



Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma *Support Vector Machine* (SVM)

Diana Puspitasari¹, Tata Sutabri²

^{1,2}Magister Teknik Informatika, Universitas Bina Dama, Palembang, Indonesia

Email: ¹dianapuspita719@gmail.com, ²tata.sutabri@gmail.com

Informasi Artikel

Diterima : 27-10-2024

Disetujui : 02-11-2024

Diterbitkan : 25-11-2024

ABSTRACT

One of the social media with 24.85 million active users is Twitter. Information published on Twitter can be in the form of user opinions on an object, such as a product or service. Therefore, the company utilizes Twitter as a medium to disseminate information. This makes the company use Twitter as a medium to disseminate information. An example is an Internet Service Provider (ISP) company such as Indihome. Through Twitter, users can discuss their complaints and satisfaction with Indihome services. A method is needed, namely sentiment analysis to understand whether the textual data includes neutral opinion, negative opinion or positive opinion. So, the authors use the Support Vector Machine (SVM) method in sentiment analysis of Indihome service user opinions on Twitter, with the aim of getting a sentiment classification model using SVM and to find out how much accuracy is produced by the SVM method applied to sentiment analysis and to find out how satisfied Indihome service users are based on Twitter. After testing with the SVM method the results are 91% accuracy. Precision 51% Recall 75% and F1-Score 59%.

Keyword: *Indihome, Twitter, Sentimen Analysis, Support Vector Machine*

ABSTRAK

Salah satu media sosial dengan pengguna aktif sebesar 24,85 juta pengguna adalah *Twitter*. Informasi yang dipublikasikan di *Twitter* dapat berupa opini pengguna terhadap suatu objek, seperti produk atau layanan. Oleh karena itu, Perusahaan memanfaatkan *Twitter* sebagai media untuk menyebarkan informasi. Hal ini membuat Perusahaan menggunakan *Twitter* sebagai media untuk menyebarkan suatu informasi. Contohnya adalah Perusahaan *Internet Service Provider (ISP)* seperti *Indihome*. Melalui *Twitter* pengguna dapat saling berdiskusi mengenai keluhan maupun kepuasan mereka terhadap layanan *Indihome*. Dibutuhkan suatu

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

metode yaitu *Analisis Sentimen* untuk memahami apakah data tekstual tersebut termasuk opini netral, opini negatif atau opini positif. Maka, penulis menggunakan algoritma *Support Vector Machine (svm)* dalam *analisis sentiment* tentang opini pengguna layanan *Indihome* pada *Twitter*, dengan tujuan untuk mendapatkan model klasifikasi sentiment menggunakan SVM dan untuk mengetahui seberapa besar akurasi yang dihasilkan oleh metode SVM yang diterapkan pada *analisis sentiment* serta untuk mengetahui seberapa puas pengguna layanan *Indihome* berdasarkan pada *Twitter*. Setelah dilakukan pengujian dengan metode *Support Vector Machine* maka didapatkan hasil nilai *accuracy* 91%, *Precision* 51%, *Recall* 75% dan *F1-Score* 59.

Kata Kunci: Indihome, Twitter, Analisis Sentimen, *Support Vector Machine*

1. PENDAHULUAN

Perkembangan teknologi pada saat ini sangatlah pesat, terutama berkaitan dengan teknologi informasi yaitu internet. Internet menjadi salah satu teknologi yang sering digunakan untuk menerima ataupun memberi suatu informasi secara mudah dan dalam waktu yang cepat dari seluruh penjuru dunia (Anggraini and Sutabri, 2024).

Pada awal 2024, jumlah pengguna internet di Indonesia mencapai 221,5 juta jiwa atau 79,5% dari total penduduk Indonesia sebanyak 282,4 juta jiwa. Jumlah ini meningkat 1,4% dibandingkan tahun sebelumnya yang sebesar 78,19%. Tentunya, jumlah ini akan terus bertambah seiring dengan berjalannya waktu, begitu pula dengan perusahaan-perusahaan yang akan berlomba-lomba untuk bergerak pada bidang usaha teknologi informasi yang menyediakan jasa layanan internet. Penggunaan internet dapat diakses menggunakan beberapa provider internet yang telah tersedia. Salah satu provider internet di Indonesia adalah *Indihome (Indonesia Digital Home)* yang mana merupakan salah satu produk layanan yang dimiliki oleh PT. Telekomunikasi Indonesia (Sulistiani and Hamka, 2024).

Salah satu media sosial dengan jumlah pengguna aktif di Indonesia mencapai 24,85 juta pengguna pada 2024 adalah Twitter. Twitter merupakan media sosial yang digunakan untuk mencari informasi tentang bisnis, hiburan, ekonomi, politik, dan lain sebagainya.

