



ARTICLE

Analisis Sentimen pada Ulasan Produk dengan SVM dan Word2Vec

Sentiment Analysis on Product Reviews with SVM and Word2Vec

Widyastuti Andriyani,¹ Yuli Astuti,² Bradika Almandin Wisesa,³ dan Hengki⁴

¹Magister Teknologi Informasi, Fakultas Teknologi Informasi, Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia

²Prodi Manajemen Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

³Study Program of Software Engineering, IDN Politeknik Manufaktur Negeri Bangka Belitung Bangka, Indonesia

⁴Department of business digital, ISB Atma Luhur, Pangkalpinang, Indonesia

*Penulis Korespondensi: widya@utdi.ac.id

(Disubmit 22-09-24; Diterima 28-12-24; Dipublikasikan online pada 05-02-25)

Abstrak

Analisis sentimen adalah salah satu cabang pemrosesan bahasa alami (NLP) yang bertujuan untuk mengidentifikasi opini dalam teks. Penelitian ini mengusulkan model analisis sentimen dengan menggunakan kombinasi Word2Vec sebagai teknik representasi fitur dan Support Vector Machine (SVM) sebagai algoritma klasifikasi. Dataset yang digunakan adalah Amazon Customer Reviews, dengan 500 ribu sampel ulasan produk yang dilabeli sebagai sentimen positif atau negatif. Model yang diusulkan dibandingkan dengan baseline seperti Naive Bayes dan Logistic Regression, yang menggunakan representasi fitur berbasis TF-IDF. Hasil evaluasi menunjukkan bahwa SVM dengan Word2Vec menghasilkan akurasi 91.3%, precision 90.8%, recall 92.1%, dan F1-score 91.4%, lebih unggul dibandingkan model baseline. Grafik Precision-Recall Curve dan ROC Curve memperkuat temuan bahwa Word2Vec memberikan representasi fitur yang lebih informatif, yang secara signifikan meningkatkan performa SVM dalam tugas klasifikasi teks. Penelitian ini membuktikan efektivitas kombinasi Word2Vec dan SVM untuk analisis sentimen pada dataset besar dan kompleks. Pendekatan ini relevan untuk berbagai domain, seperti e-commerce dan analisis opini di media sosial, serta membuka peluang untuk pengembangan lebih lanjut menggunakan model berbasis transformer.

Kata kunci: Analisis Sentimen, Ulasan Produk, Support Vector Machine, Word2Vec, Amazon Customer Reviews

Abstract

Sentiment analysis is a branch of natural language processing (NLP) that aims to identify opinions in text. This research proposes a sentiment analysis model using a combination of Word2Vec as a feature representation technique and Support Vector Machine (SVM) as a classification algorithm. The dataset used is Amazon Customer Reviews, with 500 thousand product review samples labeled as positive or negative sentiment. The proposed model is compared with baselines such as Naive Bayes and Logistic Regression, which use TF-IDF based feature representation. The evaluation results show that SVM with Word2Vec yields 91.3% accuracy, 90.8% precision, 92.1% recall, and 91.4% F1-score, superior to the baseline model. Precision-Recall Curve and ROC Curve graphs reinforce the finding that Word2Vec provides a more informative feature representation, which significantly improves SVM performance in text classification tasks. This research proves the effectiveness of the combination of Word2Vec and SVM for sentiment analysis on large and complex datasets. The approach is relevant for various domains, such as e-commerce and opinion analysis on social media, and opens up opportunities for further development using transformer-based models.

KeyWords: Sentiment Analysis, Product Reviews, Support Vector Machine, Word2Vec, Amazon Customer Reviews

This is an Open Access article - copyright on authors, distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY SA) (<http://creativecommons.org/licenses/by-sa/4.0/>)

How to Cite: W. Andriyani *et al.*, "Analisis Sentimen pada Ulasan Produk Menggunakan Metode Pembelajaran Mesin", *JIKO (JURNAL INFORMATIKA DAN KOMPUTER)*, Volume: 8, No.1, Pages 173–185, Februari 2024, doi: 10.26798/jiko.v8i1.1498.

1. Pendahuluan

Dalam era digital, ulasan produk oleh konsumen menjadi elemen penting dalam pengambilan keputusan pembelian dan pengembangan produk. Ulasan ini tidak hanya mencerminkan opini konsumen tetapi juga memberikan masukan strategis bagi produsen untuk meningkatkan kualitas produk dan layanan [1]. Namun, dengan jumlah ulasan yang terus meningkat, analisis manual menjadi tidak efisien dan rentan terhadap subjektivitas [1, 2, 3]. Oleh karena itu, diperlukan pendekatan otomatis yang memanfaatkan kecerdasan buatan (AI) untuk menganalisis sentimen dalam ulasan produk.

Analisis sentimen adalah salah satu cabang dari *Natural Language Processing* (NLP) yang bertujuan untuk mengklasifikasikan opini dalam teks menjadi kategori seperti positif, negatif, atau netral [4, 5, 6]. Dalam dekade terakhir, pendekatan berbasis pembelajaran mesin (ML) telah menjadi solusi yang dominan dalam tugas ini karena mampu mempelajari pola dari data secara otomatis dan mengatasi kekurangan metode *rule-based* atau dictionary [7, 8, 9]. Salah satu algoritma *Mechine Learning* yang sering digunakan adalah *Support Vector Machine* (SVM), yang dikenal dengan keunggulannya dalam menangani data berdimensi tinggi dan *robust* terhadap data yang tidak seimbang [10, 11, 12].

Metode baseline seperti Naive Bayes dan Logistic Regression menggunakan representasi fitur berbasis TF-IDF, yang hanya mempertimbangkan frekuensi kata tanpa memperhatikan konteks semantik. Pendekatan ini penting karena model baseline memiliki keterbatasan dalam menangani hubungan semantik antar kata, yang dapat menyebabkan penurunan performa pada data dengan pola kompleks [13].

Selain algoritma klasifikasi, representasi teks yang baik menjadi faktor penting dalam keberhasilan analisis sentimen. Pendekatan berbasis *embedding* seperti *Word2Vec* mampu menghasilkan representasi vektor dari kata-kata yang mempertimbangkan hubungan semantik dan kontekstual antar kata [13, 14, 15]. Kombinasi antara algoritma klasifikasi seperti SVM dan teknik representasi fitur berbasis *embedding* telah terbukti memberikan hasil yang lebih baik dibandingkan metode berbasis fitur manual [14, 15, 16].

