



INNOVATIVE: Journal Of Social Science Research

Volume 3 Nomor 4 Tahun 2023 Page 2173-2185

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

## Analisis Sentimen Ulasan Pengguna Klik Indomaret Pada Google Play Menggunakan Support Vector Machine

Dina Amelia Alzahra<sup>1✉</sup>, Ultach Enri<sup>2</sup>, Yuyun Umaidah<sup>3</sup>

Prodi Informatika, Universitas Singaperbangsa Karawang

Email: [dina.amelia19016@student.unsika.ac.id](mailto:dina.amelia19016@student.unsika.ac.id)

### Abstrak

Tingginya persaingan retail di Indonesia yang didukung dengan perkembangan teknologi, menjadi salah satu pemicu Indomaret untuk meluncurkan sebuah aplikasi yang memungkinkan para penggunanya untuk menikmati beragam kemudahan dalam berbelanja kebutuhan harian. Aplikasi dengan *tagline one stop online store* ini telah terunduh sebanyak 5 juta kali dan memiliki 99.000 ulasan di *Google Play*. Ulasan pengguna pada *Google Play* dapat dimanfaatkan untuk peningkatan kualitas aplikasi, pelayanan, atau bahkan produk yang disediakan. Oleh karena itu, dilakukan analisis sentimen yang bertujuan untuk memahami, mengekstrak, dan mengelola ulasan pengguna berupa data tekstual yang kemudian diketahui apakah sebuah ulasan memiliki kecenderungan pada nilai positif dan negatif. Adapun algoritme yang digunakan adalah *Support Vector Machine* (SVM) dengan membandingkan 4 kernel, diantaranya adalah *linear*, *polynomial*, *Sigmoid*, dan *RBF*. Metodologi yang digunakan pada penelitian ini adalah *Knowledge Discovery Database* (KDD) yang terdiri dari 5 tahapan yaitu *data selection*, *preprocessing*, *transformation*, *data mining*, dan terakhir *evaluation*. Data yang digunakan berjumlah 1563 yang terdiri dari 694 ulasan positif dan 869 ulasan negatif. Pada tahapan *transformation* dilakukan perbandingan antara TF-IDF dan *Bag of Words*. Terdapat tiga skenario pada tahapan *data mining*, yang membagi data *train* dan *test* menjadi 70:30, 80:90, dan 90:10. Hasil terbaik dimiliki oleh kernel *Sigmoid* menggunakan fitur ekstraksi *Bag of Words* dan rasio perbandingan data 90:10, dengan hasil *accuracy* 92%, *precision* 89%, *recall* 96%, dan *F1-Score* 92%.

Kata kunci: *Analisis Sentimen, Support Vector Machine, Klik Indomaret.*

## Abstract

The intense competition in the retail sector in Indonesia, supported by technological advancements, has become one of the driving factors for Indomaret to launch an application that allows its users to enjoy various conveniences in their daily shopping needs. This one-stop online store application has been downloaded 5 million times and has received 99,000 reviews on Google Play. User reviews on Google Play can be utilized to improve the quality of the application, services, or even the provided products. Therefore, sentiment analysis is conducted with the aim of understanding, extracting, and managing user reviews in the form of textual data to determine whether a review tends to be positive or negative. The algorithm used for this analysis is Support Vector Machine (SVM), comparing four kernels, namely linear, polynomial, Sigmoid, and RBF. The methodology employed in this research is Knowledge Discovery Database (KDD), consisting of five stages: data selection, preprocessing, transformation, data mining, and evaluation. The dataset used comprises 1563 reviews, consisting of 694 positive reviews and 869 negative reviews. During the transformation stage, a comparison is made between TF-IDF and Bag of Words. There are three scenarios in the data mining stage, where the data is divided into training and testing sets with ratios of 70:30, 80:90, and 90:10. The best result is achieved by the Sigmoid kernel using Bag of Words as the feature extraction method and a data ratio of 90:10, with an accuracy of 92%, precision of 89%, recall of 96%, and F1-Score of 92%.