Indihome merupakan operator penyedia ISP yang paling banyak digunakan oleh masyarakat Indonesia, dimana sebanyak 8,7% memilih Indihome sebagai operator penyedia ISP yang mereka gunakan dan juga sebagai layanan digital yang menyediakan internet, telepon rumah dan TV Interaktif (Indihome TV) dengan beragam pilihan paket serta layanan tambahan yang bisa dipilih sesuai kebutuhan. Melalui media sosial Twitter, pengguna dapat saling berdiskusi mengenai kritik, saran ataupun kepuasan mereka terhadap layanan Indihome. Perbedaan Twitter dari media sosial yang lainnya adalah

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

Twitter memungkinkan tiap pengguna dapat melacak tweet pengguna lain tanpa persetujuan dari pengguna tersebut.

Pandangan masyarakat atau umpan balik selalu terbukti menjadi sumber daya yang paling penting dan sangat berharga bagi perusahaan. Sehingga berdasarkan tweet yang mereka publikasi, dapat dilihat gambaran mengenai opini pengguna melalui proses penggalian informasi lebih lanjut dari setiap tweet yang dipublikasi, akan tetapi jumlah data tersebut banyak sehingga dibutuhkan sebuah metode untuk dapat mengubah data tersebut menjadi sebuah informasi atau pengetahuan. Suatu metode yang dapat digunakan untuk yaitu dengan menggunakan analisis sentimen (Ramlan, Satyahadewi and Andani, 2023).

Analisis sentimen merupakan suatu metode dalam memahami, mengekstraksi data sentimen yang biasanya akan dikategorikan berdasarkan polaritasnya, apakah positif atau negatif dengan menggunakan teknik *Natural Language Processing (NLP)*. Pada saat ini analisis sentimen menjadi salah satu topik panas yang banyak digunakan oleh para peneliti, yang memiliki tujuan untuk menyediakan informasi dari sebuah dataset yang tidak terstruktur. Dari hasil tersebut akan menjadi sebuah informasi yang memiliki makna seperti penilaian pada sebuah brand, pemberian rating, dan penyaringan opini untuk membantu perusahaan ataupun juga masyarakat.

Banyaknya ulasan yang telah diberikan oleh pelanggan terhadap *Indihome* melalui platform Twitter belum mampu untuk memberikan keterangan-keterangan terkait kepuasan atau problematika yang dialami oleh pelanggan itu sendiri. Oleh karena itu, diperlukan sebuah bantuan pengklasifikasian data ulasan tersebut dengan melakukan analisis sentimen pada *text mining* melalui metode yang mampu mengklasifikasikan secara akurat (Saputra and Sutabri, 2024).

Seperti pada penelitian analisis sentimen yang dilakukan oleh siti sumayah,dkk pada tahun 2023 dengan objek opini tentang metaverse yang ramai diperbincangkan. Penelitian tersebut bertujuan untuk mengetahui mengenai tanggapan masyarakat Indonesia terhadap Metaverse dengan *Algoritma Support Vector Machine (SVM)* yang memiliki hasil *accuracy* sebesar 81%, *precision* 79%, *recall* 63%, *f1-score* 57% (Sumayah *et al.*, 2023).

Penelitian analisis sentimen pada ulasan aplikasi KAI access yang dilakukan oleh Gracia Radiana dan Adi Nugroho pada tahun 2023. Penelitian tersebut bertujuan untuk mengetahui sentimen dari sebuah produk mobile. Opini yang terkait dengan Aplikasi KAI Access yang dapat digunakan PT Kereta Api Indonesia sebagai parameter kunci untuk mengetahui tingkat kepuasan publik sekaligus bahan evaluasi bagi PT Kereta Api Indonesia dengan *metode Support Vector Machine*. penelitian tersebut memperoleh nilai *accuracy*, *precision*, *recall*, dan *F measure* yang dihasilkan dari tiap aspek yaitu untuk *Learnability* 94.73%, 100.00%, 89.50%, dan 94.64%, *Efficiency* 94.38%, 72.00%, 100.00%, dan 94.46%, *Errors* 85.13%, 97.11%, 72.41%, dan 82.96%, *Satisfaction* 87.26%, 98.46%, 73.78%, dan 84.20% (Radiana *et al.*, 2023).

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

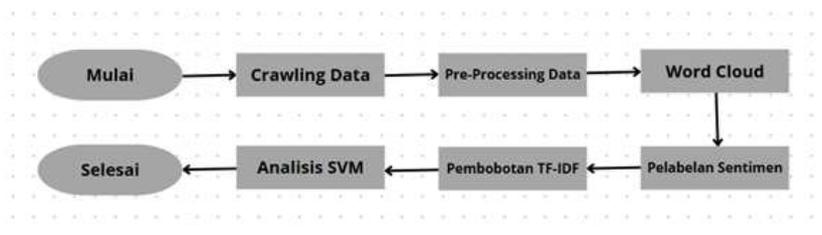
Penelitian lain mengenai analisis sentimen dengan objek opini pengguna layanan internet provider yang dilakukan oleh Fadhilah Dwi Ananda dan Yoga Pristyanto pada tahun 2021. Penelitian tersebut bertujuan untuk mengetahui persentase sentimen pengguna internet Biznet dengan Support Vector Machine, memiliki hasil pengujian melalui 3 skenario pengujian, *accuracy* untuk Kernel linear secara berurutan yaitu 76.25 %, 84.44 % dan 90 % yang terus meningkat sedangkan pada kernel RBF hasil *accuracy* yang didapat lebih kecil yaitu 76.25 %, 85.55 % dan 88 %. Untuk hasil nilai *precision* secara berurutan SVM kernel linear menghasilkan nilai 86.84%, 90.47% dan 88.63 %, sedangkan pada kernel RBF menghasilkan nilai *precision* 86.84 %, 90.69 % dan 84.78. Sedangkan untuk nilai *recall* secara berurutan SVM kernel linear menghasilkan nilai 70.21 %, 79.16 % dan 88.63 % sedangkan pada kernel RBF menghasilkan nilai *recall* 70.21 %, 81.45% dan 88.63 % (Ananda and Pristyanto, 2021).

