

SKRIPSI

**IMPLEMENTASI ALGORITMA *SUPPORT VECTOR MACHINE* TERHADAP ANALISIS SENTIMEN
PENGUNAAN APLIKASI TIKTOK SHOP SELLER CENTER**

***IMPLEMENTATION OF SUPPORT VECTOR MACHINE
ALGORITHM FOR SENTIMENT ANALYSIS OF TIKTOK SHOP
SELLER CENTER APPLICATION USAGE***



SARINA

D0219413

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS SULAWESI BARAT

MAJENE

2023

ABSTRACT

Tiktok Shop Seller Center is one of the leading e-commerce in Indonesia. This app is designed to help sellers manage and increase sales. Companies or organizations need to understand the response received from the public to the products or services they offer. Opinions that arise from the public can have an impact on the image of the company or organization. However, monitoring and managing public opinion is not a simple task. The volume of incoming opinions is too large to be processed manually. Therefore, special approaches or techniques are needed that can automatically categorize those reviews as positive or negative. One platform that provides reviews is Google Play. Data obtained from Google Play is then labeled and analyzed using the Support Vector Machine method to classify these reviews. From the results of the labeling that has been done, visualizations will be seen in each sentiment class to find information that is considered important and can be useful for decision making. Classification with the SVM method obtained an accuracy rate of 91.59%. Furthermore, visualization methods in the positive sentiment class include good, helpful, steady, easy, good, bangat, please, application and sell. Meanwhile, the class of negative sentiment that is often complained about includes applications, uploads, violations, blocks, products, difficult, registers, complicated, fails and accounts.

Keywords : *Sentiment Analysis, Support Vector Machine (SVM), E-Commerce, Tiktok Shop Seller Center, Google*

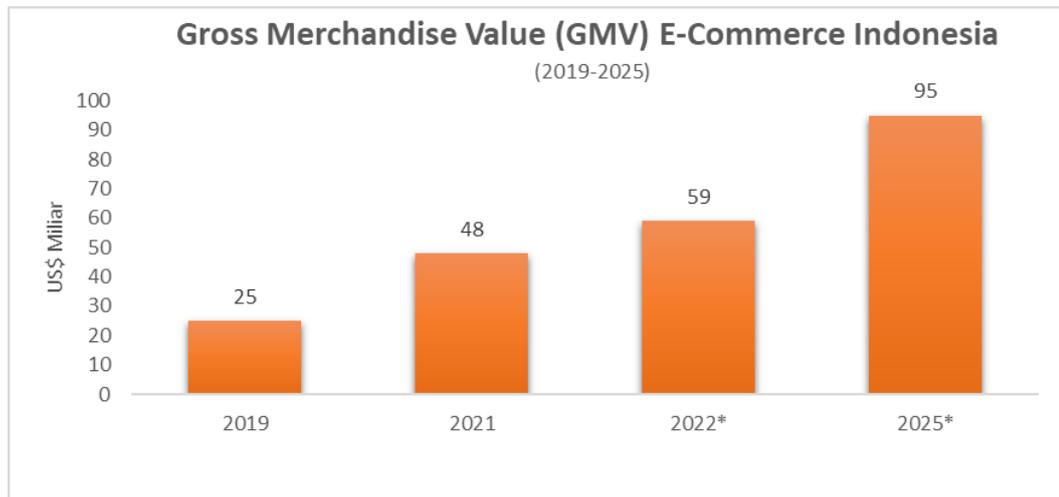
BAB I

PENDAHULUAN

A. Latar Belakang

Indonesia menjadi negara dengan tingkat pengguna *e-commerce* tertinggi di Asia Tenggara. Sejak 2019, sebanyak 90% pengguna internet Indonesia telah melakukan pembelian melalui platform *e-commerce*. Pada 2020, DBS melakukan *survey* dengan melibatkan 500 responden yang mengungkapkan bahwa setelah pandemi Covid-19 jumlah pelanggan *e-commerce* di Indonesia naik menjadi 66% dan belanja *online* meningkat 14 % sementara belanja konvensional turun signifikan sebesar 24% sejak pandemi. Seiring berkembangnya internet, minat belanja *online* menjadi tinggi yang menguntungkan perusahaan besar yang menjual produknya melalui platform *e-commerce*, selain itu juga membantu pertumbuhan usaha mikro, kecil dan menengah (DBS research, 2020).

E-commerce atau bisa disebut perdagangan elektronik atau e-dagang adalah penyebaran, pembelian, penjualan, pemasaran barang dan jasa melalui internet atau jaringan komputer. Seluruh komponen yang ada dalam perdagangan diaplikasikan ke dalam *E-commerce* seperti *customer service*, layanan produk, cara pembayaran dan cara promosi. Perkembangan internet yang semakin maju merupakan salah satu jaringan global yang menghubungkan jaringan komputer diseluruh dunia, sehingga memungkinkan terjalinnya komunikasi dan interaksi antara satu perusahaan dengan konsumen (Hermiati et al., 2021). Berikut ini adalah pertumbuhan *e-commerce* di Indonesia berdasarkan *Gross Merchandise Value* (GMV) dari tahun 2019-2025.



