

**OPTIMISASI NILAI PEFORMA ANALISIS SENTIMEN  
MENGUNAKAN HYPERPARAMETER TUNING DENGAN  
GRIDSEARCH PADA ULASAN APLIKASI SHOPEE**

**SKRIPSI**

Program Studi Sistem Informasi

Jenjang Sarjana S1



Oleh

**Muhammad Luthfi Al-Ghifari**  
**NIM 09031182025015**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
2023**

**LEMBAR PENGESAHAN**

**SKRIPSI**

**OPTIMISASI NILAI PEFORMA ANALISIS SENTIMEN MENGGUNAKAN  
HYPERPARAMETER TUNING DENGAN GRIDSEARCH PADA ULASAN APLIKASI  
SHOPEE**

Sebagai salah satu syarat untuk menyelesaikan studi di program studi  
Sistem Informasi SI

Oleh :

**Muhammad Luthfi Al-Ghifari**  
09031182025015

**Palembang, 21 Desember 2023**

**Mengetahui,**

**Ketua Jurusan Sistem Informasi**

  
**Endang Lestari Ruskan, M.T.**  
**NIP. 197811172006042001**

**Pembimbing,**

  
**Ken Ditha Tanja, M.Kom., Ph.D.**  
**NIP. 198507182013122003**

## HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan dibawah ini :

Nama : Muhammad Luthfi Al-Ghifari

NIM : 09031182025015

Program Studi : Sistem Informasi Bilingual

Judul Skripsi : Optimisasi Nilai Peforma Analisis Sentimen Menggunakan  
Hyperparameter Tuning Dengan Gridsearch Pada Ulasan Aplikasi  
Shopee

Hasil Pengecekan iThenticate/Turnitin : 7%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya saya sendiri dan bukan penjiplakan /plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku. Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 22 Desember 2023



Muhammad Luthfi Al-Ghifari  
NIM. 09031382025114

## HALAMAN PERSETUJUAN

Telah *accepted* jurnal di Indonesian Journal of Computer Science IJCS (SINTA 3) pada :

Hari : Senin

Tanggal : 18 September 2023

Nama : Muhammad Luthfi Ai-Ghifari

NIM : 09031182025015

Judul : Optimisasi Nilai Peforma Analisis Sentimen Menggunakan  
Hyperparameter Tuning Dengan Gridsearch Pada Ulasan Aplikasi  
Shopee

Tim Pembimbing :

1. Pembimbing : Ken Ditha Tania, M.Kom., PhD.



Mengetahui,

Ketua Jurusan Sistem Informasi



**Endang Lestari Ruskan, M.T.**  
**NIP. 197811172006042001**

## HALAMAN PERSEMBAHAN DAN MOTTO

*“Dan janganlah kamu (merasa) lemah, dan jangan (pula) bersedih hati, sebab kamu paling tinggi (derajatnya), jika kamu orang beriman.” (QS. Ali-Imran: 139)*

*” Great things are not done by impulse, but by a series of small things brought together” (Vincent van Gogh)*

**Skripsi ini saya persembahkan untuk :**

- 1. Allah SWT**
- 2. Diriku sendiri, Muhammad Luthfi Al-Ghifari**
- 3. Ayah, Ibuk, Cak Kiki, Uwuh beserta Seluruh Keluarga Besar**
- 4. Dosen Pembimbing Skripsi dan Dosen Pembimbing Akademik**
- 5. Dosen Sistem Informasi Universitas Sriwijaya**
- 6. Seluruh teman-teman SIBIL A Angkatan 2020 dan seperjuangan**
- 7. Seluruh teman-teman yang telah memberikan dukungan**
- 8. Fakultas Ilmu Komputer**
- 9. Almamaterku, Universitas Sriwijaya**

## **KATA PENGANTAR**

Puji dan syukur kepada Allah SWT atas ridha karunia-Nya, penulis dapat menyelesaikan laporan tugas akhir yang berjudul “Optimisasi Nilai Peforma Analisis Sentimen Menggunakan Hyperparamater Tuning Dengan Gridsearch Pada Ulasan Aplikasi Shopee” sebagai salah satu syarat menyelesaikan Pendidikan Jenjang Sarjana di program studi Sistem Infomasi pada Fakultas Ilmu Komputer Universitas Sriwijaya.

Selama penyusunan skripsi ini penulis banyak mendapatkan bantuan, bimbingan, dukungan dan doa dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis mengucapkan terima kasih kepada:

1. Allah SWT yang telah memberikan rahmat dan anugrah berupa ilmu yang bermanfaat, kesempatan, serta Kesehatan jasmani dan rohani sehingga penulis mampu menyelesaikan skripsi ini dengan tepat waktu.
2. Kedua orang tua yang selalu memberikan doa, semangat, dukungan, dan dorongan penuh kepada penulis untuk terus maju dan terus berkembang sampai kapanpun.
3. Kedua Saudara kandung yang juga selalu mendukung dan menyemangati penulis dalam menulis tugas akhir ini.
4. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Endang Lestari Ruskan, M.T. selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