Keywords: *Sentiment Analysis, Support Vector Machine, Klik Indomaret*

## PENDAHULUAN

Pada tahun 2015, Indomaret meluncurkan sebuah aplikasi *e-commerce* yang dikenal sebagai Klik Indomaret. Aplikasi tersebut memungkinkan pengguna untuk berbelanja berbagai macam produk yang dibutuhkan dengan beragam kemudahan. Akan tetapi hingga 28 maret 2023 Klik Indomaret berada di bawah pesaingnya, yaitu Alfagift yang merupakan aplikasi milik Alfamart di bawah naungan Alfa Group. Hal tersebut dibuktikan dengan data yang diambil pada laman similarweb.com terkait dengan *Top Apps Ranking* pada kategori *Most Popular Shopping Apps in Indonesia*, menunjukkan bahwa Klik Indomaret pada tanggal 28 Maret 2023 berada di bawah Alfagift.

Tabel 1. Most Popular Shopping Apps in Indonesia

<i>App</i>	<i>Publisher</i>	<i>Usage Rank</i>	<i>Change</i>	<i>Store Rank</i>	<i>Change</i>
Alfagift	Alfagift	5	=	6	=
Klik Indomaret	PT. Indomaret Perkasa	13	=	12	=

Sumber: similarweb.com

Tabel 1 menunjukkan informasi bahwa Klik Indomaret berada di peringkat ke-13 dalam *Usage*

*Rank* yang meliputi instalasi dan pengguna aktif pada dua bulan terakhir. Sedangkan untuk kategori *Store Rank* yang diambil dari peringkat aplikasi pada Google Play, Klik Indomaret menduduki peringkat ke-12. Berbeda dengan Alfacit yang secara berurutan menduduki peringkat lima dan enam pada kedua kategori tersebut.

Kondisi Klik Indomaret tersebut tentunya dapat dipengaruhi oleh beberapa faktor, salah satunya adalah kurangnya loyalitas pelanggan terhadap aplikasi tersebut. Fitur ulasan pada Google Play dapat digunakan untuk mengetahui respon dan tanggapan pengguna terhadap aplikasi. Ulasan tersebut dapat dimanfaatkan sebagai salah satu pertimbangan untuk peningkatan dan perbaikan baik untuk pelayanan maupun yang berhubungan dengan produk yang disediakan.

Untuk memahami, mengekstrak, dan mengolah ulasan pengguna yang berupa tekstual, digunakan sebuah proses yang dikenal dengan analisis sentimen. Proses tersebut dilakukan sebagai upaya untuk melihat pendapat atau opini yang terkandung pada ulasan pengguna terhadap objek yang pada akhirnya mempunyai nilai kecenderungan apakah bernilai positif atau negatif (Somantri & Apriliani, 2018).

Untuk mengetahui kecenderungan suatu teks pada nilai positif dan negatif, dapat dilakukan pengklasifikasian teks, *Support Vector Machine* (SVM) merupakan salah satu algoritme klasifikasi yang dapat digunakan untuk analisis sentimen. Pada penelitian yang dilakukan oleh (Ilmawan & Mude, 2020) terkait perbandingan metode klasifikasi *Support Vector Machine* dan *Naïve Bayes* untuk Analisis Sentimen khususnya pada data tekstual dari Google Play, didapatkan hasil bahwa *Support Vector Machine* memiliki tingkat akurasi lebih tinggi yaitu sebesar 81,46%, dibanding dengan *Naïve Bayes* yang menghasilkan akurasi sebesar 75,41%.

Kemudian pada penelitian yang membandingkan algoritme *Support Vector Machine*, K-NN, dan *Naïve Bayes* untuk analisis sentimen tanggapan masyarakat Indonesia pada media sosial twitter terkait Covid-19 oleh (Pamungkas & Kharisudin, 2021). *Support Vector Machine* mempunyai akurasi yang lebih tinggi jika dibandingkan dengan K-NN dan *Naïve Bayes*. *Support Vector Machine* menghasilkan akurasi sebesar 90,1%, *Naïve Bayes* dengan 79,2% sedangkan K-NN (k=2 dan 6) menghasilkan akurasi sebesar 62,1%.

Pada penelitian lainnya yang membandingkan algoritme *Support Vector Machine* dengan *Decision Tree* juga *Logistic Regression* oleh (Nurul Hassanah et al., 2023) dan (Putri & Kharisudin, 2022) kedua penelitian tersebut memberikan kesimpulan bahwa *Support Vector Machine* menghasilkan kinerja klasifikasi yang baik, dibuktikan dengan hasil akurasi yang paling tinggi jika dibandingkan dengan algoritme *Decision Tree* dan *Logistic Regression*.