Gambar 1.1 GVM E-commerce Indonesia

(sumber : Temasek *Bain & Company*, 2020)

Kegiatan belanja *online* melalui platform terus bertumbuh di Indonesia. Pada era digital ini, sebagian besar masyarakat pernah melakukan transaksi belanja *online*. Transaksi *online* tersebut dilakukan di berbagai *marketplace* seperti *Tiktok Shop*, *Shopee*, *Tokopedia*, *Bukalapak*, *Lazada* dan berbagai macam *marketplace* lainnya. Salah satu *marketplace* yang populer di Indonesia saat ini adalah Aplikasi *Tiktok Shop*. Dimana, pada tahun 2021 jumlah pengguna *tiktok* semakin meningkat. Pasalnya, aplikasi yang tadinya hanya sebatas media sosial tersebut, kini telah menjadi tempat belanja favorit di tanah air (Fide et al., 2021).

Tiktok merupakan salah satu media sosial yang populer di dunia. *E-commerce tiktok* yang bernama *tiktok shop*, tengah berkembang pesat terutama di wilayah Asia Tenggara. Berdasarkan data *We Are Social*, Indonesia menempati peringkat kedua di dunia sebagai pengguna aplikasi *tiktok* terbanyak sekitar 109,9 juta pengguna. Pada Januari 2023 pengguna *Tiktok* di dunia diperkirakan mencapai

1,05 miliar. Jumlah tersebut meningkat sebesar 18,8% dibandingkan tahun sebelumnya. Tahun 2021 Tiktok sudah meluncurkan fitur belanja online Tiktok Shop di Indonesia. Dengan hadirnya Tiktok Shop, pengguna bisa langsung berbelanja lewat aplikasi Tiktok, tanpa harus beralih aplikasi untuk menyelesaikan transaksi pembelian produk. Artinya semua transaksi pembelian hingga pembayaran dilakukan di aplikasi Tiktok itu sendiri. Tiktok juga telah meluncurkan aplikasi Tiktok Shop Seller Center untuk mempermudah penjual dalam manajemen toko.

Tiktok Shop Seller Center merupakan aplikasi yang dirancang untuk membantu penjual mengelola dan meningkatkan penjualan. Tiktok Seller Center adalah portal yang cukup lengkap bagi merchant untuk menjalankan bisnis *e-commerce* mereka di Tiktok. Aplikasi ini menyediakan solusi lengkap yang mencakup manajemen pengoperasian toko, berbagai fitur pembayaran, iklan, laporan penjualan, layanan dan engagement pelanggan, fitur pemantauan, perkembangan penjualan, analisis kinerja, pembaruan kebijakan dan lain sebagainya. Dengan menggunakan aplikasi ini penjual dapat meningkatkan penjualan dengan mengoptimalkan strategi pemasaran. Tiktok Shop Seller Center pada *Google Play Store* telah diunduh lebih dari 10 juta kali dan memiliki 148 ribu ulasan.

Ulasan-ulasan tersebut dapat memberikan pengaruh pada aplikasi Tiktok Shop Seller Center dalam melakukan perbaikan terhadap aplikasi. Namun ulasan – ulasan tersebut perlu dilakukan analisis untuk mengetahui pola dari tanggapan yang diberikan oleh pengguna. Metode yang dapat digunakan dalam penelitian ini adalah

Text Mining. Dimana *Text Mining* sendiri merupakan suatu analisis teks dengan tujuan menggali informasi yang berkualitas dari suatu rangkaian teks.

Pada *Google Play Store* pemberian rating aplikasi diikuti dengan ulasan dari para pengguna terhadap aplikasi, ulasan yang cukup banyak dan tidak terstruktur membuat calon pengguna sulit untuk membacanya, karena ulasan tersebut mengandung opini tentang kepuasan pengguna terhadap aplikasi sehingga sering di jadikan pertimbangan oleh calon pengguna yang memilih aplikasi yang akan digunakan. Berdasarkan hal tersebut, diperlukan suatu teknik untuk mengetahui bagaimana ulasan pengguna terhadap suatu aplikasi. Oleh karena itu, penelitian ini melakukan analisis sentimen yang merupakan salah satu teknik analisis *text mining* yang dapat diaplikasikan pada perusahaan yang menciptakan suatu produk dan jasa serta menyediakan layanan terhadap konsumen untuk memberikan pendapat (*feedback*) pada suatu produk. Kegunaan analisis sentimen yaitu untuk mengelompokkan *feedback* positif, negatif dan netral dari konsumen agar mempermudah dan mempercepat tugas perusahaan untuk meninjau kembali kekurangan suatu produk, dan bila ditemukan adanya sentimen negatif, maka perusahaan dapat dengan cepat untuk menanggulangnya.

Penelitian sebelumnya dilakukan oleh Jessica pada tahun 2021, dimana dilakukan perbandingan *Naïve Bayes*, SVM dan KNN untuk analisis sentimen pada Gadget berbasis Aspek, dan kesimpulannya adalah model klasifikasi SVM menunjukkan tingkat akurasi terbaik. Rata-rata *accuracy* SVM sebesar 96.43% dari empat aspek, yaitu aspek desain sebesar 94.40%, aspek harga sebesar 97.44%, aspek spesifikasi sebesar 96.22%, dan aspek citra merk sebesar 97.63% dalam

penelitian ini (Iskandar & Nataliani, 2021). Pada penelitian analisis sentimen ulasan aplikasi Primaku dengan membandingkan 3 kernel algoritma SVM diperoleh hasil bahwa kernel linear merupakan kernel terbaik dengan akurasi hingga 97.5% (Tinaliah & Elizabeth, 2022). Sedangkan dalam penelitian lainnya akurasi terbaik diperoleh oleh kernel RBF (*Radial Basis Function*) dalam analisis sentimen pada aplikasi Mola di *Google Play Store* dengan algoritma SVM sebesar 92.31% (Hendriyanto et al., 2022).