6. Ibu Ken Ditha Tania, S.Kom., M.Kom. selaku Dosen Pembimbing yang telah membantu penulis dengan sepenuh hati selama pengerjaan tugas akhir ini, dengan bantuan dan bimbingan Beliau lah penulis dapat menjalankan dan menyelesaikan tugas akhir ini dengan baik dan benar.
7. Ibu Putri Eka Sevtiyuni, S.SI., M.T. selaku Dosen Pembimbing Akademik yang telah membimbing penulis dari semester awal hingga semester akhir.
8. Mba Rifka selaku admin Program Studi Sistem Informasi Bilingual yang selalu memberikan kemudahan dan membantu dalam segala urusan administrasi.
9. Teman-teman seperjuangan bimbingan Bu Ken yang telah banyak membantu dan berproses bersama penulis selama ini
10. Seluruh kelas Sistem Informasi Bilingual A 2020 yang telah menjadi teman kelas dan sangat banyak membantu dari awal semester sampai akhir ini.
11. Seluruh teman jurusan Sistem Informasi angkatan 20 yang berjuang menyelesaikan skripsi dan meraih gelar sarjana.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan dengan segala kekurangannya. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran terhadap skripsi ini.

Palembang, 22 Desember 2023



Muhammad Luthfi Al-Ghifari  
NIM. 09031382025114

# **OPTIMISASI NILAI PERFORMA ANALISIS SENTIMEN MENGGUNAKAN HYPERPARAMETER TUNING DENGAN GRIDSEARCH PADA ULASAN APLIKASI SHOPEE**

**Oleh :**

**Muhammad Luthfi Al-Ghifari**

**09031182025015**

## **ABSTRAK**

Perkembangan teknologi yang begitu pesat saat ini telah memberikan kemudahan bagi kita dalam peradaban saat ini. Salah satu perkembangan tersebut adalah ditemukannya internet karena tingginya penetrasi internet dan pesatnya pertumbuhan penggunaan ponsel, belanja online pun meningkat pesat. Belanja online ini sekarang sering disebut sebagai e-commerce. E-commerce adalah salah satu model perdagangan yang telah meluas karena pengaruh penggunaan teknologi yang luas. Secara khusus, e-commerce mengacu pada penggunaan Internet atau jaringan lain. Shopee merupakan salah satu marketplace populer di Indonesia yang memiliki jumlah pengunjung tertinggi yaitu 129 juta per bulan dan dapat diunduh di Google Play Store. Play Store sendiri memiliki beberapa fitur seperti Ulasan yang dapat memungkinkan pengguna untuk memberikan pendapat. Semua keluhan dan pendapat dari pengguna shopee dapat disalurkan ke dalam fitur ini. Dengan ini dilakukan penelitian yang bertujuan untuk mengoptimalkan nilai performansi analisis sentimen dengan metode Term Frequency-Inverse Document Frequency (TF-IDF) dan Hyperparameter Tuning dengan Gridsearch pada aplikasi Shopee di Google Play Store. Berdasarkan penelitian ulasan yang dihasilkan sebanyak 3000 data dimana 2015 data pengguna bernilai positif dan 985 data bernilai negatif. Data testing dibagi dengan perbandingan 90:10 untuk 300 data testing pada setiap model klasifikasi untuk mencari nilai akurasi. Dengan hyperparameter tuning menggunakan gridsearch kita dapat melihat hasil dari masing-masing nilai akurasi KNN, DCT, RF, dan LR meningkat dari 0.73 ke 0.77, 0.823 ke 0.826, 0.856 ke 0.87, dan 0.856 ke 0.866. Hal ini mengindikasikan bahwa di antara model machine learning yang sudah di-tuning dengan gridsearch, KNN merupakan model yang paling tinggi peningkatannya.

**Kata Kunci :** Analisis Sentimen, Shopee, TF-IDF, Hyperparameter Tuning, Gridsearch



# **SENTIMENT ANALYSIS PERFORMANCE VALUE OPTIMIZATION USING HYPERPARAMETER TUNNING WITH GRID SEARCH ON SHOPEE APP REVIEWS**

**By :**

**Muhammad Luthfi Al-Ghifari**

**09031182025015**

## **ABSTRACT**

The rapid development of technology today has provided convenience for us in today's civilization. One of these developments is the invention of the internet due to high internet penetration and rapid growth in mobile usage, online shopping has increased tremendously. This online shopping is now often referred to as e-commerce. E-commerce is one of the trade models that has been widened under the effect of extensive use of technology. Specifically, e-commerce refers to the usage of the Internet or other networks. Shopee is one of the popular marketplaces in Indonesia that has the highest number of visitors of 129 million per month and can be downloaded on the Google Play Store. Play Store itself has several features such as Reviews that can allow users to give opinions. All complaints and opinions from shopee users can be channeled into this feature. With this a research aims to optimize the performance value of sentiment analysis with the Term Frequency-Inverse Document Frequency (TF-IDF) method and Hyperparameter Tuning with Gridsearch for the Shopee application on the Google Play Store. Based on research the reviews resulting in 3000 data where 2015 user data is positive and 985 data is negative. Testing data was split by a ratio of 90:10 for 300 data test in each classification model to find the accuracy score. With hyperparameter tuning using gridsearch we can see the result of each accuracy score of KNN, DCT, RF, and LR is increasing from 0.73 to 0.77, 0.823 to 0.826, 0.856 to 0.87, and 0.856 to 0.866. This indicated that among the machine learning model that had been tuning using gridsearch, KNN is the one that highly increased.