Berdasarkan latar belakang diatas dan juga berpedoman pada penelitian sebelumnya, maka didapat rumusan masalah yaitu bagaimana kinerja Algoritma *Support Vector Machine (SVM)* dalam klasifikasi sentimen pengguna *Twitter* terhadap layanan *Indihome*. Adapun tujuan dari penelitian ini adalah untuk mendapatkan pemodelan klasifikasi sentimen dengan *Support Vector Machine (SVM)*, untuk mengetahui seberapa besar akurasi yang dihasilkan oleh algoritma *SVM*, serta untuk mengetahui tingkat kepuasan pengguna layanan *Indihome* berdasarkan analisis sentimen *Twitter* dengan menggunakan Algoritma *Support Vector Machine*.

2. METODE

2.1. Alur Penelitian

Penelitian merupakan kumpulan fakta atau data yang didapat dari suatu objek penelitian. Dalam hal ini objek penelitian yang digunakan adalah data tweet mengenai *Indihome* yang kemudian datanya akan dijadikan sebagai sumber data. Sistem klasifikasi yang akan dilakukan untuk analisis sentimen ini memiliki suatu rancangan tentang bagaimana alur sistem ini akan berjalan. Dalam melakukan penelitian ini diperlukan suatu alur proses penelitian agar proses penelitian berjalan sesuai dengan rencana. Gambaran umum yang menunjukkan diagram alur proses penelitian dalam penelitian ini dapat dilihat pada Gambar 1 dibawah ini.



Gambar 1. Diagram Alur Proses Penelitian

Dapat dilihat pada Gambar 1. Terdapat beberapa tahap dalam penelitian diantaranya tahap crawling data, pre-prosesing data, Word Cloud, pelabelan sentimen, pembobotan TF-IDF, dan terakhir tahap Analisis SVM atau evaluasi. Sistem diawali dengan

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

menginputkan data tweet dengan query *Indihome* menggunakan *Twitter* yang terkoneksi dengan *API* dengan menggunakan Bahasa pemrograman Python dan bantuan Library *Tweepy*, kemudian data tersebut dilakukan preprocessing. Setelah dilakukan preprocessing kemudian dilakukan proses pembobotan dengan pembobotan TF-IDF. Lalu dilakukan perhitungan dengan menggunakan metode Support Vector Machine (SVM) sebagai algoritma klasifikasi.

2.2. Crawling Data

Crawling digunakan sebagai pengumpulan data untuk dataset (Herlambang and Avianto, 2023). Data yang diperoleh yaitu dari cuitan Twitter dengan query Indihome menggunakan Twitter yang terkoneksi dengan Twitter Application Programming Interface (*API*), untuk mendapatkan akses data diperlukan keys dan token. Dalam proses crawling data memakai python sebagai bahasa pemrograman dan *Google Colab* untuk mengeksekusi kode python arbitrer melalui browser. Setelah itu dataset tweet yang masih mentah disimpan dalam bentuk CSV (*Comma Separated Values*) yang nantinya akan di proses lebih lanjut dengan bantuan library yaitu *tweepy* (Petiwi, Triayudi and Sholihati, 2022).

2.3. Pre-Processing Data

Data ulasan pada penelitian ini merupakan data tweet atau cuitan yang diambil dari platform Twitter berupa teks yang tidak terstruktur, dikarenakan masih terdapat banyak sekali *noise*, sehingga sebelum melakukan tahapan klasifikasi data harus diubah menjadi lebih terstruktur. Data akan menjadi lebih terstruktur dan seragam dengan melalui tahapan proses sebagai berikut :

- 1) Cleaning digunakan untuk menghilangkan tanda baca, username, link dan hashtag.
- 2) Case Folding digunakan untuk menyeragamkan kata yang terdapat pada dataset untuk menjadi huruf kecil.
- 3) Stemming dalam penelitian digunakan untuk mengganti kalimat dalam dataset menjadi kata dasar dengan meniadakan kata tambahan atau imbuhan.
- 4) Tokenizing dalam penelitian berfungsi sebagai pemotongan pada kalimat yang berada pada dataset jadi perkata.
- 5) Text Normalization digunakan untuk mengganti kata singkat jadi kata baku.
- 6) Stopwords untuk meniadakan kata yang tidak berfungsi tapi seringkali terdapat pada tweet (Ilmiah and Pendidikan, 2024)