Berdasarkan penjelasan di atas, karena belum ada hasil pasti mengenai akurasi kernel terbaik dalam melakukan analisis sentimen, maka pada penelitian ini akan dilakukan pengujian antara kernel *Linear*, RBF, *polynomial* dan *sigmoid*. pada algoritma *Support Vector Machine* (SVM) dengan menggunakan metode *Knowledge Discovery Databases* (KDD), pembobotan menggunakan *Term Frequency – Inverse Document Frequency* (TD-IDF) serta evaluasi menggunakan *confusion matrix* dan melakukan analisis sentimen ulasan aplikasi Tiktok Shop Seller Center untuk mengetahui bagaimana sentimen para pengguna terhadap aplikasi Tiktok Shop Seller Center di *Google Play Store* dan akurasi kernel terbaik.

B. Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka dalam penelitian ini permasalahan yang akan dirumuskan diantaranya adalah:

1. Bagaimana gambaran umum terhadap tanggapan dari pengguna aplikasi Tiktok Shop Seller Center?

2. Bagaimana hasil penerapan metode *Support Vector Machine* (SVM) dalam melakukan klasifikasi sentimen data ulasan pengguna aplikasi Tiktok Shop Seller Center?

C. Tujuan Penelitian

Penelitian ini bertujuan untuk :

1. Mengetahui gambaran umum data ulasan tentang Tiktok Shop Seller Center berdasarkan situs *Google Play*.
2. Mengetahui hasil penerapan metode *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen data ulasan dari pengguna aplikasi Tiktok Shop Seller Center.

D. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebaga berikut:

1. Manfaat bagi masyarakat :
Analisis sentimen terhadap ulasan aplikasi Tiktok Shop Seller Center dengan algoritma SVM dapat memberikan informasi mengenai klasifikasi teks serta dapat membantu dalam pemilihan aplikasi yang akan digunakan dan pengembangan aplikasi bagi perusahaan.
2. Manfaat bagi akademik:
Hasil penelitian diharapkan dapat membantu dalam hal referensi atau memberikan informasi yang berguna untuk dunia akademis khususnya dalam

penelitian yang akan dilakukan oleh para peneliti yang akan datang, khususnya dalam bidang *text mining* tentang analisis sentimen.

3. Manfaat bagi penulis:

Untuk menambah ilmu dan pengetahuan tentang kinerja dari algoritma *Support Vector Machine* terhadap analisis sentimen dan sebagai syarat dalam memperoleh gelar sarjana (s1).

E. Batasan Masalah

Batasan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Data yang akan diklasifikasi hanya data *review* pengguna yang berbahasa Indonesia Tiktok Shop Seller Center pada *Google Play*.
2. Analisis yang digunakan untuk analisis sentimen adalah metode SVM.
3. Data ulasan yang digunakan dalam penelitian yaitu data ulasan mulai 1 Mei 2023 – 31 Juli 2023

BAB II

TINJAUAN PUSTAKA

A. *E-commerce*

E-commerce adalah kegiatan transaksi jual beli pada perusahaan atau individu yang dilakukan secara elektronik melalui internet. *E-Commerce* sendiri adalah singkatan dari *Electronic Commerce*, yang dalam Undang-undang (UU) Nomor 7 Tahun 2014 dalam pasal 65 dan pasal 66 tentang Perdagangan telah mengatur kegiatan *e-commerce*. Sesuai pasal 66 UU tersebut, ketentuan lebih lanjut mengenai transaksi pada *e-commerce* diatur dengan atau berdasarkan Peraturan Pemerintah. *E-commerce* secara singkat dapat didefinisikan sebagai mekanisme transaksi jual dan beli yang menggunakan fasilitas internet sebagai media dalam berkomunikasi. Menurut penelitian (Fernandes, 2019), *e-commerce* adalah transaksi penjualan secara online yang hanya dilakukan oleh toko itu sendiri. Dalam pengertian lain, *E-commerce* dapat diartikan sebagai rangkaian usaha perdagangan yang memungkinkan proses membeli dan menjual secara digital antar organisasi atau antar individu.

Menurut penelitian (Hermiati et al., 2021), *e-commerce* merupakan perdagangan *online* yang memungkinkan perusahaan atau individu melakukan distribusi serta transaksi yang dilakukan melalui internet. Dalam *e-commerce* terjadi proses pertukaran produk-produk secara elektronik atau penjualan dan pembelian antar dua belah pihak melalui media internet. Keunggulan *e-commerce* adalah dapat diakses dimanapun, kemampuan transaksi kapanpun, adanya

penghematan biaya dan jalur distribusi. Pada *e-commerce*, pembayaran dapat dilakukan dengan banyak cara seperti pembayaran *online banking* atau transaksi antar bank, kartu kredit serta pembayaran *via online* lainnya (Dana, bitcoin, e-money).

B. Analisis Sentimen

Analisis sentimen atau bisa juga disebut *opinion mining* merupakan cabang penelitian *text mining*. Analisis sentimen merupakan bidang yang luas yang melakukan proses pengklasifikasian dokumen teks dengan melakukan ekstraksi pendapat, emosi dan evaluasi tertulis seseorang tentang topik tertentu. Analisis sentimen dapat juga digunakan dalam berbagai kemungkinan domain, dari produk konsumen, jasa keuangan, jasa kesehatan, peristiwa sosial dan politik pada pemilu. Kecenderungan penelitian analisis sentimen berfokus pada pendapat yang menyatakan suatu sentimen positif atau negatif (Praptiwi, 2018).