Beberapa penelitian sebelumnya telah menjadi landasan untuk penelitian ini. Penelitian mengenai analisis sentimen berbasis pembelajaran mesin menunjukkan bahwa SVM memiliki keunggulan dibandingkan algoritma lain seperti *Naive Bayes* dalam tugas klasifikasi teks [5, 12]. Penelitian lain membuktikan efektivitas *Word2Vec* dalam menghasilkan representasi fitur yang lebih informatif dibandingkan metode tradisional seperti *TF-IDF* [17]. Selain itu, kombinasi SVM dan *Word2Vec* telah diaplikasikan pada berbagai domain, termasuk *e-commerce* dan media sosial, untuk meningkatkan performa analisis sentimen [18]. Namun, penelitian ini masih memerlukan eksplorasi lebih lanjut pada konteks ulasan produk, terutama dalam perbandingan kinerjanya dengan metode alternatif seperti *Logistic Regression* atau *Random Forest* [19].

Meskipun pendekatan berbasis *Word2Vec* dan SVM telah menunjukkan hasil yang menjanjikan, masih diperlukan eksplorasi lebih lanjut untuk mengevaluasi kinerjanya dalam berbagai domain, seperti ulasan produk dengan dataset besar.

Penelitian ini bertujuan untuk mengembangkan model analisis sentimen pada ulasan produk dengan mengintegrasikan SVM dan *Word2Vec*. Kombinasi ini diharapkan dapat memberikan akurasi yang lebih baik dibandingkan metode *baseline*. Selain itu, penelitian ini juga akan mengevaluasi kinerja pendekatan yang diusulkan dengan membandingkannya dengan metode pembelajaran mesin lainnya untuk menilai efektivitasnya dalam menangkap sentimen dalam data ulasan produk. Dengan pendekatan ini, penelitian diharapkan tidak hanya menawarkan solusi yang lebih akurat untuk analisis sentimen tetapi juga memberikan kontribusi baru dalam pengembangan metode berbasis AI yang dapat diaplikasikan secara luas pada berbagai domain.

2. Metode

Penelitian ini bertujuan untuk mengembangkan model analisis sentimen pada ulasan produk dengan mengintegrasikan metode *Support Vector Machine* (SVM) dan *Word2Vec*. Proses penelitian melibatkan pengumpulan dataset, preprocessing data, pembentukan embedding kata menggunakan *Word2Vec*, pelatihan model SVM, serta evaluasi performa model untuk menilai efektivitas pendekatan yang diusulkan.

2.1 Dataset

Dataset yang digunakan adalah *Amazon Customer Reviews* yang tersedia di Kaggle. *Dataset* ini berisi lebih dari 130 juta ulasan konsumen terhadap berbagai produk pada sebuah *e-commerce*. Untuk penelitian ini, digunakan sampel sebanyak 500 ribu ulasan yang dipilih secara acak. Sampel ini dipilih untuk menjaga efisiensi komputasi tanpa mengurangi keragaman data.

Dataset Amazon Customer Reviews memiliki jutaan ulasan produk dengan berbagai label sentimen, yang mencerminkan opini konsumen dalam skala besar. Untuk menjaga efisiensi komputasi, penelitian ini menggunakan sampel sebanyak 500 ribu ulasan yang dipilih secara acak. Jumlah ini dianggap cukup untuk mewakili pola sentimen dalam dataset, karena sampel acak memastikan distribusi sentimen positif dan negatif yang serupa dengan dataset asli. Selain itu, ukuran ini memungkinkan pelatihan model dengan sumber daya komputasi yang tersedia tanpa mengorbankan kualitas hasil penelitian.

Setiap ulasan terdiri dari teks ulasan, *rating* produk, dan label sentimen (positif atau negatif). Label sentimen ditentukan berdasarkan *rating*: ulasan dengan *rating* di atas 3 diklasifikasikan sebagai positif, sedangkan ulasan dengan *rating* 3 atau lebih rendah diklasifikasikan sebagai negatif. Data ini dibagi menjadi 80% data pelatihan dan 20% data pengujian.

2.2 Preprocessing Data

Tahapan *preprocessing* dilakukan untuk membersihkan data teks dan menyiapkannya menjadi *input* yang dapat digunakan oleh algoritma pembelajaran mesin. Langkah-langkah *preprocessing* meliputi:

1. Menghapus karakter khusus, angka, dan tanda baca yang tidak relevan.
2. Melakukan *tokenisasi* untuk memecah teks ulasan menjadi kata-kata.
3. Menerapkan proses *stemming* untuk mengubah kata menjadi bentuk dasarnya.
4. Menghapus kata-kata umum yang tidak memberikan informasi signifikan, seperti "dan", "atau", dan "yang".

Setiap ulasan produk dalam dataset melalui beberapa tahapan *preprocessing* untuk membersihkan dan menormalkan data teks. Langkah pertama adalah penghapusan karakter non-alfabet, seperti angka dan tanda baca, untuk mengurangi kebisingan dalam data. Ulasan kemudian diubah menjadi huruf kecil (*lowercasing*) untuk memastikan konsistensi dan mempermudah proses *tokenisasi*. *Tokenisasi* dilakukan untuk memecah ulasan menjadi kata-kata terpisah. Selanjutnya, kata-kata umum (*stop words*) seperti 'dan', 'atau', 'yang' dihapus karena tidak memberikan kontribusi signifikan terhadap analisis sentimen. Langkah-langkah ini memastikan teks yang dihasilkan bersih dan siap untuk proses representasi fitur menggunakan Word2Vec. Setelah *preprocessing*, teks ulasan siap digunakan untuk proses pembentukan *embedding*.

2.3 Model Baseline Naive Bayes dan Logistic Regression

Sebagai pembandingan, penelitian ini menggunakan dua model *baseline*, yaitu *Naive Bayes* dan *Logistic Regression*, yang diterapkan pada dataset yang sama dengan model SVM. Kedua model ini dipilih karena merupakan algoritma standar yang sering digunakan dalam tugas analisis sentimen.