**Keywords** : Sentiment Analysis, Shopee, TF-IDF, Hyperparameter Tuning, Gridsearch

## DAFTAR ISI

<b>LEMBAR PENGESAHAN .....</b>	<b>ii</b>
<b>HALAMAN PERNYATAAN BEBAS PLAGIAT .....</b>	<b>iii</b>
<b>HALAMAN PERSETUJUAN .....</b>	<b>iv</b>
<b>HALAMAN PERSEMBAHAN DAN MOTTO .....</b>	<b>v</b>
<b>KATA PENGANTAR.....</b>	<b>vi</b>
<b>ABSTRAK .....</b>	<b>viii</b>
<b>ABSTRACT .....</b>	<b>ix</b>
<b>DAFTAR ISI.....</b>	<b>x</b>
<b>DAFTAR GAMBAR.....</b>	<b>xiii</b>
<b>DAFTAR TABEL .....</b>	<b>xv</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xvi</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	5
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>6</b>
2.1 Landasan Teori .....	6
2.1.1 Shopee .....	6
2.1.2 Analisis Sentimen .....	7
2.1.3 <i>Google Play Store</i> .....	8
2.1.4 <i>Data Mining</i> .....	8
2.1.5 Klasifikasi Teks .....	9
2.1.6 <i>Web Scraping</i> .....	9
2.1.7 <i>Natural Language Toolkit (NLTK)</i> .....	10
2.1.8 Sastrawi .....	10
2.1.9 Pembobotan <i>TF-IDF (Term Frequency-Inverse Document Frequency)</i> .....	11
2.1.10 N-Gram .....	11
2.1.11 <i>Machine Learning</i> .....	11
2.1.12 <i>Confusion Matrix</i> .....	12

2.1.13 <i>K-Fold Cross Validation</i> .....	13
2.1.14 <i>Gridsearch</i> .....	13
2.2 Tahap Pengelolaan Data.....	14
2.2.1 <i>Pre-Processing</i> .....	14
2.2.2 <i>Case Folding</i> .....	14
2.2.3 <i>Stopword</i> .....	14
2.2.4 <i>Tokenizing</i> .....	14
2.2.5 <i>Stemming</i> .....	15
2.3 Penelitian Sebelumnya .....	15
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>17</b>
3.1 Desain Penelitian.....	17
3.2 Alat dan Bahan Penelitian .....	19
3.3 Metode Pengumpulan Data .....	19
3.4 Mengkategorikan Label Data .....	20
3.5 Tahapan Processing Data .....	21
3.5.1 <i>Case Folding</i> .....	21
3.5.2 <i>Stopword</i> .....	21
3.5.3 <i>Tokenizing</i> .....	22
3.5.4 <i>Stemming</i> .....	22
3.6 <i>TF-IDF (Term Frequency-Inverse Document Frequency)</i> .....	23
3.7 <i>Split Data</i> .....	23
3.8 <i>Modelling</i> .....	23
3.9 <i>Cross Validation Evaluation</i> .....	24
3.10 <i>Tuning Model</i> .....	24
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>25</b>
4.1 Hasil Proses <i>Crawling</i> Data .....	25
4.1.1 Hasil <i>Crawling</i> Data .....	25
4.2 Mengkategorikan Pelabelan Data.....	26
4.3 Hasil Preprocessing Data .....	28
4.3.1 Hasil <i>Case Folding</i> .....	28
4.3.2 Hasil <i>Stopword</i> .....	29
4.3.3 Hasil <i>Tokenizing</i> .....	30
4.3.4 Hasil <i>Stemming</i> .....	30
4.3.5 Hasil Akhir Preprocessing Data .....	31

4.4 Hasil Proses <i>TF-IDF (Term Frequency-Inverse Document Frequency)</i> .....	32
4.5 Hasil <i>Split Data</i> .....	33
4.6 Hasil Implementasi Algoritma <i>Machine Learning</i> .....	33
4.6.1 Hasil Algoritma K-Nearest Neighbor.....	34
4.6.2 Hasil Algoritma Decision Tree.....	35
4.6.3 Hasil Algoritma Random Forest.....	36
4.6.4 Hasil Algoritma Logistic Regression .....	38
4.7 Hasil Pengujian <i>Cross Validation Score</i> .....	40
4.7.1 Hasil Pengujian 3, 4, 5 K-Fold K-Nearest Neighbor.....	40
4.7.2 Hasil Pengujian 3, 4, 5 K-Fold Decision Tree.....	40
4.7.3 Hasil Pengujian 3, 4, 5 K-Fold Random Forest.....	41
4.7.4 Hasil Pengujian 3, 4, 5 K-Fold Logistic Regression .....	41
4.8 Hasil Implementasi <i>Hyperparameter Tuning</i> .....	42
4.8.1 Hasil Pengujian <i>Hyperparameter Tuning</i> K-Nearest Neighbor .....	42
4.8.2 Hasil Pengujian <i>Hyperparameter Tuning</i> Decision Tree .....	44
4.8.3 Hasil Pengujian <i>Hyperparameter Tuning</i> Random Forest .....	45
4.8.4 Hasil Pengujian <i>Hyperparameter Tuning</i> Logistic Regression.....	47
4.9 Hasil Penelitian .....	49
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>51</b>
5.1 Kesimpulan.....	51
5.2 Saran.....	52
<b>DAFTAR PUSTAKA .....</b>	<b>53</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>58</b>