C. Text Mining

1. Pengertian *Text Mining*

Text mining atau *text analytics* merupakan sebuah proses dalam menganalisis text yang tidak terstruktur yang terkait satu sama lain. Pada penelitian (Audrin & Audrin, 2022), menyatakan *text mining* telah digunakan dalam bidang pendidikan berbagai data dan mampu melakukan analisis pada volume data yang besar seperti kumpulan literatur tentang keterampilan digital dalam pendidikan. *Text mining*

adalah suatu proses mengekstrak informasi melalui identifikasi serta eksplorasi pola dari kumpulan data tekstual yang tidak terstruktur (Fide et al., 2021).

2. *Text Preprocessing*

Text preprocessing adalah proses yang dilakukan dengan menyeleksi data text menjadi terstruktur dengan melalui beberapa tahap yaitu *cleansing*, *case folding*, *tokenizing*, *stopword removal* dan *stemming*. Namun sesungguhnya tidak ada tahapan pasti dalam melakukan *text preprocessing*.

Cleansing merupakan tahap awal yang melakukan penghapusan simbol, karakter spesial, tanda baca hingga angka dan *emoticon*. Pada tahap *case folding* mengubah kata dengan mengonversi semua huruf menjadi huruf kecil untuk menghindari kata yang sama namun tidak terdeteksi karena adanya perbedaan bentuk huruf. Selanjutnya dilakukan tahap *tokenizing* yaitu proses pemotongan kalimat menjadi kata-kata dalam data ulasan. Setelah dilakukan pemotongan kalimat akan dilakukan penghapusan kata yang tidak memiliki pengaruh pada kalimat itu sendiri atau *stopword removal*. Proses terakhir yaitu *stemming* mengubah kata berimbuhan menjadi kata dasar.

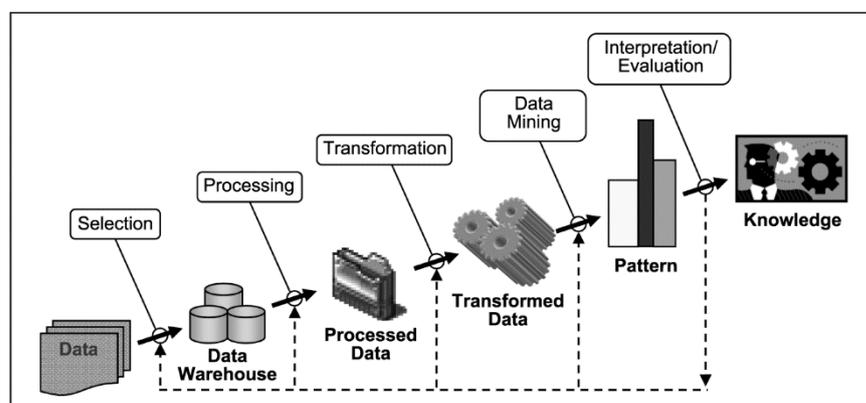
D. *Lexicon*

Metode *lexicon based* dilakukan dengan cara membuat kamus *lexicon*, dimana kamus tersebut akan digunakan untuk mengidentifikasi apakah kata tersebut mengandung opini atau tidak (Hasna, 2021). *Lexicon based* merupakan proses pelabelan otomatis apabila komentar Tiktok Seller terdapat opini akan

dilabeli dengan polaritas tinggi (positif), polaritas rendah (negatif) atau polaritas netral. Cara kerja dari *lexicon based* ini diperlukan kamus kata yang menentukan apakah kata tersebut bernilai positif, negatif atau netral.

E. Knowledge Discovery In Database (KDD)

Knowledge Discovery Databases (KDD) adalah proses penerapan metode saintifik pada data mining yang bertujuan untuk menggali serta menganalisis sejumlah besar himpunan data dan mengekstrak informasi pengetahuan. Metode KDD dapat juga diartikan sebagai proses non-trivial untuk mencari dan mengidentifikasi pola dalam data. Dimana pola bersifat baru, sah, bermanfaat dan dapat dipahami. Berikut ini penjelasan metode KDD (Arta et al., 2019)



Gambar 2.1 Tahapan KDD

(Sumber : Michael, 2019)

1. Data selection

Menciptakan dan memilih himpunan data atau memfokuskan pada subset variabel atau sample data.

2. *Pre-processing/cleaning*

Proses pembersihan dan memperbaiki kesalahan pada data

3. *Transformation*

Pencarian fitur yang berguna untuk mempresentasikan data sesuai untuk proses data mining.

4. *Data mining*

Proses pemilihan tugas data mining, seperti pemilihan algoritma untuk klasifikasi sebagai teknik atau metode

5. *Evaluation*

Proses evaluasi data dengan metode atau teknik untuk menentukan hasil akhir dari proses KDD

F. Tiktok Shop Seller Center

Tiktok Shop Seller Center merupakan aplikasi yang dirancang untuk membantu penjual mengelola dan meningkatkan penjualan. Tiktok Seller Center adalah portal yang cukup lengkap bagi merchant untuk menjalankan sebuah bisnis *e-commerce* mereka di Tiktok. Aplikasi ini menyediakan solusi lengkap yang mencakup manajemen pengoperasian toko, berbagai fitur pembayaran, iklan, laporan penjualan, layanan dan engagement pelanggan, fitur pemantauan, perkembangan penjualan, analisis kinerja, pembaruan kebijakan dan lain sebagainya. Dengan menggunakan aplikasi ini penjual dapat meningkatkan penjualan mereka dengan menganalisis sentimen pengguna dengan mengoptimalkan strategi pemasaran.