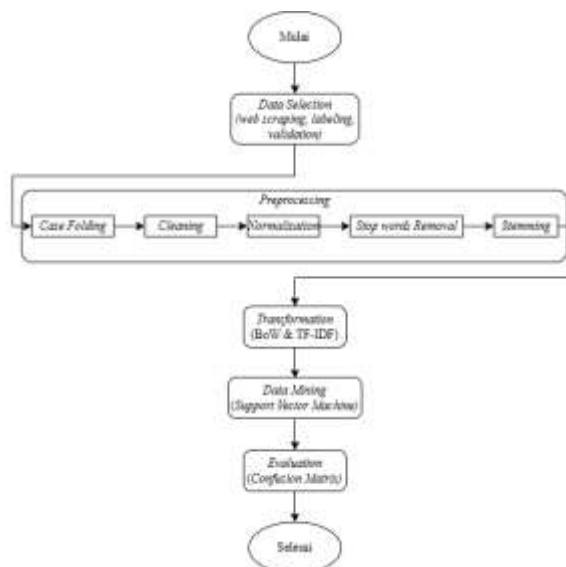
Dari beberapa penelitian di atas, dapat disimpulkan bahwa *Support Vector Machine* mempunyai tingkat akurasi yang cenderung lebih tinggi jika dibandingkan dengan beberapa algoritme klasifikasi yang ada. Akan tetapi kinerja *Support Vector Machine* bergantung pada

kernel (Rianti et al., 2021) sehingga pemilihan kernel sangat penting dilakukan untuk menghasilkan akurasi terbaik. Akan tetapi belum ada kesimpulan yang menyatakan kernel mana yang paling baik pada kasus tertentu (Aulia et al., 2021). Adapun beberapa kernel yang biasanya digunakan pada *Support Vector Machine* adalah *Linear*, *Polynomial*, *Sigmoid*, dan *Radial Basis Function* (RBF) (Hendrastuty et al., 2021). Untuk itu, penulis akan membandingkan beberapa kernel yang umum digunakan pada *Support Vector Machine* guna mengetahui kernel terbaik yang menghasilkan akurasi paling tinggi.

Dari pembahasan latar belakang yang ada, penulis melakukan analisis sentimen pada ulasan Klik Indomaret pada Google Play dengan melakukan pengklasifikasian teks menggunakan algoritme *Support Vector Machine*. Selain itu penulis juga melakukan perbandingan pada kernel *Linear*, *Polynomial*, *Sigmoid*, dan *Radial Basis Function* (RBF) untuk mengetahui akurasi terbaik pada pengklasifikasian teks.

## METODE PENELITIAN

Metodologi yang digunakan pada penelitian ini adalah *Knowledge Discovery in Database* (KDD) yang terdiri dari lima tahapan, yaitu *Selection*, *Preprocessing*, *Transformation*, *Data Mining* dan *Evaluation*. Gambar berikut merupakan alur dari KDD yang akan diaplikasikan pada penelitian ini.



Gambar 1 *Knowledge Discovery in Database*

### 1. *Data Selection*

Data diambil melalui proses *web scraping* pada platform Google Play. Data yang diambil merupakan ulasan pengguna berbahasa Indonesia pada aplikasi Klik Indomaret dengan rentan waktu antara 1 Februari 2023 hingga 25 April 2023. Setelah itu data akan diklasifikasikan ke dalam kelas negatif dan positif, dan selanjutnya divalidasi oleh ahli yang bertujuan untuk meminimalisir kesalahan dalam proses klasifikasi.

### 2. *Preprocessing*

Data yang sudah ada kemudian diolah pada tahapan *preprocessing* yang bertujuan untuk mengubah data mentah menjadi data terstruktur siap pakai dengan melewati serangkaian proses, diantaranya adalah *case folding*, *cleaning*, *normalization*, *stopwords removal*, dan *stemming*.

### 3. *Transformation*

Pada tahapan ini dilakukan pembobotan kata dengan menggunakan TF-IDF *Bag of Words*. Setiap kata-kata yang lebih sering muncul dalam satu data akan diberikan prioritas lebih tinggi dan mempunyai bobot yang lebih tinggi karena dianggap lebih berguna untuk klasifikasi.