## DAFTAR GAMBAR

<b>Gambar 2. 1</b>	Logo Aplikasi Shopee.....	6
<b>Gambar 2. 2</b>	K-Fold Cross Validation .....	13
<b>Gambar 3. 1</b>	Desain Penelitian .....	17
<b>Gambar 3. 2</b>	Hasil Ulasan Positif dan Negatif .....	20
<b>Gambar 4. 1</b>	Library Crawling Data.....	25
<b>Gambar 4. 2</b>	Crawling Data.....	26
<b>Gambar 4. 3</b>	Hasil Crawling Data Ulasan .....	26
<b>Gambar 4. 4</b>	Hasil Pelabelan Data.....	27
<b>Gambar 4. 5</b>	Hasil Pembagian Data .....	27
<b>Gambar 4. 6</b>	Grafik Hasil Pembagian Data .....	28
<b>Gambar 4. 7</b>	Kolom Data Yang Akan di Preprocessing .....	28
<b>Gambar 4. 8</b>	Data Sebelum dan Sesudah Case Folding .....	29
<b>Gambar 4. 9</b>	Hasil Proses Stopword.....	29
<b>Gambar 4. 10</b>	Hasil Proses Tokenize.....	30
<b>Gambar 4. 11</b>	Hasil Proses Stemming .....	31
<b>Gambar 4. 12</b>	Hasil Akhir Preprocessing Data.....	31
<b>Gambar 4. 13</b>	Proses Term Weighting .....	32
<b>Gambar 4. 14</b>	Pendefinisian X dan Y Untuk Modelling .....	32
<b>Gambar 4. 15</b>	Membagi Data Training dan Test .....	33
<b>Gambar 4. 16</b>	Library Yang Digunakan.....	34
<b>Gambar 4. 17</b>	Hasil Uji K-Nearest Neighbor .....	34
<b>Gambar 4. 18</b>	Confusion Matrix K-Nearest Neighbor .....	35
<b>Gambar 4. 19</b>	Hasil Uji Decision Tree .....	36
<b>Gambar 4. 20</b>	Confusion Matrix Decision Tree .....	36
<b>Gambar 4. 21</b>	Hasil Uji Random Forest .....	37
<b>Gambar 4. 22</b>	Confusion Matrix Random Forest .....	38
<b>Gambar 4. 23</b>	Hasil Uji Logistic Regression.....	39
<b>Gambar 4. 24</b>	Confusion Matrix Logistic Regression.....	39
<b>Gambar 4. 25</b>	Library dan K-Fold Cross Validation Score .....	40
<b>Gambar 4. 26</b>	Cross Validation Score K-Nearest Neighbor .....	40
<b>Gambar 4. 27</b>	Cross Validation Score Decision Tree .....	41
<b>Gambar 4. 28</b>	Cross Validation Score Random Forest .....	41
<b>Gambar 4. 29</b>	Cross Validation Score Logistic Regression.....	42
<b>Gambar 4. 30</b>	Library Gridsearch.....	42
<b>Gambar 4. 31</b>	Macam-Macam Parameter K-Nearest Neighbor .....	43
<b>Gambar 4. 32</b>	Memasukkan Best Model K-Nearest Neighbor.....	43
<b>Gambar 4. 33</b>	Skor Akurasi Setelah Hyperparameter Tuning .....	44
<b>Gambar 4. 34</b>	Macam-Macam Parameter Decision Tree .....	44
<b>Gambar 4. 35</b>	Memasukkan Best Model Decision Tree.....	45
<b>Gambar 4. 36</b>	Skor Akurasi Setelah Hyperparameter Tuning .....	45
<b>Gambar 4. 37</b>	Macam-Macam Parameter Random Forest .....	46

<b>Gambar 4. 38</b> Memasukkan Best Model Random Forest.....	46
<b>Gambar 4. 39</b> Skor Akurasi Setelah Hyperparameter Tuning .....	47
<b>Gambar 4. 40</b> Macam-Macam Parameter Logistic Regression.....	48
<b>Gambar 4. 41</b> Memasukkan Best Model Logistic Regression .....	48
<b>Gambar 4. 42</b> Skor Akurasi Setelah Hyperparameter Tuning .....	49

## DAFTAR TABEL

<b>Tabel 2. 1</b> Confusion Matrix.....	12
<b>Tabel 2. 2</b> Tabel Penelitian Terdahulu .....	15
<b>Tabel 3. 1</b> Contoh Case Folding .....	21
<b>Tabel 3. 2</b> Contoh Stopword.....	21
<b>Tabel 3. 3</b> Contoh Tokenizing .....	22
<b>Tabel 3. 4</b> Contoh Stemming.....	22