2.3.1 Naive Bayes

Naive Bayes digunakan dengan asumsi probabilitas bersyarat independen antar fitur. Representasi teks ulasan dilakukan menggunakan TF-IDF, sesuai dengan pendekatan tradisional. Model ini dilatih menggunakan data pelatihan yang telah di proses sebelumnya, dengan parameter *default* untuk memastikan keadilan perbandingan.

Fungsi probabilitas Naive Bayes adalah seperti terlihat pada rumus (1):

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)} \quad (1)$$

- di mana :
- $P(y|x)$: probabilitas kelas y (positif/negatif) diberikan ulasan x .
 - $P(x|y)$: probabilitas teks x diberikan kelas y .
 - $P(y)$: probabilitas prior untuk kelas y .
 - $P(x)$: probabilitas teks x , dianggap konstan.

2.3.2 Logistic Regression

Logistic Regression menggunakan representasi fitur TF-IDF yang sama. Model ini dioptimalkan untuk memaksimalkan *log-likelihood* dari kelas positif dan negatif, menggunakan fungsi sigmoid untuk menghasilkan probabilitas seperti terlihat pada rumus (2):

$$P(y|x) = \frac{1}{1 + e^{-z}} \quad (2)$$

dengan

$$z = w \cdot x + b$$

di mana w adalah vektor bobot, x adalah vektor fitur, dan b adalah bias.

Data pelatihan yang sama digunakan untuk melatih kedua model baseline ini, sehingga hasilnya dapat dibandingkan secara langsung dengan model SVM yang diusulkan.

2.4 Representasi Fitur dengan Word2Vec

Word2Vec digunakan dalam penelitian ini untuk menghasilkan representasi numerik dari teks ulasan produk. Pendekatan yang digunakan adalah Continuous Bag of Words (CBOW), yang mempelajari hubungan antar kata dalam konteks lokal. Setiap kata dalam dataset dipetakan ke dalam vektor dengan dimensi 300, yang merepresentasikan hubungan semantik antar kata. Untuk setiap ulasan, representasi vektor diperoleh dengan menghitung rata-rata vektor dari semua kata dalam ulasan tersebut. Secara matematis, jika sebuah ulasan terdiri dari n kata ($w_1, w_2, w_3, \dots, w_n$), maka representasi ulasan v_{ulasan} didefinisikan seperti terlihat pada rumus (3)

$$v_{ulasan} = \frac{1}{n} \sum_{i=1}^n v(w_i) \quad (3)$$

di mana $v(w_i)$ adalah vektor embedding dari kata w_i . Parameter Word2Vec yang digunakan adalah dimensi embedding sebesar 300, ukuran jendela (window size) 5, dan jumlah kata unik (vocabulary size) 50 ribu. Representasi ini memastikan bahwa informasi semantik dan kontekstual dari kata-kata dalam ulasan tetap terjaga, yang sangat penting untuk analisis sentimen.

Data teks yang telah di-*preprocessing* diubah menjadi representasi numerik menggunakan Word2Vec, yang memanfaatkan pendekatan probabilistik untuk menghasilkan *embedding* kata. Dalam penelitian ini, digunakan pendekatan Skip-Gram, yang memprediksi kata konteks berdasarkan kata target. Pendekatan ini dipilih karena lebih efektif dalam menangkap hubungan semantik pada data yang jarang.

Fungsi objektif Skip-Gram dirumuskan pada rumus (4) sebagai berikut:

$$\mathcal{J} = \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log P(w_{t+j} | w_t) \quad (4)$$

- di mana :
- t : Jumlah kata dalam korpus
 - w_t : Kata target pada waktu t
 - w_{t+j} : kata konteks dalam jendela c

Probabilitas $P(w_{t+j}|w_t)$ dihitung menggunakan *softmax* seperti terlihat pada rumus (5):

$$P(w_{t+j}|w_t) = \frac{\exp(v_{w_{t+j}} \cdot v'_{w_t})}{\sum_{w=1}^W \exp(v_{w_{t+j}} \cdot v'_{w_t})} \tag{5}$$

- di mana :
- v_{w_t} : vektor kata target.
- $v'_{w_{t+j}}$: vektor kata konteks.
- W : jumlah total kata dalam kosakata.

Untuk meningkatkan efisiensi, digunakan pendekatan *negative sampling*, yang memodifikasi fungsi objektif menjadi seperti terlihat pada rumus (6):

$$J = \log \sigma(v_{w_t} \cdot v'_{w_{context}}) + \sum_{i=1}^k E_{w_i \sim P(w)} [\log \sigma(-v_{w_t} \cdot v'_{w_i})] \tag{6}$$

- di mana :
- $\sigma(x)$: adalah fungsi sigmoid
- w_i : kata negatif yang diambil sampelnya
- $P(w)$: distribusi probabilitas kata dalam kosakata

Model Word2Vec dilatih menggunakan data pelatihan dengan dimensi vektor sebesar 100 dan jendela konteks sebesar 5 kata. Setiap ulasan direpresentasikan sebagai rata-rata vektor dari kata-kata yang menyunnnya.

2.5 Klasifikasi dengan *Support Vector Machine* (SVM)

Support Vector Machine (SVM) digunakan sebagai algoritma klasifikasi dalam penelitian ini. SVM dipilih karena keunggulannya dalam menangani data berdimensi tinggi, seperti representasi vektor dari Word2Vec, dan *robust* terhadap data yang tidak seimbang. Penelitian ini menggunakan kernel *Radial Basis Function* (RBF), yang dikenal fleksibel dalam menangani data non-linear.

Untuk mengoptimalkan performa model, parameter C dan γ dioptimalkan menggunakan teknik *grid search*. Parameter C merepresentasikan trade-off antara kesalahan klasifikasi pada data pelatihan dan margin keputusan, sedangkan γ menentukan pengaruh satu sampel data terhadap keputusan model. Optimasi *grid search* dilakukan dengan rentang $C = \{0.1, 1, 10, 100\}$ dan $\gamma \{0.001, 0.01, 0.1, 1\}$.

