

Universitas Katolik Santo Thomas

Repositori Unika Santo Thomas

<http://eprints.ust.ac.id>

---

Fakultas Ilmu Komputer (FIKOM)  
Program Studi Teknik Informatika

in the Undergraduate Papers

---

Purba, Brian Marthin  
2025

# Perbandingan Akurasi Analisis Sentimen Kepuasan Pengguna Shopee Dengan Metode Naïve Bayes Dan Support Vector Machine

---

<http://eprints.ust.ac.id/id/eprint/999>

*Downloaded from Repositori Institusi UST, Universitas Katolik Santo Thomas*

**PERBANDINGAN AKURASI ANALISIS SENTIMEN  
KEPUASAN PENGGUNA SHOPEE DENGAN  
METODE NAÏVE BAYES DAN SUPPORT  
VECTOR MACHINE**

**SKRIPSI**

**Diajukan Untuk Melengkapi dan Memenuhi Persyaratan  
Memperoleh Gelar Sarjana Komputer Pada  
Program Studi Teknik Informatika**

**Oleh :**

**BRIAN MARTHIN PURBA**

**190840024**



**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS KATOLIK SANTO THOMAS  
MEDAN  
2025**

## PERNYATAAN PENULIS

Sebagai sivitas akademik Universitas Katolik Santo Thomas, saya yang bertanda tangan dibawah ini:

Nama : Brian Marthin Purba  
NPM : 190840024  
Program Studi : Teknik Informatika  
Fakultas : Ilmu Komputer  
Jenis Karya : Skripsi

1. Menyatakan dan bertanggungjawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri, kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut.
2. Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Katolik Santo Thomas Hak Bebas Royalti Nonesklusif (Non Exclusive Royalty-FreeRight) atas karya ilmiah saya tersebut, beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Nonesklusif ini, Universitas Katolik Santo Thomas berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Universitas Katolik Santo Thomas  
Tanggal: 8 April 025  
Yang menyatakan

\

Brian Marthin Purba  
Penulis



## LEMBAR PERSETUJUAN SKRIPSI

NAMA : BRIAN MARTHIN PURBA  
NIM : 190840024  
PRGRAM STUDI : TEKNIK INFORMATIKA  
JENJANG : STRATA SATU (S-1)  
JUDUL : PERBANDINGAN AKURASI ANALISIS  
SENTIMEN KEPUASAN PENGGUNA SHOPEE  
DENGAN METODE NAÏVE BAYES DAN  
SUPPORT VECTOR MACHINE





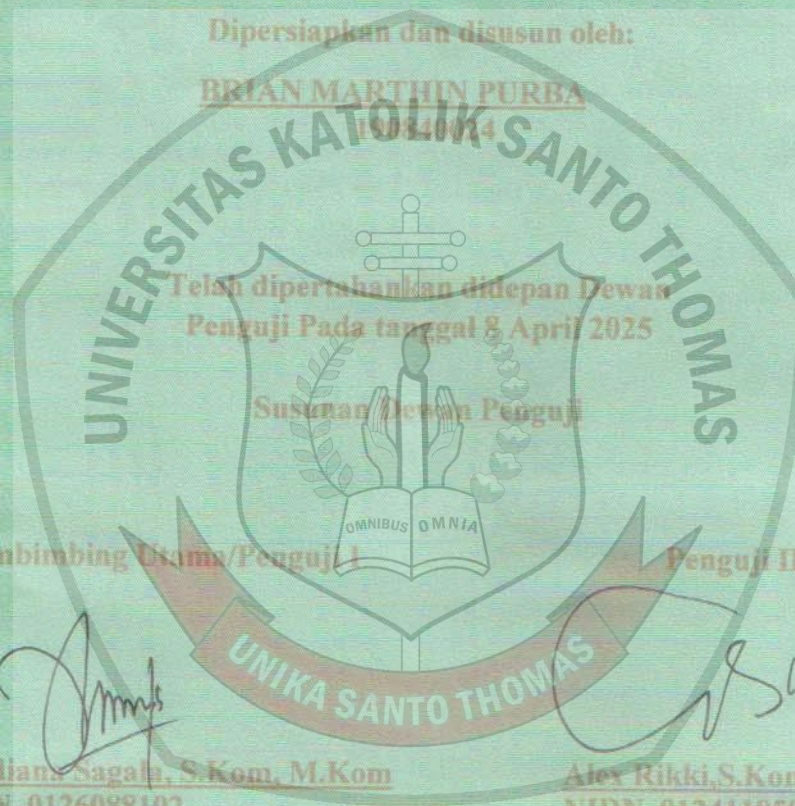
## SKRIPSI

# PERBANDINGAN AKURASI ANALISIS SENTIMEN KEPUASAN PENGGUNA SHOPEE DENGAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