2.4. Word Cloud

Word Cloud merupakan bentuk visual dari keberadaan kata. semakin banyak istilah yang sering muncul dalam dokumen yang dianalisis, maka kata yang muncul pada gambar tentu akan semakin besar (Wati *et al.*, 2021). *Word Cloud* akan diproses setelah tahapan dalam pre-processing selesai. Pada penelitian ini akan diketahui kata-kata yang banyak diperbincangkan mengenai Indihome (Cantika Larasati, Dewi and Juli, 2024).

2.5. Pelabelan Sentimen

Labeling Sentimen pada tahap ini Bertujuan untuk mengetahui pendapat masyarakat Indonesia mengenai Layanan Indihome. pelabelan atau Labeling ialah proses pemberian terhadap kelas berdasarkan dari ciri atau karakteristik yang terkandung di dalam dokumen atau kalimat. Hal ini karena tujuan utama dari analisis sentimen yaitu untuk mengidentifikasi apakah data atau teks memiliki sentimen positif, negatif, atau netral (Alviani, Alam and Kurniawan, 2023). Penandaan dilakukan secara otomatis sehingga dapat memudahkan pengelolaan data dengan membaca arti kalimat dalam suatu cuitan untuk menandai opini publik tersebut. Pada proses ini pelabelan data menggunakan pustaka NLTK. Pustaka *NLTK (Natural Language Toolkit)* adalah salah satu pustaka Python yang dirancang untuk mendukung pemrosesan bahasa alami (*Natural Language Processing/NLP*). NLTK menyediakan berbagai alat, data, dan sumber daya untuk membantu para pengembang, peneliti, dan pelajar dalam menganalisis dan memahami teks dalam berbagai bahasa. NLTK digunakan untuk menganalisis sentimen teks. Hasilnya akan menampilkan sentimen dari teks dalam bahasa Indonesia beserta score sentimen. Setelah mendapatkan score sentiment, Langkah selanjutnya adalah mengklasifikasi score sentiment tersebut ke dalam kategori negatif, netral, atau positif. Pada tahap ini label kelas opini publik terhadap Layanan Indihome dipisahkan menjadi tiga kategori yaitu label negatif, netral, dan positif. Apabila kalimat bernilai >0 berada pada klasifikasi kelas positif, apabila kalimat bernilai $= 0$ berada pada klasifikasi kelas netral, lalu apabila kalimat bernilai < 0 berada pada klasifikasi kelas negatif (Alexander, Bria and Witanti, 2024).

2.6. Pembobotan TF-IDF

Pembobotan terhadap *Term Frequency-Inverse Document Frequency (TF-IDF)*, yakni gabungan dua proses dari TF dan IDF (Syahputra, Rahayudi and Muflikhah, 2022). TF-IDF digunakan saat kita ingin merubah atau transformasi data teks menjadi vektor (angka). TF untuk mengetahui jumlah pada suatu kata yang sering muncul pada kalimat atau bahasan sedangkan IDF mengukur seberapa penting suatu kata dalam dokumen. Berikut ini merupakan rumus pembobotan kata TF-IDF (Robison Manalu *et al.*, 2022) :

$$TF_{t,d} = \begin{cases} 1 + \log_{10} tf_{t,d} & > 0 \\ 0 & < 0 \end{cases} \quad (1)$$

$$IDF_t = \log \left(\frac{N}{df_t} \right) \quad (2)$$

$$TFIDF_{t,d} = TF_{t,d} \times IDF_t \quad (3)$$

Keterangan:

$TF_{t,d}$ = Nilai kemunculan term

IDF_t = Pendistribusian term pada seluruh dokumen

$F_{t,d}$ = Nilai kemunculan term pada dokumen

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

N = Jumlah seluruh dokumen

dft = Jumlah dokumen yang terdapat term d

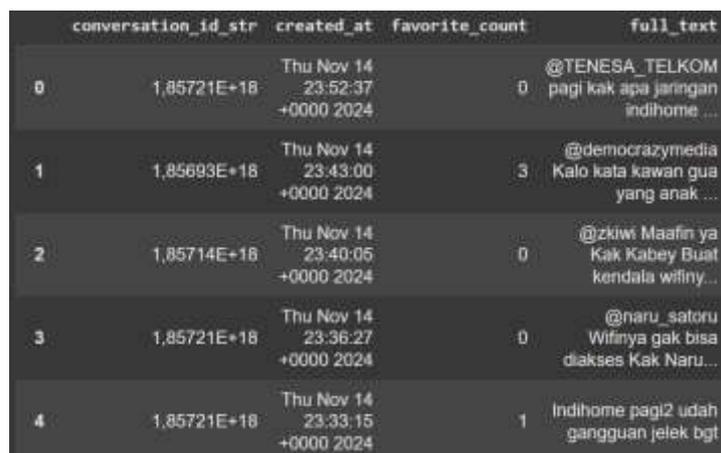
2.6. Klasifikasi *Support Vector Machine*