Kombinasi Word2Vec dan SVM relevan untuk analisis sentimen karena Word2Vec menghasilkan representasi fitur yang kaya akan informasi semantik, sementara SVM memberikan kemampuan klasifikasi yang kuat, terutama pada data dengan pola kompleks seperti ulasan produk. Pendekatan ini memastikan bahwa model dapat menangkap hubungan antar kata dalam teks ulasan dengan performa yang optimal.

Fungsi objektif dari *Soft-Margin* SVM adalah terlihat pada rumus (7):

$$\min_{m,b,\epsilon} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \epsilon_i \tag{7}$$

Parameter C berfungsi sebagai pengendali regulasi, yang menentukan keseimbangan antara margin maksimum dan toleransi terhadap kesalahan klasifikasi. Nilai C yang lebih besar memberikan penalti lebih besar terhadap kesalahan klasifikasi, sementara nilai C yang lebih kecil memberikan margin yang lebih fleksibel.

Fungsi keputusan SVM dirumuskan seperti pada rumus (8) berikut:

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b) \tag{8}$$

- di mana :
- x : vektor input
 - y_i : label kelas
 - $K(x_i, x)$: kernel fungsi untuk menghitung hubungan antar vektor
 - α_i : bobot hasil optimasi
 - b : bias

Model dilatih menggunakan data pelatihan, dan parameter seperti *cost* (C) dan *gamma* dioptimalkan menggunakan *grid search* untuk memastikan performa terbaik.

2.6 Proses Metode SVM dengan Word2Vec

Proses penggunaan algoritma SVM dengan Word2Vec dalam penelitian ini terdiri dari beberapa tahapan utama yang terstruktur untuk memudahkan replikasi oleh pembaca. Alur ini diuraikan sebagai berikut:

1. Pengumpulan Dataset

Dataset yang digunakan adalah Amazon Customer Reviews, tersedia di platform Kaggle. Dari dataset ini, diambil 500.000 sampel ulasan yang dipilih secara acak untuk menjaga keragaman data dan efisiensi komputasi.

Setiap ulasan mencakup teks, rating, dan label sentimen (positif/negatif). Label ditentukan berdasarkan rating: rating di atas 3 dianggap positif, sedangkan 3 atau di bawahnya dianggap negatif.

2. Preprocessing Data

Data ulasan diproses melalui langkah-langkah berikut: (1)Pembersihan teks: Menghapus karakter khusus, angka, dan tanda baca. (2)Tokenisasi: Memecah teks ulasan menjadi kata-kata individu. (3)Stemming dan stop-word removal: Mengubah kata menjadi bentuk dasarnya (contoh: "berjalan" menjadi "jalan") dan menghapus kata-kata umum yang tidak informatif, seperti "dan", "yang". (3)Normalisasi: Semua teks diubah ke huruf kecil untuk konsistensi.

3. Representasi Fitur dengan Word2Vec

Model Word2Vec: Penelitian ini menggunakan pendekatan Skip-Gram untuk menghasilkan embedding kata. Skip-Gram dipilih karena lebih efektif dalam menangkap hubungan semantik antara kata-kata dalam dataset besar.

Parameter Word2Vec: Dimensi embedding: 300, Ukuran jendela konteks: 5 kata, Kosakata: 50.000 kata unik.

Representasi Teks:Setiap ulasan direpresentasikan sebagai rata-rata vektor embedding dari semua kata dalam ulasan tersebut. Secara matematis seperti terlihat pada rumus (9):

$$v_{ulasan} = \frac{1}{n} \sum_{i=1}^n v(w_i) \quad (9)$$

4. Pelatihan Model SVM

Input Data: Vektor rata-rata dari Word2Vec digunakan sebagai input untuk algoritma Support Vector Machine (SVM). **Algoritma SVM:** Kernel *Radial Basis Function* (RBF) digunakan karena fleksibilitasnya dalam menangkap pola non-linear. **Optimasi Parameter:** Parameter C dan γ dioptimalkan menggunakan teknik *grid search*, dengan rentang sebagai berikut:

$$C = 0.1, 1, 10, 100, \gamma 0.001, 0.01, 0.1, 1$$

Fungsi Keputusan SVM dapat dilihat pada rumus (10):

$$f(x) = \text{sign} \left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right) \quad (10)$$

5. Evaluasi Model

Performa model dievaluasi menggunakan metrik akurasi, precision, recall, F1-score, dan AUC.

Diagram pada Gambar 1 menjelaskan alur utama proses yang digunakan dan menyajikan langkah-langkah utama proses penelitian, memberikan gambaran visual dari alur data mulai dari pengumpulan dataset

hingga evaluasi model dalam penelitian ini, mulai dari pengumpulan dataset hingga evaluasi performa model. Setiap tahap dirancang untuk memastikan hasil optimal dalam analisis sentimen.

2.7 Evaluasi Model

Kinerja model dievaluasi menggunakan beberapa metrik utama untuk memberikan penilaian yang menyeluruh. Akurasi digunakan untuk mengukur proporsi prediksi yang benar terhadap total data. *Precision* mengukur tingkat ketepatan model dalam memprediksi ulasan positif yang benar. *Recall* menilai sejauh mana model mampu menangkap semua ulasan positif. *F1-score* digunakan sebagai metrik keseimbangan antara *precision* dan *recall*, terutama pada data yang tidak seimbang.

Hasil evaluasi dibandingkan dengan metode *baseline* seperti *Naive Bayes* dan *Logistic Regression*. Perbandingan ini bertujuan untuk menunjukkan keunggulan pendekatan Word2Vec-SVM dalam tugas analisis sentimen.

Evaluasi performa model dilakukan dengan menggunakan empat metrik utama, yaitu akurasi, *precision*, *recall*, dan *F1-score*.

Akurasi mengukur proporsi prediksi yang benar terhadap total data. *Precision* mengukur tingkat ketepatan model dalam memprediksi ulasan positif yang benar. *Recall* mengukur sejauh mana model mampu menangkap semua ulasan positif. *F1-score* digunakan untuk mengevaluasi keseimbangan antara *precision* dan *recall*, terutama pada dataset yang tidak seimbang. Metrik ini memberikan pandangan menyeluruh tentang performa klasifikasi dalam tugas analisis sentimen.