Dipersiapkan dan disusun oleh:

BRIAN MARTIIN PURBA

1008240014



Masdiana Sagala, S.Kom, M.Kom  
NIDN. 0126088102

Alex Rikki, S.Kom, M.Kom  
NIDN. 0124118502

Penguji III

Zekson Arizona Matondang, S.Kom, M.Kom.  
NIDN. 0125048801

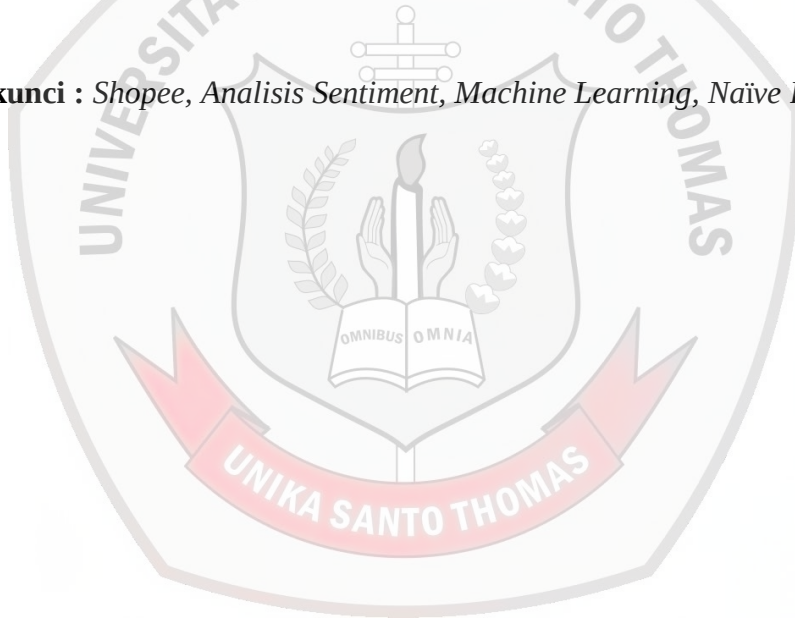


## ABSTRAK

Perkembangan zaman yang semakin pesat banyak teknologi yang semakin berkembang dari tahun ke tahun. Salah satunya yaitu masyarakat yang dulunya berbelanja dengan pergi ke toko beralih dengan berbelanja Online shop atau yang disebut dengan toko online. sistem belanja online yang mana pembeli dapat menanyakan harga atau pertanyaan apapun terkait produk kepada penjual secara langsung, dapat melalui BBM, LINE, Facebook, Instagram atau Whatsapp. Shopee merupakan salah satu toko online yang digunakan oleh masyarakat.

Pada penelitian ini, penulis akan melakukan analisis sentimen pada shopee dengan menggunakan tipe pendekatan metode machine learning. Metode klasifikasi Machine Learning yang digunakan dalam penelitian ini menggunakan metode algoritma Naïve Bayes dan Support Vector Machine. Hal ini bertujuan untuk mendapatkan informasi yang berguna bagi pengguna maupun pihak toko di shopee. Dataset yang akan digunakan pada penelitian ini diambil dari twitter. Untuk melakukan perbandingan metode algoritma Naïve Bayes dan Support Vector Machine menggunakan tools yaitu aplikasi RapidMiner.

**Kata kunci :** *Shopee, Analisis Sentiment, Machine Learning, Naïve Bayes*



## KATA PENGANTAR

Dengan mengucapkan puji dan syukur atas kehadiran Tuhan Yang Maha Esa atas berkat dan kasih-Nya, penulis dapat menyelesaikan penyusunan Skripsi. Laporan Skripsi dengan judul **“PERBANDINGAN AKURASI ANALISIS SENTIMEN KEPUASAN PENGGUNA SHOPEE DENGAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE”**.