G. Google Play

Google play adalah layanan konten digital milik *google* yang terdiri dari toko-toko *online* yang dikunjungi pengguna untuk mendapatkan aplikasi, film, acara TV, edukasi, game, buku dan konten lainnya. *Google Play* dapat diakses melalui aplikasi android (*Play Store*), *web* dan *Google TV*. Pada *Google Play* dilengkapi fitur yang berisi ulasan terhadap suatu aplikasi dari para pengguna yang dapat digunakan untuk melihat ranting aplikasi.

H. Support Vector Machine

Vapnik pada tahun 1992, memperkenalkan *Support Vector Machine* (SVM) sebagai suatu metode klasifikasi yang efisien untuk masalah linear maupun nonlinear. *Support Vector Machine* (SVM) juga dikenal sebagai salah satu algoritma *machine learning* berbasis *supervised learning* yang mampu menganalisis data untuk masalah klasifikasi dan regresi. Cara kerja SVM adalah dengan menemukan *hyperplane* terbaik yang berfungsi untuk memisahkan antara kelas opini positif dan opini negatif pada input space. SVM berusaha mencari *Hyperplane* atau fungsi pemisah dengan memaksimalkan jarak antar kelas. Penelitian (Siringoringo & Jamaluddin, 2019), menyatakan bahwa SVM termasuk metode terbaik di antara metode lainnya, karena mampu melakukan pengkomputasian data dengan dimensi tinggi yang menghasilkan tingkat akurasi yang lebih baik. SVM mempunyai sebuah fungsi yang mampu mentransformasikan data ke dimensi yang tinggi yaitu ruang kernel yang sering disebut fungsi kernel

trick. Ada beberapa fungsi kernel pada SVM yaitu kernel *Linear*, *Polynomial*, *Radial Basis Function* (RBF), dan *Sigmoid* (Zuriel & Fahrurozi, 2021).

1. Kernel *Linear*

Fungsi kernel *linear* adalah fungsi kernel yang paling sederhana, digunakan untuk klasifikasi data linear yang digunakan ketika data yang dianalisis telah terpisah secara linear. Kernel ini tepat digunakan ketika terdapat banyak fitur, karena pemetaan ke ruang dimensi yang lebih tinggi tidak benar-benar meningkatkan kinerja. Berikut persamaan fungsi kernel linear:

$$K(\vec{x}_i, \vec{x}_j) = \text{sum}(x \times x_i) \quad (2.1)$$

2. Kernel *Polynomial*

Fungsi Kernel *polynomial* adalah bentuk umum dari kernel *linear*, yang digunakan ketika data tidak terpisah secara linear. Kernel *Polynomial* juga berfungsi untuk memecahkan suatu masalah klasifikasi pada dataset pelatihan yang dinormalisasi. Dalam *machine learning*, kernel *polynomial* adalah fungsi kernel yang tepat digunakan dalam SVM dan kernelisasi lainnya, karena kernel mewakili kesamaan vektor sampel pelatihan dalam ruang fitur. Parameter derajat (d) pada kernel *polynomial* berfungsi untuk mencari nilai optimal pada setiap dataset. Dimana nilai default $d = 2$, semakin besar nilai d yang dihasilkan akan fluktuatif dan kurang stabil. Dikarenakan semakin tinggi nilai d maka

semakin melengkung garis hyperplane yang dihasilkan. Berikut persamaan kernel *polynomial*:

$$K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i, \vec{x}_j + 1)^d \quad (2.2)$$

3. Kernel *Radial Basic Function* (RBF)

Kernel *Radial Basic Function* atau juga disebut kernel *Gaussian* yang digunakan untuk klasifikasi data non-linear. Konsep kernel yang paling banyak digunakan dalam memecahkan masalah klasifikasi data yang tidak dipisahkan secara linear. Fungsi ini dikenal memiliki performa yang baik, dan hasil dari pelatihan memiliki nilai error yang kecil di banding kernel lainnya. Rumus persamaan kernel RBF :

$$K(\vec{x}_i, \vec{x}_j) = \exp(-\gamma \|\vec{x}_i - \vec{x}_j\|^2) \quad (2.3)$$

4. Kernel *Sigmoid*

Fungsi kernel *Sigmoid* merupakan kernel trick SVM yang merupakan pengembangan dari jaringan saraf tiruan.

$$K(\vec{x}_i, \vec{x}_j) = \tanh(\alpha \vec{x}_i + \vec{x}_j + \beta) \quad (2.4)$$

I. *Term Frequency – Inverse Document Frequency (TF-IDF)*

Term Frequency – Inverse Document Frequency (TF-IDF) Merupakan suatu proses pembobotan term secara statistic. TF-IDF adalah metode pembobotan

term yang banyak digunakan sebagai metode pembandingan terhadap metode pembobotan yang baru. Konsep TF-IDF adalah melakukan perhitungan bobot term t dalam sebuah dokumen yang dilakukan dengan mengalikan nilai *Term Frequency* dengan *Inverse Document Frequency*.

Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk *transformation* dengan melakukan perubahan pada kalimat atau kata menjadi angka sehingga dapat diklasifikasi dan menggunakan *Cosine Similarity* agar bisa melihat kemiripan antar kata atau dokumen (Haranto & Sari, 2019).