Validasi silang 5-fold diterapkan untuk memastikan model tidak *overfitting* pada data pelatihan. Teknik ini membagi data pelatihan menjadi lima lipatan, di mana setiap lipatan secara bergantian digunakan sebagai data validasi, sementara sisanya digunakan untuk pelatihan. Pendekatan ini memastikan hasil evaluasi lebih general terhadap data yang belum pernah dilihat sebelumnya.

Selain itu, *Precision-Recall Curve* dan *ROC Curve* digunakan untuk menggambarkan performa model pada berbagai *threshold*. *Precision-Recall Curve* memberikan pandangan tentang keseimbangan antara *precision* dan *recall*, sedangkan *ROC Curve* menunjukkan kemampuan model dalam membedakan antara kelas positif dan negatif. Nilai AUC (*Area Under Curve*) dari *ROC Curve* digunakan untuk mengukur kemampuan diskriminasi model secara keseluruhan.

Gambar 1 menunjukkan tahapan-tahapan utama dalam penelitian ini. Setiap tahap dirancang untuk memastikan pengolahan data dan implementasi model berjalan secara sistematis dan terstruktur, sehingga menghasilkan model analisis sentimen yang optimal. Berikut penjelasan setiap tahapan:

1. Pengumpulan Dataset

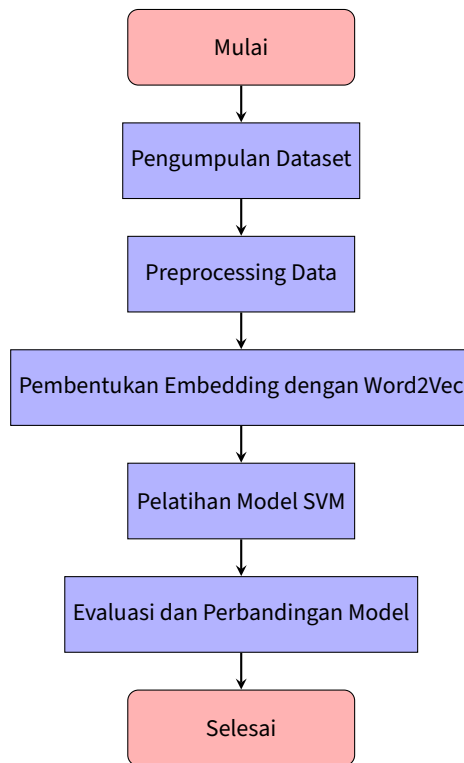
Tahap pertama dimulai dengan pengumpulan dataset ulasan produk, yaitu *Amazon Customer Reviews* yang diperoleh dari platform Kaggle. Dataset ini berisi ulasan konsumen yang mencakup teks ulasan, rating, dan metadata terkait. Dataset disiapkan dengan memilih sampel sebanyak 500 ribu ulasan secara acak untuk menjaga efisiensi pemrosesan.

2. Preprocessing Data

Data ulasan yang telah terkumpul di proses melalui tahapan *preprocessing* untuk membersihkan data teks dari karakter yang tidak relevan seperti tanda baca, angka, atau simbol. Tokenisasi dilakukan untuk memecah ulasan menjadi kata-kata. Kata-kata umum yang tidak memiliki nilai informasi signifikan, seperti "dan", "atau", dan "yang", dihapus. Selain itu, setiap kata disederhanakan menjadi bentuk dasarnya melalui proses *stemming* atau *lemmatization*.

3. Pembentukan *Embedding* dengan Word2Vec

Data teks hasil *preprocessing* diubah menjadi representasi numerik menggunakan Word2Vec. Proses ini menghasilkan *embedding* kata yang mencerminkan hubungan semantik antar kata dalam teks ulasan. Model Word2Vec dilatih menggunakan pendekatan Skip-Gram, yang memprediksi kata konteks berdasarkan kata target. Setiap ulasan direpresentasikan sebagai rata-rata dari vektor *embedding* kata-kata yang ada dalam ulasan tersebut.



Gambar 1. Alur Penelitian

4. Pelatihan Model SVM

Representasi numerik dari ulasan digunakan sebagai *input* untuk model *Support Vector Machine* (SVM). Model SVM dilatih untuk memisahkan ulasan menjadi dua kategori, yaitu positif dan negatif, menggunakan *kernel Radial Basis Function* (RBF) untuk menangkap pola non-linear. Parameter seperti C (*cost*) dan γ dioptimalkan menggunakan metode *grid search*.

5. Evaluasi dan Perbandingan Model

Setelah pelatihan, kinerja model diuji menggunakan data pengujian. Evaluasi dilakukan berdasarkan metrik seperti akurasi, *precision*, *recall*, dan *F1-score* untuk memberikan gambaran menyeluruh tentang performa model. Model yang diusulkan, yaitu SVM dengan Word2Vec, dibandingkan dengan metode *baseline* seperti *Naive Bayes* dan *Logistic Regression* untuk menunjukkan keunggulan pendekatan ini.

3. Hasil

Penelitian ini mengevaluasi performa model analisis sentimen yang dikembangkan menggunakan SVM dengan representasi fitur Word2Vec. Untuk memberikan perbandingan, hasil model *baseline* menggunakan *Naive Bayes* dan *Logistic Regression* dengan representasi fitur TF-IDF juga disajikan. Evaluasi dilakukan menggunakan data pengujian yang terdiri dari 100.000 ulasan produk.

Tabel 1 menyajikan hasil evaluasi model menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Data *baseline* diperoleh dengan menggunakan TF-IDF sebagai representasi fitur untuk model *Naive Bayes* dan *Logistic Regression*, sementara model yang diusulkan menggunakan Word2Vec.