Klasifikasi SVM merupakan algoritma klasifikasi data linear maupun non linear, terutama pada masalah non linear dengan menggunakan konsep kernel ke dalam ruang berdimensi yang lebih tinggi (Arfat, Nurkholis and Kurniawan, 2022). Klasifikasi termasuk kedalam model supervised. Terdapat beberapa fungsi kernel $K(x_i, x_d)$ yang sering digunakan secara umum, yaitu kernel linear, sigmoid, polynomial, dan RBF. Kernel linier digunakan untuk penelitian ini. Kernel linier adalah salah satu jenis fungsi kernel yang digunakan dalam algoritma pembelajaran mesin, terutama dalam *Support Vector Machines (SVM)* dan metode berbasis kernel lainnya. Fungsi kernel adalah teknik untuk memetakan data ke ruang berdimensi lebih tinggi sehingga data menjadi lebih mudah dipisahkan atau dianalisis secara linear. Dengan menggunakan evaluasi *Confusion Matrix* pada tahap pengujian yaitu untuk menghitung dan dapat digunakan untuk menghasilkan matrik evaluasi seperti *accuracy*, *precision*, *Recall* dan *F1-score* yang dihasilkan semenjak proses pelabelan dan klasifikasi *Support Vector Machine* pada proses testing data (Farasqa Nauval Akbar *et al.*, 2023).

3. HASIL DAN PEMBAHASAN

3.1. Crawling

Crawling data atau pengumpulan data digunakan selama fase pengumpulan data dengan bantuan paket *tweepy*. Crawling data melibatkan banyak langkah dapat dimulai dengan memasukkan kata kunci. Kemudian, data cuitan didapatkan menggunakan *API Twitter* dengan bantuan paket *tweepy* dan dapat diunduh sebagai file berbentuk csv.



	conversation_id_str	created_at	favorite_count	full_text
0	1,85721E+18	Thu Nov 14 23:52:37 +0000 2024	0	@TENESA_TELKOM pagi kak apa jaringan indihome ...
1	1,85693E+18	Thu Nov 14 23:43:00 +0000 2024	3	@democrazymedia Kalo kata kawan gua yang anak ...
2	1,85714E+18	Thu Nov 14 23:40:05 +0000 2024	0	@zkiwi Maafin ya Kak Kabey Buat kendala wifiny...
3	1,85721E+18	Thu Nov 14 23:36:27 +0000 2024	0	@naru_satoru Wifinya gak bisa diakses Kak Naru...
4	1,85721E+18	Thu Nov 14 23:33:15 +0000 2024	1	Indihome pagi2 udah gangguan jelek bgt

Gambar 2. Tampilan antar muka hasil Crawling Data

Terlihat pada Gambar 2. yaitu hasil pengumpulan data cuitan. Hasil pengumpulan data itu mendapatkan data cuitan mengenai sentimen Layanan Indihome di Indonesia.

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

Dari hasil *Word Cloud* pada gambar 3. Diatas dapat dilihat kata yang berukuran besar, artinya kata yang sering muncul. Semakin besar kata yang muncul dalam teks maka semakin besar frekuensi kata tersebut muncul dalam sentiment analysis, begitupun sebaliknya. Hasilnya dalam data penelitian yang digunakan kata yang sering muncul diantaranya Indihome, Telkomsel, Indihomecare, DM, dan Kak.

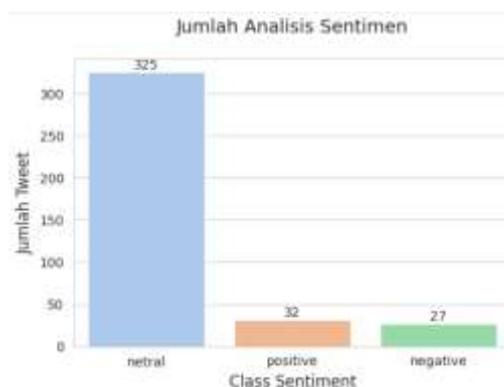
3.4. Pelabelan Sentimen

Langkah selanjutnya adalah melakukan pelabelan data menggunakan Pustaka nltk. Dataset tersebut terbagi kedalam 3 kategori yaitu opini positif, opini negative serta opini netral. Labeling sentimen menggunakan bahasa pemrograman *python* dengan membuat kamus negatif, positif dan netral dengan bahasa Indonesia dan membuat sistem. Berikut ini merupakan hasil labelling pada Gambar 4.

	stemming_data	sentiment_score	sentiment
0	sewreverse udah full kah	0.0000	netral
1	wokskyyy wokskyyy mohon maaf buat kendals jari...	0.0000	netral
2	telkomindonesia lagi kalau kita memang pake a...	0.0000	netral
3	dutagoldluck dutagoldluck malam kak mohon maaf...	0.0000	netral
4	telkomindonesia telkomindonesia kenapa engga b...	0.0000	netral
5	indihome brngssek	0.0000	netral
6	pusing wak indihome aing masih eror	0.0000	netral
7	gorenganbuku pliss sabar kakkk sini sini mm...	0.0000	netral
8	maltharu maltharu malam kak kyow mohon maaf b...	0.0000	netral
9	reminder my friend buat pake indihome biar ga ...	0.4838	positive

Gambar 4. Hasil Labelling Data

Adapun jumlah perbandingan data labeling dari banyak data yang berbeda terdapat dalam gambar 5.