### 4. *Data Mining*

Pada tahap ini akan dilakukan proses klasifikasi sentimen pada data hasil processing ke dalam kelas negatif dan positif dengan menggunakan algoritme Support Vector Machine dan dilakukan perbandingan kernel Linear, Polynomial, Sigmoid, dan Radial Basis Function (RBF) untuk mengetahui akurasi paling baik. Adapun penggunaan parameter pada penelitian ini disesuaikan dengan tiap kernel ((Linear: cost =1), (Polynomial: cost=1, gamma= scale/auto, coefficient=0.0, degree=3), (Sigmoid: cost=1, gamma= scale/auto, coefficient=0.0), (RBF= cost=1, gamma= scale/auto)).

### 5. *Evaluation*

Selanjutnya dilakukan tahapan pengujian dengan menggunakan *Confusion Matrix*. Hasil evaluasinya akan berupa nilai akurasi, presisi, *recall*, dan *F1-Score*.

## HASIL DAN PEMBAHASAN

### A. Hasil

Alur dari penelitian ini akan mengikuti tahapan yang ada pada metodologi *Knowledge Discovery Database* (KDD) yang terdiri dari lima tahapan, diantaranya adalah sebagai berikut:

#### 1. *Data Selection*

Data diambil dengan menggunakan teknik *web scraping* pada laman Google Play. Data yang diambil dari rentan waktu 1 Februari 2023 hingga 25 April 2023 sebanyak 1750 data, yang kemudian diseleksi dengan menghapus beberapa data yang hanya mengandung bahasa asing, dan kata yang tidak relevan lainnya, sehingga data akhir yang tersisa sebanyak 1563. Selanjutnya data diberikan label dan diklasifikasikan ke dalam kelas negatif (0) atau positif (1), dimana hasil dari klasifikasi tersebut kemudian divalidasi oleh ahli untuk meminimalisir kesalahan. Hasil dari proses validasi tersebut kemudian akan digunakan pada tahapan selanjutnya. Tabel 2 merupakan perbandingan antara jumlah data negatif dan positif dari pelabelan awal dan hasil validasi:

Tabel 2 Perbandingan Label dan Validasi

Label Ulasan	Pelabelan Awal	Validasi Data
Negatif	696	694
Positif	867	869
Total	1563	1563

## 2. *Preprocessing*

Terdapat enam proses yang dilalui pada proses ini, diantaranya adalah *case folding* yang bertujuan untuk menyeragamkan bentuk kalimat atau teks dengan mengubah seluruh kalimat menjadi huruf kecil. Pada proses *Cleaning* data dibersihkan dari berbagai macam karakter yang dianggap tidak dibutuhkan seperti tanda baca, tautan, angka, dan karakter lainnya. Kemudian dilakukan perubahan kata ke dalam bentuk awal seperti penggantian kata non baku ke dalam kata baku, beberapa kata bahasa Inggris ke dalam bahasa Indonesia, dan kata singkatan pada proses *normalization*. Selain itu, kata yang sering muncul pada kalimat namun dianggap tidak bermakna dan tidak mempunyai pengaruh yang besar pada proses sentimen akan dihapus di *stopword removal*, sedangkan proses *stemming* akan menghapus imbuhan yang ada pada kata. Berikut merupakan perbedaan antara data awal dengan data yang sudah melewati *preprocessing*.

Tabel 3 Hasil Preprocessing

Data Awal	Hasil <i>Stemming</i>
Lama, udah psn nungguin lama gk dtg udah di email juga ttp lamaa gabisa di batalin, beli 4 produk dr jam 12 eh jam 3 baru di konfirmasi gaada 1 barang	pesan tunggu surat elektronik batal beli produk jam eh jam konfirmasi barang

## 3. *Transformation*

Tahapan selanjutnya adalah *transformation* guna menghitung bobot kemunculan kata dengan menggunakan dua teknik yaitu TF-IDF dan *Bag of Words*. TF-IDF akan menghitung bobot kata berdasarkan frekuensi kemunculannya, kemudian dihitung seberapa penting

sebuah kata dalam seluruh Kumpulan dokumen dengan cara membagi total jumlah dokumen dengan jumlah dokumen yang mengandung kata tersebut. Tabel 4 merupakan contoh hasil pembobotan dari TF-IDF.