## DAFTAR LAMPIRAN

<b>Lampiran 1</b> Surat Ketersediaan Membimbing .....	A-1
<b>Lampiran 2</b> Form Pengajuan Topik Skripsi .....	B-1
<b>Lampiran 3</b> Surat Keputusan Tugas Akhir (SK TA) .....	C-1
<b>Lampiran 4</b> Form Desk Evaluasi.....	D-1
<b>Lampiran 5</b> Table Dataset Awal Hasil Penelitian .....	E-1
<b>Lampiran 6</b> Table Dataset Hasil Preprocessed Penelitian.....	F-1
<b>Lampiran 7</b> Bukti Proses Submit Jurnal.....	G-1
<b>Lampiran 8</b> Bukti LoA .....	H-1
<b>Lampiran 9</b> Hasil Pengecekan Turnitin Skripsi .....	I-1
<b>Lampiran 10</b> Surat Keterangan Pengecekan Similarity .....	J-1
<b>Lampiran 11</b> Kartu Konsultasi .....	K-1



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Peradaban modern telah menjadi lebih nyaman bagi kita berkat kemajuan teknologi yang cepat. Terciptanya internet merupakan salah satu kemajuan tersebut (Wahyudi et al., 2021). Dalam beberapa tahun terakhir, akibat tingginya angka pertumbuhan internet dan pesatnya pertumbuhan penggunaan ponsel, belanja online mengalami peningkatan yang luar biasa dan menjadi sangat populer (Karthik & Ganapathy, 2021). Teknologi internet telah membuat semua kegiatan manusia menjadi sangat praktis dan cepat, belanja online ini sekarang sering disebut sebagai *e-commerce*.

Salah satu konsep bisnis yang telah mendapatkan popularitas sebagai akibat dari meningkatnya penggunaan teknologi adalah *e-commerce*. *E-commerce* secara khusus mengacu pada pembelian, penjualan, dan perdagangan barang, jasa, atau data melalui Internet atau jaringan lain (seperti intranet). Ini benar-benar mengubah cara pelanggan berbelanja dan bagaimana bisnis berjalan. (Turban et al., 2017). Dengan kemajuan teknologi, platform *e-commerce* seperti Shopee, Tokopedia, Bukalapak telah muncul sebagai pasar online populer yang menghubungkan penjual dan pembeli dengan cara yang nyaman dan efisien (Salsabila Putri & Zakaria, 2020). Dengan demikian, *e-commerce* adalah salah satu media belanja yang sering digunakan orang akhir-akhir ini untuk memudahkan mereka membeli barang dari jarak jauh tanpa mengunjungi toko secara fisik. Mereka yang ingin menjual barang atau jasa dapat melakukannya hanya dengan mengklik layar perangkat, kemudian mereka dapat berbelanja dan bersantai sambil menunggu barang tiba di rumah mereka.

Dengan 129 juta kunjungan bulanan, Shopee adalah salah satu pasar paling terkenal di Indonesia. Shopee adalah aplikasi mobile marketplace (iPrice, 2017). *Google Play Store* menawarkan aplikasi jual beli elektronik Shopee untuk diunduh. *Play Store* adalah situs web yang menawarkan berbagai aplikasi online untuk diunduh, termasuk musik, film, buku, dan kategori lainnya. (Herlinawati et al., 2020). *Play Store* sendiri memiliki beberapa fitur seperti *Reviews* yang dapat membuat pengguna dapat memberikan pendapat atau komentar terhadap karya orang lain. *Rating* yang dapat membuat pengguna memberikan nilai pada suatu aplikasi. Semua keluhan dan pendapat dari pengguna shopee dapat disalurkan ke dalam fitur ini pada *Google Play Store* (Rezki et al., 2020). Analisis sentimen digunakan untuk menentukan bagaimana reaksi pengguna terhadap aplikasi Shopee dan apakah ulasan positif atau negatif tergantung pada komentar atau peringkat yang diterima setiap ulasan.

Penelitian serupa pernah dilakukan oleh Christian dan Yessica (Cahyaningtyas et al., 2021). Menggunakan model Decision Tree menggunakan SMOTE. Penelitian ini mengenai analisis sentiment pada rating aplikasi Shopee yang menunjukkan nilai akurasi yang dicapai adalah 0,02% dan AUC adalah 0,049 saat menggunakan SMOTE tanpa menggunakan SMOTE. menunjukkan bagaimana alat analisis sentimen sesuai dan berguna untuk memberikan referensi dan umpan balik kepada pengguna di aplikasi Shopee.

Menurut riset Sri Rahayu (Rahayu et al., 2022) tentang pemilihan algoritma dan metode analisis sentimen, sentimen pengguna terhadap aplikasi teknologi finansial FLIP dapat dianalisis menggunakan metode K-Nearest Neighbor (K-NN) dengan pembobotan berdasarkan metode Term Frequency-Inverse Document Frequency (TF-IDF). Empat belas ratus tiga puluh tujuh data uji ditemukan

diidentifikasi secara akurat berdasarkan temuan pemeriksaan klasifikasi. Algoritma K-NN memiliki tingkat akurasi klasifikasi 76,68%. Hasil pemeriksaan klasifikasi juga mengungkapkan bahwa kelas data dengan ulasan positif memiliki presisi 82,67%, sedangkan kelas dengan ulasan negatif memiliki presisi 65,46% dan kelas dengan ulasan netral memiliki presisi 57,45%. Kelas data dengan evaluasi baik memiliki nilai recall 86,92%, 46,92% untuk kelas dengan ulasan negatif, dan 67,23% untuk kelas dengan ulasan netral. Temuan ini menunjukkan bahwa, ketika datang untuk menilai sentimen pengguna terhadap aplikasi FLIP, pendekatan K-NN dengan pembobotan TF-IDF dapat menghasilkan beberapa temuan yang sangat positif.