Gambar 5. Perbandingan Data

Terlihat pada gambar 5. Diagram garis yang digunakan berfungsi untuk menyajikan atau menampilkan dalam perkembangan suatu informasi. Dapat dilihat bahwa kelas yang lebih unggul pada penelitian sentimen analisis Layanan Indihome dengan banyak data yang berbeda adalah kelas sentimen netral, peringkat kedua yaitu kelas positif dan kelas ketiga yaitu negatif.

3.5. Pembobotan TF-IDF

Pembobotan kata pada penelitian ini mengimplementasikan TF-IDF yaitu dengan menghitung banyak kata yang muncul dalam dokumen dan tahapan pembobotan pada kata dilakukan setelah pre-processing. TF-IDF untuk memberi nilai pada term lalu nilai term untuk input pada proses klasifikasi *Support Vector Machine (SVM)*. Berikut ini merupakan proses pembobotan TF-IDF python sebagai dengan bahasa pemrograman dibantu oleh library Scikit learn, TfidfVectorizer pada gambar 6.

```
from sklearn.feature_extraction.text import
TfidfVectorizer

tfidf_vectorizer = TfidfVectorizer()

X_train_tfidf =
tfidf_vectorizer.fit_transform(X_train)
X_test_tfidf = tfidf_vectorizer.transform(X_test)
print("hasil transformasi dari data pelatihan:")
print(X_train_tfidf)

print("\nfitur dari vektor TF-IDF:")
print(features_names)
```

Gambar 6. Script TF-IDF

Terlihat pada gambar 6. Proses pengerjaan TfidfVectorizer merupakan sebuah proses transformasi dokumen berbentuk teks menjadi vector. Adapun hasil dari pembobotan adalah sebagai berikut:

Tabel 2. Hasil dari TF-IDF

Urutan pada kata	Tf-IDF
(0, 1125)	0.275246225280136
(0, 283)	0.13852558134702897
(0, 211)	0.17546474883807095
(0, 317)	0.2261260138230409
(0, 644)	0.275246225280136

Seperti yang ada pada Tabel 2. Bahwa angka 0 pada kolom urutan kata (0, 1125) yaitu merupakan kalimat atau tweet yang pertama, lalu 1125 merupakan kata yang berada dalam urutan 1125 dalam urutan pada penggabungan. Adapun penjelasan yang berada di kolom hasil TF-IDF yaitu 0.275246225280136 merupakan nilai pembobotan terhadap kata dengan TF-IDF.

3.5. Klasifikasi *Support Vector Machine*

Klasifikasi terhadap penelitian yang telah dilakukan, data yang didapat dari tweet mengenai Layanan Indihome pada media sosial twitter. Tahap klasifikasi SVM ini sudah melalui tahap crawling data, pre-processing, Labelling sentimen, pembobotan terhadap kata menggunakan TF-IDF lalu data yang akan memasuki pada tahap klasifikasi dibagi 2 atau split data menjadi data latih serta data uji. Algoritma yang dipakai yakni *Support Vector machine* (SVM). Begitupun dengan persentase dari *accuracy* yang didapat dari nilai pelabelan. Sentimen dari setiap label itu mempunyai nilai yaitu nilai bobot yang dihitung. Kemudian dilakukanlah evaluasi *confusion matrix* multiple class, untuk memperoleh nilai terhadap *accuracy*, *precision*, *recall* dan *F1-score*. Data yang diproses pengujian klasifikasi menggunakan split data yaitu 90% pada data latih dan 10% pada data uji dengan banyak data berbeda. Adapun script untuk proses Algoritma *Support Vector Machine* (SVM) kernel linear pada pengerjaan klasifikasi dan dilakukan evaluasi *confusion matrix* terlihat pada gambar 6.

```
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
from sklearn.metrics import confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt
svm_model = SVC(kernel='linear', random_state=42)
svm_model.fit(X_train_tfidf, y_train)
print("Parameter model SVM:")
print(f"Kernel: {svm_model.kernel}")
print(f"C: {svm_model.C}")
print(f"Intercept: {svm_model.intercept_}")
print(f"Support Vectors: {svm_model.support_vectors_}")
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy:.2f}")
print(classification_report(y_test, y_pred))
conf_matrix = confusion_matrix(y_test, y_pred)
print("Confusion Matrix:")
print(conf_matrix)
```

Gambar 7. Script pada klasifikasi SVM

Terlihat pada gambar 6. Peneliti melakukan klasifikasi algoritma SVM menggunakan python yang dimana hasil dari script tersebut menghasilkan *accuracy* dari pengujian banyak data yang digunakan peneliti. Hasil dari klasifikasi ini yang sudah dilakukan dengan melalui tahapan-tahapan yang sudah dilalui. Berikut hasilnya dapat dilihat pada tabel 3.