Tabel 4 Hasil TF-IDF

index	abal	abang	abis	acara
1233	0.0	0.0	0.0	0.0
50	0.0	0.0	0.0	0.0
261	0.0	0.0	0.0	0,323
1373	0.0	0.0	0,652	0.0
1240	0.0	0.0	0,587	0.0
1259	0.0	0,4623	0.0	0.0
1088	0,764	0.0	0.0	0.0

Dari pembobotan di atas dihasilkan sebuah matriks dimana terdapat 1518 fitur yang terurut sesuai dengan abjad. Dimana fitur tersebut akan dicocokkan pada data untuk dilihat frekuensi kemunculannya. Dari tabel di atas, dinyatakan bahwa semakin sering kata muncul pada teks, maka memiliki nilai yang cenderung kecil. sedangkan kata yang jarang muncul memiliki nilai yang cenderung lebih besar. Kata acara dengan nilai 0,323 dianggap lebih penting jika dibandingkan dengan kata abal yang memiliki nilai 0,764. Hal tersebut dikarenakan kata acara memiliki frekuensi kemunculan yang lebih sedikit apabila dibandingkan dengan kata abal. Adapun apabila term bernilai 0, maka kata tersebut tidak muncul dalam dokumen.

Sama halnya dengan *Bag of Words* (BoW) akan mengubah teks menjadi vektor agar bisa dipahami oleh komputer, BoW akan menghitung frekuensi kemunculan suatu kata pada keseluruhan dokumen. Tabel 5 merupakan hasil pembobotan kata dengan menggunakan BoW.

Tabel 5 Hasil BoW

index	abal	abang	abis	acara
1233	0	0	0	0
50	0	0	0	0
261	0	0	0	1
1240	0	0	1	0
1373	0	0	1	0
1259	0	1	0	0
1088	2	0	0	0

Kata abal pada dokumen 1088 dihitung memiliki frekuensi yang lebih besar jika dibandingkan dengan kata abang pada dokumen 1259 yang bernilai 1. Semakin tinggi nilainya, maka semakin sering juga sebuah kata muncul dalam sebuah dokumen. Adapun apabila kata bernilai 0, maka kata tersebut tidak muncul dalam satu dokumen tersebut.

#### 4. Data Mining

Pada tahapan Data Mining digunakan algoritme *Support Vector Machine* dengan membandingkan empat kernel untuk mengetahui perbandingan dari hasil akurasi, diantaranya adalah *Linear*, *Polynomial*, *Sigmoid*, dan RBF. Adapun parameter yang digunakan disesuaikan dengan tiap-tiap kernel. Kemudian, semua parameter di atur ke dalam *default*. Terdapat total tiga skenario pembagian data, skenario satu membagi data menjadi 70% data *train* dan 30% data *test*, skenario dua 80% data *train* dan 20% data *test*, terakhir skenario tiga membagi data menjadi 90% data *train* dan 10% data *test*. Tabel 6 merupakan pembagian antara data *train* dan *test*.

Tabel 6 Pembagian Data

Data	Skenario		
	70:30	80:20	90:10
<i>Train</i>	1094	1250	1406
<i>Test</i>	469	313	157
Jumlah	1563		

#### 5. Evaluation

Tahap terakhir dari penelitian ini adalah evaluasi, yang bertujuan untuk mengetahui akurasi, presisi, *recall*, dan *f1-score* dari model dengan menggunakan *Confusion Matrix*. Berikut merupakan hasil evaluasi dari TF-IDF:

Tabel 7 Evaluasi Menggunakan Fitur Ekstraksi TF-IDF

Fitur Ekstraksi	Rasio Perbandingan	Kernel	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
TF/IDF	70:30	<i>Linear</i>	91%	91%	93%	92%
		<i>Polynomial</i>	82%	94%	72%	82%
		<i>Sigmoid</i>	90%	90%	92%	91%
		RBF	90%	92%	89%	91%
	80:20	<i>Linear</i>	90%	88%	94%	91%
		<i>Polynomial</i>	84%	92%	76%	83%
		<i>Sigmoid</i>	89%	88%	93%	90%
		RBF	89%	90%	90%	90%

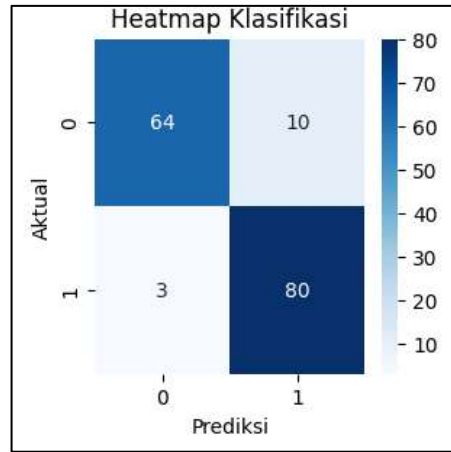
	90:10	<i>Linear</i>	91%	89%	95%	82%
		<i>Polynomial</i>	82%	90%	75%	82%
		<i>Sigmoid</i>	91%	89%	95%	92%
		RBF	90%	90%	92%	91%