Ucapan Terimakasih sebesar-besarnya penulis ucapkan kepada semua orang yang telah membantu dalam proses pembuatan Laporan Skripsi ini baik dari awal hingga selesai, di antaranya :

1. Kedua orangtua yang sangat disayangi dan dicintai penulis, Bapak E. Purba dan Mama R. Saragih yang selalu memberikan doa yang tiada henti, kasih sayang, nasehat, motivasi, dukungan serta selalu memenuhi kebutuhan penulis selama melalui proses perkuliahan. Dalam keadaan sesulit apapun namun mereka selalu bekerja keras demi memberikan kesempatan pendidikan yang lebih baik kepada ketiga anaknya. Terima kasih atas segala pengorbanan, kesabaran, dan kepercayaan yang telah diberikan kepada penulis selama menempuh pendidikan hingga ke tahap ini. Setiap doa dan restu yang diberikan menjadi kekuatan besar bagi penulis dalam menyelesaikan tugas akhir ini.
2. Bapak Prof.Dr.Maidin Gultom, SH., M.Hum, selaku Rektor Universitas Katolik Santo Thomas Medan.
3. Ibu Desinta Purba,ST.,M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Katolik Santo Thomas Medan.
4. Ibu Masdiana Sagala, S.Kom,M.Kom selaku Kepala Prodi Teknik Informatika.
5. Bapak Doni El Rezen Purba, S.Kom, M.Kom, selaku dosen pembimbing yang telah meluangkan waktu, serta memberikan arahan dan bimbingan selama penyelesaian Skripsi ini.
6. Bapak/Ibu Dosen Fakultas Ilmu Komputer Program Studi S1 serta seluruh Pegawai.

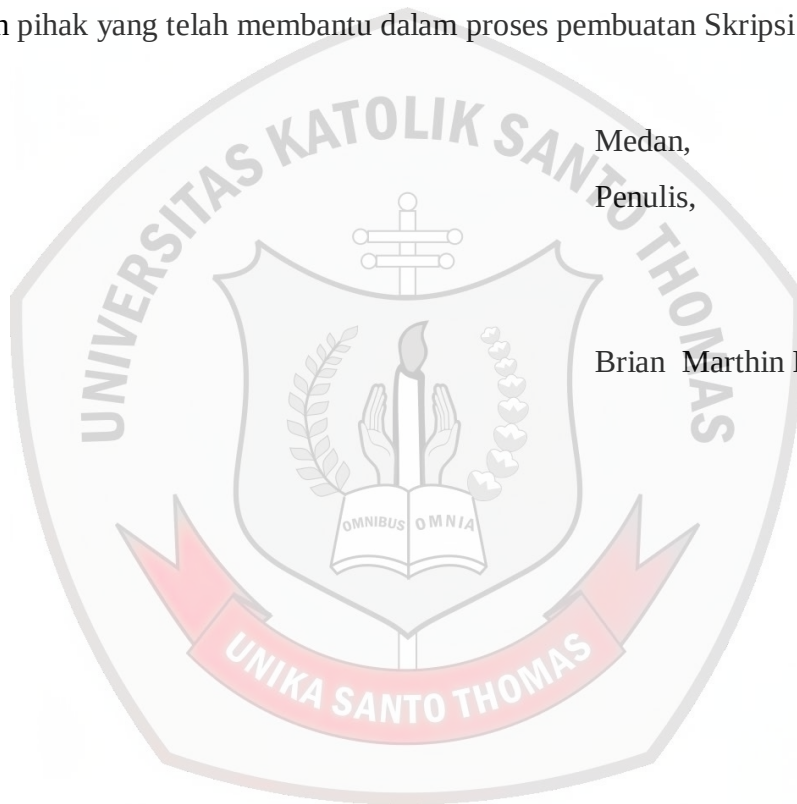
7. Kepada saudara saudari penulis, Nopsa Purba, Olivia Purba dan Sepria Saragih, terimakasih selalu memberi semangat dalam penyelesain tugas akhir ini.
8. Kepada Cikiki Manurung yang selalu mendukung dan menyemangati penulis.
9. Saudara-saudara serta rekan-rekan mahasiswa Fakultas Ilmu Komputer stambuk 19 dengan keakraban dan persaudarannya banyak membantu dalam meyelesaikan Skripsi ini.

Semoga Tuhan Yang Maha Esa selalu melimpahkan berkatnya kepada seluruh pihak yang membantu. Sekali lagi terima kasih penulis ucapkan kepada seluruh pihak yang telah membantu dalam proses pembuatan Skripsi ini.