Tabel 1. Hasil Evaluasi Model dengan Validasi Silang 5-Fold

Model	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
Naive Bayes	81.4 ± 0.45	80.2 ± 0.48	82.6 ± 0.50	81.4 ± 0.47
Logistic Regression	84.9 ± 0.40	83.7 ± 0.42	85.5 ± 0.43	84.6 ± 0.41
SVM + Word2Vec	91.3 ± 0.35	90.8 ± 0.38	92.1 ± 0.37	91.4 ± 0.42

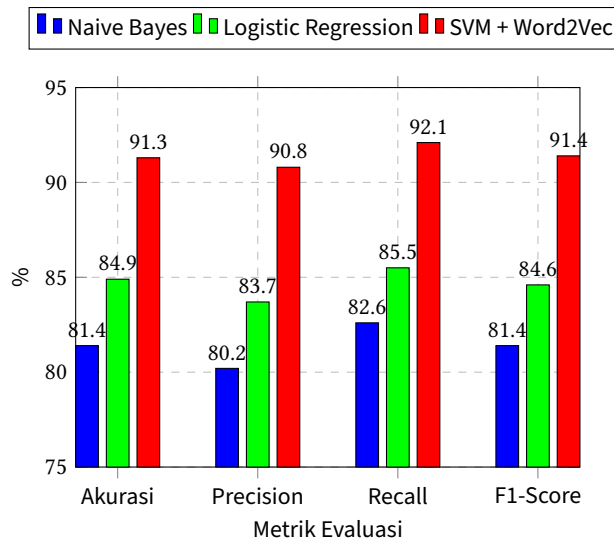
Dari tabel 1, terlihat bahwa model SVM dengan Word2Vec memiliki performa yang lebih unggul diban-

dingkan Naive Bayes dan Logistic Regression. Akurasi model yang diusulkan mencapai 91.3%, yang menunjukkan peningkatan signifikan dibandingkan Naive Bayes (81.4%) dan Logistic Regression (84.9%). Nilai *precision*, *recall*, dan *F1-score* juga lebih tinggi, yang menunjukkan konsistensi model dalam mengklasifikasikan sentimen ulasan produk.

Performa *precision* yang tinggi pada model SVM+Word2Vec, seperti yang ditunjukkan dalam Tabel 1, sangat relevan dalam domain seperti rekomendasi produk. Dalam aplikasi ini, prediksi positif yang salah (*false positives*) dapat mengganggu pengalaman pengguna atau menyebabkan rekomendasi yang tidak sesuai. Sebaliknya, *recall* yang tinggi memastikan bahwa opini penting, seperti ulasan negatif yang kritis, tidak terlewatkan. Kombinasi *precision* dan *recall* yang seimbang pada SVM+Word2Vec membuatnya cocok untuk analisis opini di berbagai domain, termasuk *e-commerce* dan media sosial.

Hasil evaluasi pada Tabel 1 berasal dari proses validasi silang 5-fold. Model SVM+Word2Vec menunjukkan stabilitas performa yang baik, dengan deviasi standar untuk akurasi sebesar 0.35% dan F1-score sebesar 0.42%. Nilai ini menunjukkan bahwa model menghasilkan hasil yang konsisten di semua lipatan data, sehingga dapat diandalkan untuk analisis pada dataset yang belum pernah dilihat sebelumnya. AUC sebesar 0.92 menunjukkan bahwa model ini mampu membedakan kelas positif dan negatif dengan akurasi tinggi, bahkan pada data yang sulit diklasifikasikan.

Gambar 2 memberikan visualisasi perbandingan performa model berdasarkan metrik evaluasi.



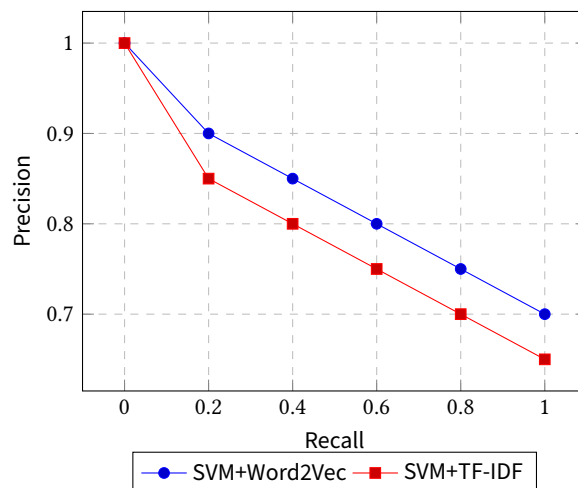
Gambar 2. Perbandingan Performa Model Berdasarkan Metrik Evaluasi

Grafik ini mempertegas bahwa model yang menggunakan Word2Vec menghasilkan performa yang lebih baik pada semua metrik dibandingkan dua *baseline*. Hal ini menunjukkan bahwa representasi fitur berbasis Word2Vec memberikan keunggulan signifikan dibandingkan TF-IDF.

Selain metrik individu, grafik Precision-Recall Curve pada Gambar 3 menggambarkan hubungan antara *precision* dan *recall* untuk model yang diuji. Grafik ini memberikan pandangan lebih jelas tentang keseimbangan performa pada berbagai *threshold*.

Keunggulan SVM+Word2Vec dalam *precision* dan *recall* dibandingkan *Logistic Regression* dapat dijelaskan melalui cara kedua metode menangani representasi fitur. Word2Vec menghasilkan vektor kata yang mempertimbangkan hubungan semantik antar kata, sehingga memberikan representasi fitur yang lebih informatif untuk SVM. Sebaliknya, TF-IDF hanya mengandalkan frekuensi kata tanpa memperhatikan konteks semantik, yang menyebabkan penurunan performa *Logistic Regression* pada dataset dengan pola yang kompleks. Selain itu, SVM lebih efektif dalam menangani data berdimensi tinggi, seperti vektor yang dihasilkan oleh Word2Vec, dibandingkan *Logistic Regression* yang sensitif terhadap redundansi fitur.

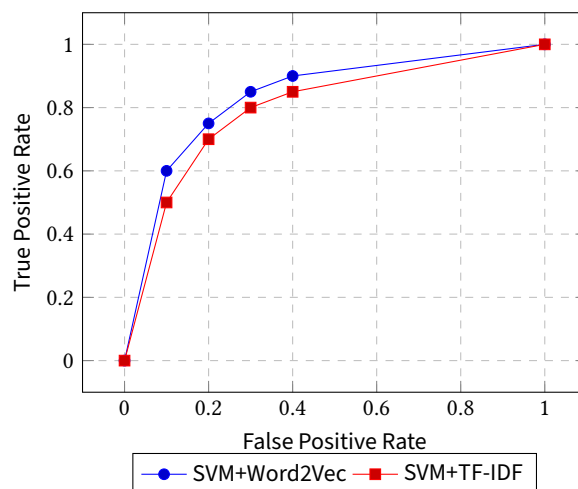
Gambar 4 menunjukkan grafik ROC Curve, yang memberikan gambaran menyeluruh tentang performa klasifikasi dengan menggunakan metrik *True Positive Rate* dan *False Positive Rate*. AUC (*Area Under Curve*)



Gambar 3. Precision-Recall Curve for SVM Models

yang lebih besar menunjukkan kemampuan model dalam membedakan kelas dengan lebih baik.