Perhitungan metode TF-IDF dinyatakan dalam persamaan berikut:

$$w_t = TF_{t,d} \times IDF_t = TF_{t,d} \times \log \frac{N}{DF_t} \quad (2.5)$$

Keterangan :

w_t : Bobot TF-IDF

$TF_{t,d}$: Jumlah frekuensi kata

IDF_t : Jumlah inverse frekuensi dokumen tiap kata

DF_t : Jumlah frekuensi dokumen tiap kata

N : Jumlah total dokumen

J. Confusion Matrix

Confusion Matrix merupakan salah satu metode *Data Mining* yang dapat digunakan untuk mengukur kinerja suatu masalah klasifikasi *machine learning*. Selain itu, *Confusion Matrix* memiliki label yang dapat dijadikan sebagai acuan

untuk pengukuran performa pada sebuah model dan dapat mewakili jumlah dari nilai prediksi dan nilai aktual. Cara kerja metode *Confusion Matrix* adalah matriks dari prediksi dibandingkan dengan kelas asli yang berisi informasi nyata dan nilai klasifikasi (Styawati et al., 2021). Kondisi untuk melakukan pengujian klasifikasi menggunakan metode Matrix dapat dilihat pada **Tabel 2.1**.

Terdapat beberapa rumus umum yang dapat digunakan untuk menghitung performa klasifikasi. Hasil dari nilai akurasi, presisi, dan recall biasa ditampilkan dalam persentase.

Tabel 2.1 *Confusion Matrix*

Class	Positif	Negatif
Positif	True Positif (TP)	False Negatif (FN)
Negatif	False Positif (FP)	True Negatif (TN)

1. *Accuracy*

Akurasi adalah jumlah proporsi prediksi yang benar. Rumus dapat dilihat pada persamaan (Manik et al., 2021).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.6)$$

2. *Precision*

Precision mengukur sejauh mana prediksi dari model yang benar. Presisi berfokus pada jumlah prediksi positif yang benar dibandingkan dengan total prediksi yang dilakukan oleh model (Manik et al., 2021).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2.7)$$

3. *Recall*

Recall mengukur sejauh mana model mampu mengidentifikasi semua instance yang seharusnya positif. *Recall* berfokus pada jumlah prediksi positif yang benar dibandingkan dengan total instance yang seharusnya positif (Manik et al., 2021).

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2.8)$$

4. *F1-Score*

F1-Score adalah ukuran yang menggabungkan presisi dan recall menjadi satu nilai tunggal. Ini memberikan gambaran yang lebih seimbang tentang kinerja model daripada hanya menggunakan presisi atau *recall* secara terpisah (Manik et al., 2021).

$$F1\ Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (2.9)$$

K. Kuantitatif

Metode kuantitatif merupakan salah satu jenis penelitian yang spesifikasinya adalah sistematis, terencana, dan terstruktur dengan baik dari awal hingga proses pembuatan desain penelitian. Selain itu, metode kuantitatif dapat juga diartikan sebagai penelitian yang banyak menggunakan angka, mulai dari pengumpulan data, penafsiran data hingga hasilnya (Ortega & Brotosaputro, 2019).

L. Python

Python dikenal sebagai bahasa pemrograman interpreter, karena *Python* dieksekusi dengan sebuah interpreter (Hasna, 2021). *Python* dikatakan *reliability* karena *Python* memiliki *source code* yang sederhana sehingga mudah ditulis, mudah diingat dan mudah digunakan secara berulang. Hal tersebut memudahkan pengembang aplikasi dalam melakukan coding, *testing*, perbaikan apabila terjadi kesalahan, *bug* atau *error* lainnya. *Python* dikatakan efisien, dikarenakan *python* memiliki *library* yang lengkap dan *python* code akan lebih sederhana apabila dibandingkan kode yang ditulis pada pemrograman lainnya.

M. Penelitian Terkait

Berikut ini merupakan penelitian terkait dalam penelitian ini :

Tabel 2.2 Penelitian Terkait

No	Nama dan Tahun	Judul Penelitian	Hasil Penelitian	Perbedaan Penelitian	Persamaan Penelitian
1	Geyessella Manik, Iin Ernawati, Ika Nurlaili 2021	Analisis Sentimen pada review pengguna e-commerce Bidang Pangan Menggunakan Metode Support Vector Machine	Metode SVM pada klasifikasi sentimen berjalan dengan baik. Dengan penerapan pembagian data train dan data test dengan perbandingan 80 : 20 didapatkan hasil akurasi klasifikasi pada data sayurbox dengan metode SVM dengan kernel linear yaitu sebesar 91,4%, presisi 93,1%, dan recall yaitu sebesar 94,4%. Hasil akurasi pada data sayurbox, disimpulkan bahwa dari 210 review yang diuji didapatkan 192 review yang sesuai dengan	Sumber dataset yang digunakan, perbandingan skenario yang berbeda. Dan perbandingan kernel.	e-commerce dengan analisis sentimen menggunakan SVM