Sebuah studi deteksi kanker payudara berbasis K-Nearest Neighbor yang dioptimalkan telah diterbitkan (Assegie, 2021) menggunakan dataset dalam penelitian yang berasal dari repositori data kanker payudara Wisconsin. *Hyperparameter tuning* memiliki pengaruh yang signifikan terhadap kinerja model KNN untuk deteksi kanker payudara. Penelitian yang dilakukan oleh Tsehay Admassu Assegie menunjukkan bahwa performa model KNN meningkat secara signifikan ketika nilai hyper-parameter atau nilai K terbaik digunakan untuk pelatihan. Akurasi model menggunakan *default* parameter KNN menghasilkan nilai akurasi sebesar 90,10%. Sedangkan akurasi menggunakan *hyperparameter tuning* dengan *gridsearch* meningkat sebesar 94,35%. Hal ini menunjukkan bahwa *hyperparameter tuning* meningkatkan akurasi dan kinerja model KNN untuk deteksi kanker payudara.

Mengingat informasi latar belakang yang diberikan di atas, penulis akan masuk ke detail tambahan dalam proyek akhir di bawah judul **“Optimisasi Nilai Performa Analisis Sentimen Menggunakan Hyperparameter Tuning Dengan**

## **Gridsearch Pada Ulasan Aplikasi Shopee”**

### **1.2 Rumusan Masalah**

Berikut rumusan masalah untuk dibahas dapat dibuat berdasarkan uraian latar belakang :

1. Cara menggunakan *gridsearch* dan metode *hyperparameter tuning* untuk menganalisis sentimen di ulasan aplikasi Shopee dalam menaikkan nilai performa.
2. Berapa tingkat nilai keakuratan analisis sentimen Dalam ulasan aplikasi shopee sebelum menggunakan *hyperparameter tuning* dan sesudahnya.
3. Apakah tingkat nilai keakuratannya menurun atau meningkat.

### **1.3 Tujuan Penelitian**

Tujuan dari penelitian ini termasuk pada hal-hal berikut :

1. Memeriksa kemampuan analisis sentimen pada evaluasi aplikasi shopee *Google Play Store*.
2. Mengetahui seberapa akurat evaluasi analisis sentimen aplikasi Shopee di *Google Play Store* dan membuatnya lebih akurat.
3. Mengetahui komentar dan umpan balik mengenai aplikasi Shopee di *Google Play Store*.

### **1.4 Manfaat Penelitian**

Berikut ini adalah beberapa keuntungan yang akan dihasilkan oleh penelitian ini :

1. Untuk membaca evaluasi pengguna di *Google Play Store* mengenai layanan aplikasi Shopee.
2. Untuk menilai tingkat kebenaran ulasan di *Google Play Store* untuk aplikasi Shopee.

3. Untuk meningkatkan tingkat akurasi supaya ulasan yang didapat semakin tepat sasaran.

### **1.5 Batasan Masalah**

Berikut ini adalah Batasan masalah dalam penelitian ini :

1. Ulasan aplikasi Shopee dan data komentar dari *Google Play store* digunakan dalam penelitian ini.
2. Ada 3.000 data ulasan dalam bahasa Indonesia yang digunakan dalam penyelidikan ini.
3. Ada dua jenis klasifikasi data yang digunakan : positif dan negatif.
4. Excel dan Google Colab adalah perangkat lunak yang digunakan untuk pemrosesan data.

## DAFTAR PUSTAKA

- Aji Andika, L., & Amalia Nur Azizah, P. (2019). *Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier*.
- Albab, M. U., Karuniawati P, Y., & Fawaiq, M. N. (2023). *Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic*. 20(2), 1–10. <https://doi.org/10.26623/transformatika.v20i2.5374>
- Alhamid, M. (2020, December 25). *What is Cross-Validation?* Towards Data Science. <https://towardsdatascience.com/what-is-cross-validation-60c01f9d9e75>
- Anthony, L., & Sama, H. (2021). *Studi Kualitatif Mengenai Faktor Penerimaan Aplikasi E-Commerce Shopee Dan Fintech Shopeepay Bagi Masyarakat Senior* (Vol. 1, Issue 1). <https://journal.uib.ac.id/index.php/combines>
- Apriliwanto, R. E., Sanjaya, A., & Widodo, D. W. (2021). *Identifikasi Pola Kalimat Bahasa Indonesia Pada Siswa Sekolah Dasar Menggunakan Metode LALR dan Stemming*.
- Arifin, N., Enri, U., & Sulistiyowati, N. (2021). *STRING (Satuan Tulisan Riset dan Inovasi Teknologi) PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) DENGAN TF-IDF N-GRAM UNTUK TEXT CLASSIFICATION*.
- Assegie, T. A. (2021). An optimized K-Nearest neighbor based breast cancer detection. *Journal of Robotics and Control (JRC)*, 2(3), 115–118. <https://doi.org/10.18196/jrc.2363>
- Cahyaningtyas, C., Nataliani, Y., & Widiasari, I. R. (2021). Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE. *AITI: Jurnal Teknologi Informasi*, 18(Agustus), 173–184.
- Cahyanti, D., Rahmayani, A., & Ainy Husniar, S. (2020). *Indonesian Journal of Data and Science Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara*. 1(2), 39–43.
- Djufri, M. (2020). *JURNAL BPPK PENERAPAN TEKNIK WEB SCRAPING UNTUK PENGALIAN POTENSI PAJAK (Studi Kasus pada Online Market Place Tokopedia, Shopee dan Bukalapak)* (Vol. 13).
- Fatmawati, & Rifai, N. A. K. (2023). Klasifikasi Penyakit Diabetes Retinopati Menggunakan Support Vector Machine dengan Algoritma Grid Search Cross-validation. *Jurnal Riset Statistika*, 79–86. <https://doi.org/10.29313/jrs.v3i1.1945>
- George, S., & Sumathi, B. (2020). Grid Search Tuning of Hyperparameters in Random Forest Classifier for Customer Feedback Sentiment Prediction. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 11, Issue 9). [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- Gunawan, K. I., & Santoso, J. (2021). Multilabel Text Classification Menggunakan SVM dan Doc2Vec Classification Pada Dokumen Berita Bahasa Indonesia. *Journal of Information System, Graphics, Hospitality and Technology*, 3(01), 29–38. <https://doi.org/10.37823/insight.v3i01.126>