Gambar 4 menunjukkan ROC Curve untuk model SVM+Word2Vec dan baseline. Kurva yang lebih dekat ke sudut kiri atas menunjukkan kemampuan model untuk membedakan antara kelas positif dan negatif. Nilai AUC sebesar 0.92 untuk SVM+Word2Vec mengonfirmasi performa yang sangat baik, dibandingkan dengan AUC sebesar 0.87 untuk Logistic Regression dengan TF-IDF. Dalam konteks data tidak seimbang, AUC yang lebih tinggi memastikan bahwa model mampu mengidentifikasi sentimen minoritas dengan lebih andal, yang penting dalam domain seperti ulasan produk.



Gambar 4. ROC Curve for SVM Models

4. Pembahasan

Hasil penelitian ini menunjukkan bahwa integrasi Word2Vec sebagai teknik representasi fitur dengan SVM memberikan kontribusi signifikan dalam meningkatkan performa model analisis sentimen. Model yang diusulkan mencapai akurasi, *precision*, *recall*, dan *F1-score* yang lebih tinggi dibandingkan *baseline*, yaitu *Naive Bayes* dan *Logistic Regression*, yang menggunakan representasi fitur tradisional berbasis TF-IDF.

Kinerja yang lebih rendah pada *Naive Bayes* dapat dijelaskan oleh asumsi independensi antar fitur yang tidak sepenuhnya berlaku dalam teks ulasan. TF-IDF hanya mempertimbangkan frekuensi kata dalam dokumen dan tidak mampu menangkap hubungan semantik antar kata. Sebagai contoh, kata-kata sinonim seperti "baik" dan "bagus" dianggap entitas yang berbeda dalam TF-IDF, sedangkan Word2Vec dapat memetakan kedua kata tersebut ke dalam ruang vektor yang berdekatan.

Logistic Regression memberikan hasil yang lebih baik dibandingkan *Naive Bayes* karena pendekatannya

yang lebih fleksibel untuk memodelkan hubungan linier antara fitur dan kelas. Namun, pendekatan ini tetap kalah dari SVM dengan Word2Vec, yang mampu menangkap pola non-linear dan hubungan semantik antar kata dalam teks.

Peningkatan *F1-score* pada model yang diusulkan mencerminkan keseimbangan yang baik antara *precision* dan *recall*, yang penting dalam tugas analisis sentimen. Model ini mampu menangkap pola sentimen positif dan negatif dengan akurasi tinggi tanpa banyak kesalahan prediksi.

Precision-Recall Curve (Gambar 3) menunjukkan hubungan antara *precision* dan *recall* untuk model yang diuji, yaitu SVM+Word2Vec dan SVM+TF-IDF. Grafik ini menggambarkan performa model dalam menangani prediksi positif pada berbagai *threshold*.

Pada Gambar 3, terlihat bahwa model SVM+Word2Vec mempertahankan *precision* yang lebih tinggi dibandingkan SVM+TF-IDF pada berbagai tingkat *recall*. Sebagai contoh, ketika *recall* berada pada tingkat menengah hingga tinggi (0.6 hingga 1.0), *precision* SVM+Word2Vec stabil di atas 75%, sedangkan SVM+TF-IDF menunjukkan penurunan yang lebih tajam. Hal ini mengindikasikan bahwa Word2Vec sebagai teknik representasi fitur mampu menangkap hubungan semantik antar kata dengan lebih baik, sehingga meningkatkan ketepatan prediksi positif bahkan ketika *recall* tinggi.

Grafik ini juga menunjukkan bahwa pada tingkat *recall* rendah (0.2 hingga 0.4), kedua model memiliki *precision* yang hampir serupa. Namun, performa Word2Vec tetap unggul secara keseluruhan, yang memperkuat argumen bahwa Word2Vec menghasilkan representasi fitur yang lebih informatif dibandingkan TF-IDF.

ROC Curve (Gambar 4) memberikan gambaran kemampuan model dalam membedakan antara kelas positif dan negatif. Hubungan antara True Positive Rate (TPR) dan False Positive Rate (FPR) pada berbagai *threshold* membantu mengevaluasi performa model secara keseluruhan.

Dari Gambar 4, terlihat bahwa kurva SVM+Word2Vec lebih dekat ke sudut kiri atas dibandingkan kurva SVM+TF-IDF, menunjukkan performa yang lebih baik. AUC (Area Under Curve) untuk SVM+Word2Vec lebih besar, yang mencerminkan kemampuan diskriminasi yang lebih tinggi dibandingkan SVM+TF-IDF. Hal ini menunjukkan bahwa SVM+Word2Vec lebih andal dalam membedakan kelas positif dan negatif, bahkan pada data yang sulit diklasifikasikan.

Salah satu alasan utama untuk hasil ini adalah kemampuan Word2Vec dalam menghasilkan vektor yang mempertimbangkan hubungan semantik antar kata, yang meningkatkan kualitas input untuk model SVM. Sebaliknya, TF-IDF hanya mempertimbangkan frekuensi kata tanpa memperhatikan konteks, yang menyebabkan penurunan performa pada data dengan pola yang lebih kompleks.

Secara keseluruhan, hasil ini menunjukkan bahwa pendekatan Word2Vec dengan SVM lebih unggul untuk analisis sentimen ulasan produk, khususnya pada dataset yang besar dan beragam seperti *Amazon Customer Reviews*. Keunggulan ini membuat pendekatan yang diusulkan relevan untuk diterapkan pada berbagai domain lain yang membutuhkan analisis opini dalam jumlah besar.

Hasil ini menunjukkan bahwa pendekatan SVM + Word2Vec tidak hanya efektif untuk dataset *Amazon Customer Reviews*, tetapi juga memiliki potensi untuk diterapkan dalam analisis opini di media sosial atau ulasan layanan publik.

