10.1109/IJCNN48605.2020.9206709.
- [2] R. Mehra, M. K. Bedi, G. Singh, R. Arora, T. Bala, and S. Saxena, "Sentimental analysis using fuzzy and

- Naive Bayes,” in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, Jul. 2017, pp. 945–950. doi: 10.1109/ICCMC.2017.8282607.
- [3] B. Andrian, T. Simanungkalit, I. Budi, and A. F. Wicaksono, “Sentiment Analysis on Customer Satisfaction of Digital Banking in Indonesia,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 3, 2022, doi: 10.14569/IJACSA.2022.0130356.
- [4] M. S. Hossain and M. F. Rahman, “Customer Sentiment Analysis and Prediction of Insurance Products’ Reviews Using Machine Learning Approaches,” *FIIB Bus. Rev.*, p. 231971452211157, Aug. 2022, doi: 10.1177/23197145221115793.
- [5] J. Kim and C. Lim, “Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses,” *Adv. Eng. Informatics*, vol. 49, p. 101304, Aug. 2021, doi: 10.1016/j.aei.2021.101304.
- [6] N. A. M. Razali *et al.*, “Opinion mining for national security: techniques, domain applications, challenges and research opportunities,” *J. Big Data*, vol. 8, no. 1, p. 150, Dec. 2021, doi: 10.1186/s40537-021-00536-5.
- [7] K. A. B. Permana, M. Sudarma, and W. G. Ariastina, “Analisis Rating Sentimen pada Video di Media Sosial Youtube Menggunakan STRUCT-SVM,” *Maj. Ilm. Teknol. Elektro*, vol. 18, no. 1, p. 113, May 2019, doi: 10.24843/MITE.2019.v18i01.P17.
- [8] P. Yang and Y. Chen, “A survey on sentiment analysis by using machine learning methods,” in *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Dec. 2017, pp. 117–121. doi: 10.1109/ITNEC.2017.8284920.
- [9] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” *Found. Trends® Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008, doi: 10.1561/1500000011.
- [10] M. N. M. Ibrahim and M. Z. M. Yusoff, “The impact of different training data set on the accuracy of sentiment classification of Naïve Bayes technique,” in *2017 IEEE Conference on Open Systems (ICOS)*, Nov. 2017, pp. 17–20. doi: 10.1109/ICOS.2017.8280267.