Tabel 3. Hasil Klasifikasi SVM

Data				
<i>Accuracy</i>	Kelas pada sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
91%	Negatif	1.00	0.43	0.60
	Netral	0.90	1.00	0.95
	Positif	1.00	0.12	0.22

Dari hasil pengujian yang sudah dilakukan, yaitu bertujuan untuk mencari terhadap nilai pada *accuracy* begitupun nilai rata-rata pada *Precision*, *Recall* dan *F1-Score*, dan hasilnya ditemukan dengan variasi nilai. Nilai pada *accuracy* yang tinggi terdapat pada pengujian klasifikasi menghasilkan *accuracy* sebanyak 91%. *Precision* sebanyak 51% *Recall* sebanyak 75% dan *F1-Score* 59%, dengan label sentimen positif 15.86%, netral 70.68%, negatif 13.46%. menunjukkan bahwa pada penelitian yang dilakukan jika banyak kata yang terdeteksi dalam kelas sentimen netral dan negatif maka nilai akurasi semakin tinggi, dan jika data pada label sentimen positif semakin kecil maka atau semakin sedikit maka nilai akurasi akan semakin kecil. Untuk menghasilkan akurasi yang tinggi tergantung variasi data yang didapat. Faktor-faktor yang memengaruhi akurasi SVM melibatkan data, parameter, dan model. Pendekatan yang sistematis, seperti preprocessing data, tuning hyperparameter, dan pemilihan kernel, dapat meningkatkan akurasi secara signifikan. Jika dibandingkan dengan penelitian-penelitian sebelumnya, pada penelitian yang penulis lakukan, kelas sentimen dibagi menjadi tiga kelompok, yaitu sentimen positif, negatif, dan netral. Penelitian ini lebih berfokus pada percobaan setiap metode kernel pada algoritma SVM. Hal ini dilakukan agar dapat mengetahui metode kernel terbaik pada algoritma SVM yang dapat mengklasifikasikan sentimen dengan lebih akurat. Pada penelitian ini juga melakukan uji lanjutan untuk merepresentasikan hasil klasifikasi dalam bentuk *Word Cloud* dan juga asosiasi kata agar lebih menarik.

4. PENUTUP

4.1. Kesimpulan

Berdasarkan dari hasil yang diperoleh setelah melakukan penelitian ini, maka dapat diambil kesimpulan penerapan model dengan menggunakan algoritma *Support Vector Mahine* (SVM) terhadap dataset yang sudah didapatkan, Penggunaan *Support Vector Mahine* (SVM) dapat dijadikan opsi dalam melakukan analisis sentimen level aspek untuk penilaian terhadap Layanan *Indihome*. Penerapan metode SVM dalam analisa sentimen tentang Layanan *Indihome* berhasil mencapai tingkat *Accuracy* sebesar menghasilkan *accuracy* sebanyak 91%. *Precision* sebanyak 51% *Recall* sebanyak 75% dan *F1-Score* 59%, dengan label sentimen positif 15.86%, netral 70.68%, negatif 13.46%. Secara keseluruhan, hasil ini mengindikasikan bahwa model klasifikasi analisis sentimen menggunakan metode SVM memiliki performa yang relatif baik dalam mengenali dan mengklasifikasikan sentimen terkait Layanan *Indihome* di Indonesia.

4.2. Saran

Pada penelitian selanjutnya, diharapkan dapat melakukan pengambilan data dengan menambah waktu dalam pengumpulan data agar bisa mendapatkan hasil yang lebih bagus. Bagi penelitian selanjutnya, sebaiknya menggunakan algoritma klasifikasi lain dengan melakukan perbandingan antar klasifikasi tersebut, sehingga kita dapat mengetahui algoritma dengan kinerja terbaik. Dapat juga menggunakan *algoritma support vector machine* untuk analisis sentimen pada penelitian ini, baik berupa sentimen positif, sentiment netral maupun sentimen negatif dapat dijadikan sebagai bahan evaluasi untuk dipertahankan ataupun lebih ditingkatkan di masa yang akan datang.