No	Nama dan Tahun	Judul Penelitian	Hasil Penelitian	Perbedaan Penelitian	Persamaan Penelitian
			pengklasifikasi annya.		
2	Heart Parasian, Achmad Fahrurozi 2021	Implementasi Algoritma Klasifikasi Support Vector Machine untuk Analisa Sentimen Pengguna Twitter terhadap Kebijakan PSBB	Model klasifikasi SVM dengan kernel RBF mendapatkan hasil keseluruhan yang paling baik. Nilai <i>accuracy</i> sebesar 95.94%, <i>precision</i> sebesar 94.41%, <i>recall</i> sebesar 97.8%, dan <i>F1-score</i> sebesar 96.08%	Dataset yang digunakan dan sumber data pada penelitian ini dari aplikasi Twitter	Algoritma Support Vector Machine dengan analisis sentimen
3	Sola Fide, Suparti, Sudarno 2021	Analisis Sentimen Ulasan Aplikasi Tiktok di Google Play menggunakan Metode Support Vector Machine dan Asosiasi	Klasifikasi sentimen dari hasil <i>sentiment scoring</i> ulasan aplikasi <i>Tiktok</i> di <i>Google Play</i> menggunakan metode <i>Support Vector Machine</i> (SVM) <i>kernel</i> RBF dengan perbandingan data latih dan uji sebesar 80:20 menghasilkan tingkat <i>accuracy</i> dan <i>kappa</i> terbaik sebesar 90,62% dan 81,24% yang termasuk	Dataset yang digunakan, kernel serta perbandingan skenario dan metode penelitian	Analisis sentimen dengan algoritma SVM dengan menggunakan sumber dataset dari Google Play Store

No	Nama dan Tahun	Judul Penelitian	Hasil Penelitian	Perbedaan Penelitian	Persamaan Penelitian
			hasil klasifikasi sempurna.		
4	Styawati, Nirwana, Auliya, Ari 2021	Analisis Sentimen Masyarakat Terhadap Program Prakerja pada Twitter dengan Metode Support Vector Machine	Hasil dari klasifikasi menggunakan metode Support Vector Machine yaitu netral sebanyak 98,34%, negatif sebanyak 0,99%, positif sebanyak 0,66%. dapat disimpulkan bahwa sentiment masyarakat dari pengguna twitter terhadap program kartu prakerja dimasa pandemi lebih condong ke netral sebesar 98,34%	Metode pengumpulan data, Sumber data pada penelitian ini yaitu dari twitter	Analisis sentimen, pembobotan, <i>Confusion Matrix</i> , SVM
5	Primandani Arsi, Retno Waluyo 2019	Analisis Sentimen Wacana Pemandangan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)	Berdasarkan hasil pengujian yang dilakukan terhadap <i>tweets</i> sentimen pemindahan ibu kota dari media sosial twitter sebanyak 1.236 <i>tweets</i> (404 positif dan 832 negatif) menggunakan SVM diperoleh	Sumber data serta dataset, tidak melakukan perbandingan skenario dan kernel SVM	Analisis sentimen dengan menggunakan algoritma SVM

No	Nama dan Tahun	Judul Penelitian	Hasil Penelitian	Perbedaan Penelitian	Persamaan Penelitian
			akurasi=96,68 %, <i>precision</i> =95,82%, <i>recall</i> =94,04% dan AUC = 0,979		
6	Rimbun Siringoringo, Jamaluddin 2019	Text mining dan Klasifikasi Sentimen pada Ulasan Produk Toko Online	Dengan menerapkan pengujian pada tiga jenis data sentimen yang berbeda di peroleh bahwa Support Vector Machine dapat bekerja dengan baik pada data ulasan tidak seimbang. Hasil klasifikasi sentimen menggunakan SVM lebih unggul dari K-Nearest Neighbor	Sumber Dataset yang digunakan serta pada penelitian ini melakukan perbandingan klasifikasi algoritma	E-commerce dengan analisis sentimen menggunakan SVM
7	Muhammad Diki Hendriyanto, Azhari Ali Ridha, Ultach Enri 2022	Analisis Sentimen Ulasan Aplikasi Mola pada Google Play Store Menggunakan Algoritma Support Vector Machine	Algoritma SVM dalam menganalisis sentimen ulasan aplikasi MOLA dengan 3 skenario di peroleh hasil terbaik 90:10 dengan akurasi 92,31%	Dasaset yang digunakan, perbandingan skenario dan pada penelitian ini tidak melakukan pelabelan data.	Menggunakan algoritma SVM dengan 4 perbandingan kernel
8	Tinaliah, Triana Elizabeth 2022	Analisis Sentimen Ulasan Aplikasi	SVM dapat melakukan analisis sentimen	Dataset yang digunakan, labeling	Klasifikasi analisis sentimen dengan

No	Nama dan Tahun	Judul Penelitian	Hasil Penelitian	Perbedaan Penelitian	Persamaan Penelitian
		Primaku Menggunakan Metode Support Vector Machine	ulasan pengguna aplikasi PrimaKu dengan baik menggunakan linear kernel dengan nilai akurasi, yaitu 97.5%	dengan menggunakan skor ulasan, perbandingan kernel.	menggunakan algoritma SVM
9	Rizki Wahyudi, Gilang Kusumawardhana 2021	Analisis Sentimen pada Review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine	Hasil dari analisis menggunakan Support Vectore Machine menghasilkan akurasi 85,54%	Data yang digunakan, penelitian ini tidak melakukan labeling, dan perbandingan rasio	Analisis sentimen dengan menggunakan algoritma Support Vector Machine
10	Alfio Kusuma, Ermatita, Helena Nurramdhani Irmanda 2022	Analisis Sentimen pada Ulasan Aplikasi Indodax di Google Play Store Menggunakan Metode Support Vector Machine	Klasifikasi menggunakan metode Support Vector Machine menghasilkan akurasi sebesar 85% pada rasio pembagian data sebesar 80:20	Pada penelitian ini tidak melakukan labeling dan menggunakan dataset yang berbeda	Sumber dataset yang sama dan klasifikasi dengan menggunakan algoritma SVM