- Gunawan, M. I., Sugiarto, D., & Mardianto, I. (2020). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Peningkatan Kinerja Akurasi Prediksi Penyakit Diabetes Mellitus Menggunakan Metode Grid Search pada Algoritma Logistic Regression. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*.
- Hartono, F., Lim, R., & Dewi, L. P. (2020). *Pembuatan Sistem Rumah Pintar dengan Voice Assistant di Raspberry Pi*.
- Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., Komputer STMIK Nusa Mandiri Jl Damai No, I., Jati Barat, W., & Selatan, J. (2020). *ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE* (Vol. 5, Issue 2).
- Homepage, J., Roihan, A., Abas Sunarya, P., & Rafika, A. S. (2019). IJCIT (Indonesian Journal on Computer and Information Technology) Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. In *IJCIT (Indonesian Journal on Computer and Information Technology)* (Vol. 5, Issue 1).
- iPrice, T. R. (2017, December 21). *[ANALISIS] Kilas Balik Persaingan E-Commerce Indonesia Tahun 2017*. IPrice. <https://iprice.co.id/trend/insights/kilas-balik-e-commerce-indonesia-2017/>
- Kabir Hashi, E., & Shahid Uz Zaman, M. (2020). Developing a Hyperparameter Tuning Based Machine Learning Approach of Heart Disease Prediction. *Journal of Applied Science & Process Engineering*, 7(2).
- Karthik, R. V., & Ganapathy, S. (2021). A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. *Applied Soft Computing*, 108. <https://doi.org/10.1016/j.asoc.2021.107396>
- Kartika, M., & Ganarsih, R. L. (2019). *ANALISIS E-WOM, ONLINE SHOPPING EXPERIENCE DAN TRUST TERHADAP KEPUTUSAN PEMBELIAN DAN KEPUASAN KONSUMEN E-COMMERCE SHOPEE PADA MAHASISWA PASCASARJANA UNIVERSITAS RIAU*.
- Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 406. <https://doi.org/10.30865/mib.v5i2.2835>
- Kosasih, R., & Alberto, A. (2021). Sentiment analysis of game product on shopee using the TF-IDF method and naive bayes classifier. *ILKOM Jurnal Ilmiah*, 13(2), 101–109. <https://doi.org/10.33096/ilkom.v13i2.721.101-109>
- Mardiana, L., Kusnandar, D., & Satyahadewi, N. (2022). ANALISIS DISKRIMINAN DENGAN K FOLD CROSS VALIDATION UNTUK KLASIFIKASI KUALITAS AIR DI KOTA PONTIANAK. In *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)* (Vol. 11, Issue 1).
- Mulyani, E., Muhamad, F. P. B., & Cahyanto, K. A. (2021). Pengaruh N-Gram terhadap Klasifikasi Buku menggunakan Ekstraksi dan Seleksi Fitur pada Multinomial Naïve

- Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 264.  
<https://doi.org/10.30865/mib.v5i1.2672>
- Mutawalli, L., Taufan, M., Zaen, A., & Bagye, W. (2019). KLASIFIKASI TEKS SOSIAL MEDIA TWITTER MENGGUNAKAN SUPPORT VECTOR MACHINE (Studi Kasus Penusukan Wiranto). In *Jurnal Informatika & Rekayasa Elektronika* (Vol. 2, Issue 2).  
<http://e-journal.stmiklombok.ac.id/index.php/jire>
- Novianti, D., Nusa, S., Jakarta, M., & Sitasi, C. (2019). *Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner*. 21(1).  
<https://doi.org/10.31294/p.v20i2>
- Oryza Habibie Rahman, Gunawan Abdillah, & Agus Komarudin. (2021). Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 17–23.  
<https://doi.org/10.29207/resti.v5i1.2700>
- Pahlevi, R., & Ziveria, M. (2022). *Perancangan Deteksi Kemiripan pada Abstrak Artikel Ilmiah Informatika* (Vol. 8, Issue 2).
- Pratama, R. R. (2020). Analisis Model Machine Learning Terhadap Pengenalan Aktifitas Manusia. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 19(2), 302–311. <https://doi.org/10.30812/matrik.v19i2.688>
- Priyadarshini, I., & Cotton, C. (2021). A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis. *Journal of Supercomputing*, 77(12), 13911–13932.  
<https://doi.org/10.1007/s11227-021-03838-w>
- Purbaya, M. E., Rakhmadani, D. P., Maliana Puspa Arum, & Luthfi Zian Nasifah. (2023). Implementation of n-gram Methodology to Analyze Sentiment Reviews for Indonesian Chips Purchases in Shopee E-Marketplace. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(3), 609–617. <https://doi.org/10.29207/resti.v7i3.4726>
- Rachman, F. F., & Pramana, S. (2020). Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. In *Health Information Management Journal ISSN* (Vol. 8, Issue 2).
- Rahayu, S., MZ, Y., Bororing, J. E., & Hadiyat, R. (2022). Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 98–106.  
<https://doi.org/10.29408/edumatic.v6i1.5433>
- Rais, Z., Said, R. N., & Ruliana, R. (2022). Text Classification on Sentiment Analysis of Marketplace SHOPEE Reviews On Twitter Using K-Nearest Neighbor (KNN) Method. *JINAV: Journal of Information and Visualization*, 3(1), 1–8.  
<https://doi.org/10.35877/454ri.jinav1389>
- RAMADHAN, M. M., SITANGGANG, I. S., NASUTION, F. R., & GHIFARI, A. (2017). Parameter Tuning in Random Forest Based on Grid Search Method for Gender Classification Based on Voice Frequency. *DEStech Transactions on Computer Science and Engineering, cece*. <https://doi.org/10.12783/dtscse/cece2017/14611>



- Reddy, G. T., Bhattacharya, S., Siva Ramakrishnan, S., Chowdhary, C. L., Hakak, S., Kaluri, R., & Praveen Kumar Reddy, M. (2020, February 1). An Ensemble based Machine Learning model for Diabetic Retinopathy Classification. *International Conference on Emerging Trends in Information Technology and Engineering, Ic-ETITE 2020*. <https://doi.org/10.1109/ic-ETITE47903.2020.235>
- Rezki, M., Kholifah, D. N., Faisal, M., Suryadithia, R., Bina, U., & Informatika, S. (2020). Analisis Review Pengguna Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes. In *Jurnal* (Vol. 2, Issue 2). <http://ejournal.bsi.ac.id/ejurnal/index.php/infortech264>
- Salsabila Putri, A., & Zakaria, R. (2020). *Analisis Pemetaan E-Commerce Terbesar Di Indonesia Berdasarkan Model Kekuatan Ekonomi Digital*.
- Sari, F. V., & Wibowo, A. (2019). ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI. *Jurnal SIMETRIS*, 10(2).
- Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). *JOURNAL OF INTELLIGENT SYSTEMS AND COMPUTATION* 43. <https://t.co/9Wl0aWpfd5>
- Setiawan, D. F., Tristiyanto, T., & Hijriani, A. (2020). APLIKASI WEB SCRAPING DESKRIPSI PRODUK. *Jurnal Teknoinfo*, 14(1), 41. <https://doi.org/10.33365/jti.v14i1.498>
- Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). ANALISIS DATA MINING DATA NETFLIX MENGGUNAKAN APLIKASI RAPID MINER. *JBASE - Journal of Business and Audit Information Systems*, 4(1). <https://doi.org/10.30813/jbase.v4i1.2729>
- Suryani, P. S. M., Linawati, L., & Saputra, K. O. (2019). Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia. *Majalah Ilmiah Teknologi Elektro*, 18(1), 145. <https://doi.org/10.24843/mite.2019.v18i01.p22>
- Tarigan, V. (2023). *Pembuatan Aplikasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma Naive Bayes*. 11(1).
- Thomas, V. W. D., & Rumaisa, F. (2022). Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(3), 1767. <https://doi.org/10.30865/mib.v6i3.4218>
- Turban, E., Whiteside, J., King, D., & Outland, J. (2017). *Introduction to Electronic Commerce and Social Commerce*. <http://www.springer.com/series/10099>
- Wahyudi, R., Kusumawardhana, G., Purwokerto, A., Letjend, J., Soemarto, P., Purwanegara, K., Purwokerto, T., & Banyumas, K. (2021). Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *JURNAL INFORMATIKA*, 8(2). <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- Wijanarto, W., & Brilianti, S. P. (2020). Peningkatan Performa Analisis Sentimen Dengan Resampling dan Hyperparameter pada Ulasan Aplikasi BNI Mobile. *Jurnal Eksplora Informatika*, 9(2), 140–153. <https://doi.org/10.30864/eksplora.v9i2.333>

- Witanti, A., Yogyakarta Jl Raya Wates-Jogjakarta, B., Sedayu, K., Bantul, K., & Istimewa Yogyakarta, D. (2022). ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM). *Jurnal Sistem Informasi Dan Informatika (Simika) P-ISSN*, 5, 2622–6901.
- Yutika, C. H., Adiwijaya, A., & Faraby, S. Al. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 422. <https://doi.org/10.30865/mib.v5i2.2845>