Sedangkan hasil evaluasi menggunakan BoW dapat dilihat pada Tabel 8 berikut:

Tabel 8 Evaluasi Menggunakan Fitur Ekstraksi BoW

Fitur Ekstraksi	Rasio Perbandingan	Kernel	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
BoW	70:30	<i>Linear</i>	86%	83%	94%	88%
		<i>Polynomial</i>	66%	62%	98%	76%
		<i>Sigmoid</i>	89%	85%	96%	90%
		RBF	89%	90%	91%	90%
	80:20	<i>Linear</i>	83%	79%	93%	85%
		<i>Polynomial</i>	64%	60%	98%	75%
		<i>Sigmoid</i>	89%	85%	96%	90%
		RBF	89%	87%	93%	90%
	90:10	<i>Linear</i>	89%	84%	98%	90%
		<i>Polynomial</i>	67%	62%	99%	76%
		<i>Sigmoid</i>	92%	89%	96%	92%
		RBF	89%	88%	93%	90%

Dari Tabel 7 dan Tabel 8 di atas, dapat diketahui bahwa hasil akurasi tertinggi dimiliki oleh kernel *Sigmoid* dengan rasio perbandingan data 90:10 dan menggunakan fitur ekstraksi *Bag of Words*, yaitu sebesar 92%, presisi 89%, *recall* 96%, dan *F1-score* 92%. Adapun hasil akurasi terendah dimiliki oleh kernel *Polynomial* dengan rasio perbandingan 80:20 dan menggunakan fitur ekstraksi BoW, yaitu sebesar 64%, presisi 60%, *recall* 98%, dan *F1-Score* 75%.



Gambar 2 *Heatmap* Klasifikasi

Gambar 2 merupakan hasil evaluasi dari *confusion matrix* dari hasil akurasi terbesar yang dimiliki oleh kernel Sigmoid dengan rasio perbandingan data 90:10 dan menggunakan fitur ekstraksi *Bag of Words*. Terdapat 80 data yang diprediksi positif pada kelas positif, dan 4 data yang tidak diprediksi negatif pada kelas positif. Kemudian terdapat 64 data yang diprediksi negatif pada kelas negatif, dan 10 data diprediksi positif pada kelas negatif. Adapun perhitungan akurasi, presisi, *recall*, dan *f1-score* secara manual dari matrix adalah sebagai berikut:

$$\begin{aligned}
 \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} = \frac{80 + 64}{80 + 64 + 10 + 3} = 0,91 \\
 \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{80}{80 + 10} = 0,8 \\
 \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{80}{80 + 3} = 0,96 \\
 \text{F1 - Score} &= 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} = 2 \times \frac{(0,89 \times 0,96)}{(0,89 + 0,96)} = 0,92
 \end{aligned}$$

## B. Pembahasan

Dari 24 percobaan dengan menggunakan dua fitur ekstraksi, tiga skenario pembagian data, dan empat kernel yang berbeda, diketahui kernel dengan akurasi paling baik yang dimiliki oleh kernel Sigmoid dengan rasio perbandingan 90:10 dan menggunakan fitur ekstraksi *Bag of Words*, yaitu sebesar 92%, presisi 89%, *recall* 96%, dan *F1-Score* 92%. Adapun hasil akurasi terendah dimiliki oleh kernel *Polynomial* dengan rasio perbandingan 80:20 dan menggunakan fitur ekstraksi yang sama yaitu BoW. Akurasi yang didapatkan sebesar 64%, presisi 60%, *recall* 98%, dan *F1-Score* 75%. Kemudian hasil klasifikasi dari sentimen yang didapatkan dari ulasan pengguna Klik Indomaret divisualisasikan ke dalam *word cloud* untuk mengetahui gambaran atau informasi, khususnya kata-kata yang sering muncul dalam dataset yang ada. Gambar 3 merupakan visualisasi dari kelas negatif dan positif



Gambar 3 Visualisasi Negatif

Terdapat beberapa kata yang sering muncul pada kelas negatif. Semakin besar ukuran katanya, maka semakin sering muncul kata tersebut muncul. Pada kelas negatif aplikasi, belanja, pesan, barang, bayar, batal, masuk, kirim, dan toko merupakan kata yang frekuensi kemunculannya paling tinggi. Dari visualisasi tersebut, dapat ditarik kesimpulan bahwa beberapa pengguna mengalami permasalahan pada proses masuk ke dalam aplikasi, barang yang mungkin tidak sesuai antara toko dan aplikasi, pembayaran yang susah, permasalahan pada pembatalan pesanan dan proses kirim barang.