DAFTAR PUSTAKA

- Alexander, N., Bria, R. and Witanti, A. (2024) 'ANALISIS SENTIMEN MASYARAKAT INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE TENTANG PILPRES 2024', 7(6).
- Alviani, V., Alam, S. and Kurniawan, I. (2023) 'Analisis Sentimen Review Aplikasi Wetv Pada Platform Twitter Menggunakan Support Vector Machine', *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, 2(3), pp. 143–149. doi: 10.55123/storage.v2i3.2351.
- Ananda, F. D. and Pristyanto, Y. (2021) 'Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Twitter Users on Internet Service Providers Using Support Vector Machine Algorithm', 20(2), pp. 407–416. doi: 10.30812/matrik.v20i2.1130.
- Anggraini, D. and Sutabri, T. (2024) 'IJM: Indonesian Journal of Multidisciplinary Pengembangan Aplikasi Penyaringan Spam e-Mail Menggunakan Teknik Machine Learning dengan Metode Support Vector Machines', 2, pp. 106–114.
- Arfat, M. F., Nurkholis, A. and Kurniawan, I. (2022) 'Analisis Sentimen Masyarakat Indonesia Terkait Vaksin Covid-19 Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine (SVM)', 7(2), pp. 96–104.
- Cantika Larasati, R., Dewi, C. and Juli, C. H. (2024) 'Analisis sentimen produk kecantikan jenis moisturizer di twitter menggunakan algoritma super vector machine', *Tekinkom*, 7(1), pp. 124–134. doi: 10.37600/tekinkom.v7i1.1243.
- Farasqa Nauval Akbar, B. *et al.* (2023) 'Analisis Sentimen Terhadap Anies Baswedan Menggunakan Metode Support Vector Machine Studi Kasus Media Sosial Twitter Sentiment Analysis of Anies Baswedan Using the Support Vector Machine Method Case Study of Twitter Social Media', *Jurnal Smart Teknologi*, 4(6), pp. 2774–1702. Available at: <http://jurnal.unmuhjember.ac.id/index.php/JST>.
- Herlambang, F. P. and Avianto, D. (2023) 'Analisis Sentimen Opini Pengguna Twitter Terhadap Tragedi Kanjuruhan Malang dengan Metode Support Vector Machine', 7, pp. 1727–1739. doi: 10.30865/mib.v7i4.6332.

Analisis Sentimen Berdasarkan pada Twitter (X) terhadap Layanan Indihome Menggunakan Algoritma Support Vector Machine (SVM)

- Ilmiah, J. and Pendidikan, W. (2024) 'Analisis Sentimen Isu Childfree Di Media Sosial Twitter Menggunakan Algoritma Support Vector Machine Lidya Nurhidayati 1 , Yuyun Umaidah 2 , Ultach Enri Universitas Singaperbangsa Karawang', 10(4), pp. 422–430.
- Saputra, R. R., & Sutabri, T. (2024). "EVALUASI TEKNOLOGI METAVERSE ROBLOX MENGGUNAKAN METODE SUPPORT VECTOR MACHINE," vol. 2, pp. 68–72, 2024.
- Petiwi, M. I., Triayudi, A. and Sholihati, I. D. (2022) 'Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine', *Jurnal Media Informatika Budidarma*, 6(1), p. 542. doi: 10.30865/mib.v6i1.3530.
- Radiena, G. *et al.* (2023) 'APLIKASI KAI ACCESS MENGGUNAKAN METODE SUPPORT VECTOR MACHINE', (April), pp. 1–10.
- Ramlan, R., Satyahadewi, N. and Andani, W. (2023) 'Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM', *Jambura Journal of Mathematics*, 5(2), pp. 431–445. doi: 10.34312/jjom.v5i2.20860.
- Robison Manalu, D. *et al.* (2022) 'METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi ANALISIS SENTIMEN TWITTER TERHADAP WACANA PENUNDAAN PEMILU DENGAN METODE SUPPORT VECTOR MACHINE', *Menthomika*, 6(2), pp. 149–156. Available at: <https://doi.org/10.46880/jmika.Vol6No2.pp149-156>.
- Sulistiani, V. A. and Hamka, M. (2024) 'Analisis Sentimen Pengguna Media Sosial Terhadap Identitas Kependudukan Digital Menggunakan Metode Support Vector Machine (SVM)', *Journal of Information System ...*, 5(4), pp. 1323–1332. doi: 10.47065/josh.v5i4.5211.
- Sumayah, S. *et al.* (2023) 'ANALYSIS OF SENTIMENT OF INDONESIAN COMMUNITY ON METAVERSE ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP METAVERSE', 4(1), pp. 143–150.
- Syahputra, D. W., Rahayudi, B. and Muflikhah, L. (2022) 'Analisis Sentimen Twitter terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat menggunakan Metode Support Vector Machine', *Fakultas Ilmu Komputer, Universitas Brawijaya*, 6(3), pp. 1067–1072.
- Wati, R. *et al.* (2021) 'Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python', 06, pp. 240–247.
- Sutabri, Tata. (2012). *Analisis Sistem Informasi*. Yogyakarta: Andi Offset.
- Sutabri, Tata. (2004). *Pemrograman Terstruktur*. Yogyakarta: Andi Offset.
- Sutabri, Tata dan Darmawan Napitupulu. (2019). *Sistem Informasi Bisnis*. Yogyakarta: Andi Offset.