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan hasil analisis dan pembahasan pada bab sebelumnya, penulis dapat menarik beberapa kesimpulan sebagai berikut:

1. Berdasarkan hasil pelabelan dari 3.207 ulasan, yang termasuk sentimen positif yaitu sebanyak 2.244 dan ulasan yang termasuk kedalam sentimen negatif sebanyak 963. Dapat dilihat bahwa jumlah sentimen positif lebih banyak dibanding dengan sentimen negatif.
2. Berdasarkan proses klasifikasi dengan menggunakan metode SVM yang dilakukan bekerja dengan baik dan menunjukkan nilai akurasi tertinggi dimiliki oleh kernel *linear* dengan perbandingan skenario data latih 80% dan data uji 20% sebesar 94,39% dengan parameter $C = 20$ dan $\gamma = 20$. Sedangkan hasil *confusion matrix* dari 642 data yang diujikan terdapat 606 data yang benar pengklasifikasiannya oleh metode SVM.

DAFTAR PUSTAKA

- Arta, K. J., Indrawan, G., & Dantes, G. R. (2019). Data mining rekomendasi calon mahasiswa berprestasi di STMIK Denpasar menggunakan metode technique form others reference by similarity to ideal solution. *Jurnal Ilmu Komputer Indonesia*, 4, 11–21.
- Audrin, C., & Audrin, B. (2022). Key factors in digital literacy in learning and education : a systematic literature review using text mining. *Education and Information Technologies*, 7395–7419. <https://doi.org/10.1007/s10639-021-10832-5>
- DBS. (2020). *E-commerce Indonesia menempati urutan pertama di Asia Tenggara*.
https://www.dbs.com/newsroom/Indonesias_ecommerce_ranks_first_in_Southeast_Asia
- Fernandes, A. L. (2019). *Perancangan e-commerce untuk memperluas produk komunikasi di PT . Golden Communication berbasis web mobile*. 4, 95–100.
- Fide, S., Suparti, & Sudarno. (2021). Analisis sentimen ulasan tiktok di google play menggunakan metode support vector machine (SVM) dan asosiasi. *JURNAL GAUSSIAN*, 10, 346–358.
- Haranto, F. F., & Sari, B. W. (2019). Implentasi support vector machine untuk analis sentimen pengguna twitter terhadap pelayanan telkom dan biznet. *Jurnal PILAR Nusa Mandiri Vol.*, 15(2), 171–176.
<https://doi.org/10.33480/pilar.v15i2.699>
- Hasna, S. K. (2021). *Analisis Sentimen Data Ulasan Menggunakan Algoritma Support Vector Machine (Studi Kasus: Aplikasi Iflix)*. 17522136.
- Hendriyanto, M. D., Ridha, A. A., & Enri, U. (2022). Analisis sentimen ulasan

- aplikasi mola pada google play store menggunakan algoritma support vector machine. *Journal of Information Technology and Computer Science (INTECOMS)*, 5, 1–7.
- Hermiati, R., Asnawati, & Kanedi, I. (2021). Pembuatan e-commerce pada raja komputer menggunakan bahasa pemrograman PHP dan database MYSQL. *Jurnal Media Infotama*, 17(1), 54–66.
- Iskandar, J. W., & Nataliani, Y. (2021). *Perbandingan naive bayes, SVM, dan KNN untuk analisis sentimen gadget berbasis aspek*. 5(158), 1120–1126.
- Manik, G., Ernawati, I., & Nurlaili, I. (2021). Analisis sentimen pada review pengguna e-Commerce bidang pangan menggunakan metode support vector machine (studi kasus : review sayurbox dan tanihub pada google play). *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya*, 64–74.
- Ortega, J., & Brotosaputro, G. (2019). Analisis sentimen tokoh politik pada situs berita menggunakan NER . Studi Kasus : IMMC. *Sistem Informasi Dan Teknologi*, 3, 212–216.
- Praptiwi, D. Y. (2018). *Analisis sentimen online review pengguna e-commerce menggunakan metode support vector dan maximum entropy*.
- Siringoringo, R., & Jamaluddin. (2019). Text mining dan klasterisasi sentimen pada ulasan produk toko online. *Jurnal Teknologi Dan Ilmu Komputer Prima*, 2, 314–319.
- Styawati, Hendrastuty, N., Isnain, A. R., & Rahmadhani, A. Y. (2021). Analisis sentimen masyarakat terhadap program kartu prakerja twitter dengan metode support vector machine. *Jurnal Informatika: Jurnal Pengembangan IT*, 6(3), 150–155.
- Tinaliah, & Elizabeth, T. (2022). Analisis sentimen ulasan aplikasi primaku menggunakan metode support vector machine. *Jurnal Teknik Informatika Dan Sistem Informasi*, 9(4), 3436–3442.

- Wahyudi, R., & Kusumawardana, G. (2021). Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *Jurnal Informatika*, 8(2), 200–207. <https://doi.org/10.31294/ji.v8i2.9681>
- Zuriel, H. P. P., & Fahrurozi, A. (2021). Implementasi algoritma klasifikasi support vector machine untuk analisa sentimen pengguna twitter terhadap kebijakan PSBB. 26, 149–162. <https://doi.org/10.35760/ik.2021>