Gambar 4 Visualisasi Positif

Gambar 4 merupakan visualisasi kelas positif. Kata bantu, belanja, cepat, mudah, dan bagus merupakan kata dengan frekuensi kemunculan paling tinggi. Dari visualisasi di atas, dapat ditarik kesimpulan bahwa sebagian pengguna merasa puas, aplikasi dapat membuat belanja menjadi lebih mudah dan praktis, sehingga dapat membantu para pengguna. Pelayanan terbilang ramah dan baik, serta pengiriman barang yang cepat.

### SIMPULAN

Dari uraian pembahasan tersebut di atas, maka peneliti dapat mengambil kesimpulan dari hasil penelitian yang telah dilakukan, Akurasi tertinggi dimiliki oleh kernel Sigmoid dengan rasio perbandingan 90:10 dan menggunakan fitur ekstraksi *Bag of Word*, yaitu sebesar 92%,

presisi 89%, *recall* 96%, dan F1-*Score* 92%. Adapun hasil akurasi terendah dimiliki oleh kernel *Polynomial* dengan rasio perbandingan 80:20 dan menggunakan fitur ekstraksi yang sama yaitu BoW. Akurasi yang didapatkan sebesar 64%, presisi 60%, *recall* 98%, dan F1-*Score* 75%. Pada sentimen positif dapat diketahui bahwa pengguna puas dengan layanannya yang dapat membuat pengguna berbelanja dengan cepat dan mudah. Sedangkan pada sentiment negatif dapat disimpulkan bahwa beberapa pengguna mengalami permasalahan pada proses masuk ke dalam aplikasi, barang yang mungkin tidak sesuai antara toko dan aplikasi, pembayaran yang susah, permasalahan pada pembatalan pesanan dan proses kirim barang.

#### DAFTAR PUSTAKA

- Aulia, T. M. P., Arifin, N., & Mayasari, R. (2021). Perbandingan Kernel Support Vector Machine (SVM) Dalam Analisis Sentimen Vaksinasi Covid-19. *Science And Information Technology*, 4(2), 140–145. <https://doi.org/https://doi.org/10.31598>
- Hendrastuty, N., Rahman Isnain, A., & Yanti Rahmadhani, A. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, 6(3). <http://situs.com>
- Ilmawan, L. B., & Mude, M. A. (2020). Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 12(2), 154–161. <https://doi.org/10.33096/ilkom.v12i2.597.154-161>
- Nurul Hassanah, I., Faisal, S., Mutoi Siregar, A., Buana Perjuangan Karawang Jl HSRonggo Waluyo, U., Timur, T., & Barat, J. (2023). Perbandingan Algoritma Support Vector Machine Dengan Decision Tree Pada Aplikasi Ruangguru. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 10, 39–50.
- Pamungkas, F. S., & Kharisudin, I. (2021). Analisis Sentimen dengan SVM, Naive Bayes, dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 628–634. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- Putri, M. I., & Kharisudin, I. (2022). Penerapan Synthetic Minority Oversampling Technique (SMOTE) Terhadap Analisis Sentimen Data Review Pengguna Aplikasi Marketplace Tokopedia. *PRISMA, Prosiding Seminar Nasional Matematika*, 5, 759–766. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- Rianti, D. L., Umaidah, Y., & Voutama, A. (2021). *Tren Marketplace Berdasarkan Klasifikasi Ulasan Pelanggan Menggunakan Perbandingan Kernel Support Vector Machine*.
- Somantri, O., & Apriliani, D. (2018). Support Vector Machine Berbasis Feature Selection Untuk Sentimen Analysis Kepuasan pelanggan Terhadap Pelayanan Warung Dan Restoran Kuliner Kota Tegal. *Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, 5, 537–548. <https://doi.org/10.25126/jtiik20185867>