5. Simpulan

Penelitian ini mengembangkan model analisis sentimen menggunakan kombinasi Word2Vec untuk representasi fitur dan Support Vector Machine (SVM) sebagai algoritma klasifikasi. Model yang diusulkan menunjukkan performa unggul dibandingkan baseline, dengan akurasi 91.3%, *precision* 90.8%, *recall* 92.1%, dan *F1-score* 91.4%.

Temuan ini mengindikasikan bahwa Word2Vec mampu menangkap hubungan semantik antar kata yang sulit ditangkap oleh pendekatan tradisional seperti TF-IDF. Kombinasi Word2Vec dan SVM juga menunjukkan kemampuan yang lebih baik dalam menangani konteks ulasan yang kompleks, seperti sentimen implisit atau kombinasi positif dan negatif dalam satu ulasan.

Relevansi praktis dari penelitian ini meliputi aplikasi pada analisis opini konsumen di berbagai domain, seperti e-commerce, layanan publik, dan media sosial. Pendekatan ini dapat memberikan wawasan yang lebih mendalam tentang persepsi konsumen terhadap produk atau layanan tertentu.

Namun, penelitian ini memiliki keterbatasan pada penggunaan rata-rata vektor Word2Vec, yang mungkin kurang optimal untuk menangkap konteks pada teks yang lebih panjang atau kompleks. Penelitian masa depan dapat mengatasi keterbatasan ini dengan mengeksplorasi model berbasis transformer, seperti BERT atau GPT, untuk menangkap hubungan semantik yang lebih dalam serta menggunakan data dari berbagai bahasa atau domain.

Pustaka

- [1] L. E. Putri and M. A. Humaira, "Pengaruh ulasan produk terhadap keputusan pembelian konsumen," *karimahtauhid*, vol. 3, no. 1, pp. 694–702, Jan. 2024.
- [2] S. Hartini, M. Kurniawati, and M. Ihwanudin. (2022) Customer review: Impact on choice confidence, product attitude, and purchase intention. [Online]. Available: <https://journalppw.com/index.php/jpsp/article/view/10868/7018>.
- [3] C. F. Riska, Y. Hendayana, and M. Wijayanti. (2023) Pengaruh ulasan konsumen, kualitas produk dan marketing influencer terhadap keputusan pembelian produk erigo (studi kasus mahasiswa universitas bhayangkara jakarta raya).
- [4] B. Saju, S. Jose, and A. Antony, "Comprehensive study on sentiment analysis: Types, approaches, recent applications, tools and apis," in *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*, Cochin, India, 2020, pp. 186–193.
- [5] N. D. Putranti and E. Winarko. (n.d.) Analisis sentimen twitter untuk teks berbahasa indonesia dengan maximum entropy dan support vector machine. [Online]. Available: <https://doi.org/10.22146/ijccs.3499>.
- [6] S. Kumawat, I. Yadav, N. Pahal, and D. Goel, "Sentiment analysis using language models: A study," in *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 2021, pp. 984–988.
- [7] M. S. Islam, M. N. Kabir, N. Ab Ghani, K. Z. Zamli, N. S. A. Zulkifli, M. M. Rahman, and M. A. Mo-ni, "Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach," 2024.
- [8] R. Chahar and A. K. Dubey, "Machine learning-based sentiment analysis: A comprehensive review," in *2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS)*, Coim-batore, India, 2024, pp. 52–57.
- [9] S. Rezvani, F. Pourpanah, C. P. Lim, and Q. M. J. Wu, "Methods for class-imbalanced learning with support vector machines: a review and an empirical evaluation," 2024.
- [10] (n.d.) Sentiment analysis methods: Overview, pros & cons. [Online]. Available: <https://research.aimultiple.com/sentiment-analysis-methods/>.
- [11] A. Nurdin, B. Seno Aji, A. Bustamin, and Z. Abidin, "Perbandingan kinerja word embedding word2vec, glove, dan fasttext pada klasifikasi teks," *Jurnal Tekno Kompak*, vol. 14, p. 74, 2020.
- [12] I. Rifky, E. Utami, and A. Dwi, "Analisis perbandingan metode tf-idf dan word2vec pada klasifikasi teks sentimen masyarakat terhadap produk lokal di indonesia," *Smart Comp: Jurnalnya Orang Pintar Komputer*, vol. 11, 2022.
- [13] S. Styawati, A. Nurkholis, A. A. Aldino, S. Samsugi, E. Suryati, and R. P. Cahyono, "Sentiment analysis on online transportation reviews using word2vec text embedding model feature extraction and support vector machine (svm) algorithm," in *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, Jakarta, Indonesia, 2022, pp. 163–167.

- [14] J. Lilleberg, Y. Zhu, and Y. Zhang, "Support vector machines and word2vec for text classification with semantic features," in *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*, Beijing, China, 2015, pp. 136–140.
- [15] E. M. Alshari, A. Azman, S. Doraisamy, N. Mustapha, and M. Alkeshr, "Improvement of sentiment analysis based on clustering of word2vec features," in *2017 28th International Workshop on Database and Expert Systems Applications (DEXA)*, Lyon, France, 2017, pp. 123–126.
- [16] H. Tao, J. Zhou, S. A. B. Zadeh, O. A. Adetiloye, M. Li, I. Uhumuavbi, and C. Iwendi, "Evaluation of text classification using support vector machine compare with naive bayes, random forest decision tree and k-nn," in *Springer International Publishing*, n.d.
- [17] A. Supian, B. T. Revaldo, N. Marhadi, Rahmaddeni, and L. Efrizoni, "Penerapan svm dan word2vec untuk analisis sentimen ulasan pengguna aplikasi dana," *jikstik*, vol. 23, no. 3, pp. 321–332, Sep. 2024.
- [18] S. A. P., "Analisis sentimen berbasis aspek dengan pendekatan machine learning menggunakan dataset bahasa indonesia," 2022, [Online]. Available: <https://etd.repository.ugm.ac.id/penelitian/detail/209326>.
- [19] M. Hitesh, V. Vaibhav, Y. J. A. Kalki, S. H. Kamtam, and S. Kumari, "Real-time sentiment analysis of 2019 election tweets using word2vec and random forest model," in *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, Jaipur, India, 2019, pp. 146–151.