Medan,  
Penulis,

2025

Brian Marthin Purba





## DAFTAR ISI

|                                         |            |
|-----------------------------------------|------------|
| <b>PERNYATAAN PENULIS .....</b>         | <b>i</b>   |
| <b>LEMBAR PERSETUJUAN SKRIPSI .....</b> | <b></b>    |
| ..... Error! Bookmark not defined.      |            |
| <b>SKRIPSI.....</b>                     | <b></b>    |
| ..... Error! Bookmark not defined.      |            |
| <b>ABSTRAK.....</b>                     | <b>ii</b>  |
| <b>KATA PENGANTAR.....</b>              | <b>v</b>   |
| <b>DAFTAR ISI .....</b>                 | <b>vii</b> |
| <b>DAFTAR TABEL .....</b>               | <b>x</b>   |
| <b>DAFTAR GAMBAR.....</b>               | <b>xi</b>  |
| <b>BAB I PENDAHULUAN .....</b>          | <b>1</b>   |
| I.1 Latar Belakang Masalah.....         | 1          |
| I.2 Rumusan Masalah.....                | 1          |
| I.3 Batasan Masalah .....               | 1          |
| I.4 Maksud dan Tujuan Penelitian.....   | 2          |
| I.5 Manfaat Penelitian .....            | 2          |
| I.6 Metodologi Penelitian .....         | 2          |
| I.7 Sistematika Penulisan.....          | 3          |
| <b>BAB II LANDASAN TEORI.....</b>       | <b>4</b>   |
| II.1 Analisis Sentimen .....            | 4          |
| II.2 Machine Learning.....              | 4          |
| II.3 Shopee.....                        | 5          |
| II.4 Twitter .....                      | 5          |
| II.5 Confusion Matrix.....              | 6          |
| II.6 Pembobotan TF_IDF .....            | 7          |
| II.7 Naïve Bayes.....                   | 8          |
| II.8 Support Vector Machine .....       | 9          |
| II.9 Flowchart.....                     | 12         |

|                                                            |           |
|------------------------------------------------------------|-----------|
| II.10 Microsoft Excel .....                                | 13        |
| II.11 RapidMiner .....                                     | 14        |
| <b>BAB III ANALISIS DAN PERANCANGAN .....</b>              | <b>15</b> |
| III.1 Analisis Model .....                                 | 15        |
| III.1.1 Pengumpulan Data.....                              | 16        |
| III.1.2 Cleansing.....                                     | 17        |
| III.1.3 Pre-processing Data .....                          | 18        |
| III.1.3 Pembobotan Kata dengan TF-IDF .....                | 24        |
| III.1.5 Pemodelan Data.....                                | 26        |
| III.1.6 Evaluasi.....                                      | 32        |
| III.2 Perancangan Desain Model .....                       | 32        |
| III.2.1 Rancangan Pengumpulan Data.....                    | 32        |
| III.2.2 Rancangan Cleansing.....                           | 33        |
| III.2.3 Rancangan Pre-preprocessing .....                  | 33        |
| III.2.4 Pembobotan TF-IDF dan Pemilihan Fitur Seleksi..... | 34        |
| III.2.5 Rancangan Pemodelan Data.....                      | 34        |
| <b>BAB IV IMPLEMENTASI MODEL .....</b>                     | <b>36</b> |
| IV.1 Implementasi Model.....                               | 36        |
| IV.1.1 Tampilan Pengumpulan Data.....                      | 36        |
| IV.1.2 Tampilan Data Cleansing.....                        | 36        |
| IV.1.3 Tampilan data Pre-processing .....                  | 36        |
| IV.1.4 Tampilan pemodelan Data .....                       | 36        |
| IV.2 Evaluasi Model .....                                  | 37        |
| IV.2.1 Data Awal .....                                     | 37        |
| IV.2.2 Data Cleansing .....                                | 37        |
| IV.2.3 Hasil Transform Cases.....                          | 38        |
| IV.2.4 Filter Stopword.....                                | 39        |
| IV.2.5 Tokenizing .....                                    | 39        |



|                             |           |
|-----------------------------|-----------|
| IV.2.6 Filter Token.....    | 40        |
| IV.2.7 Stem.....            | 40        |
| <b>BAB VI PENUTUP .....</b> | <b>42</b> |
| VI.1 Kesimpulan .....       | 42        |
| VI.2 Saran .....            | 42        |
| <b>DAFTAR PUSTAKA .....</b> | <b>43</b> |



## DAFTAR TABEL

|                                                              |           |
|--------------------------------------------------------------|-----------|
| <b>Tabel II. 1 Confusion Matrix.....</b>                     | <b>6</b>  |
| <b>Tabel II. 2 Simbol Flowchart.....</b>                     | <b>13</b> |
| <b>Tabel III. 1 Proses Cleansing.....</b>                    | <b>18</b> |
| <b>Tabel III. 2 Proses Cleansing.....</b>                    | <b>19</b> |
| <b>Tabel III. 3 Contoh Proses Transform Cases .....</b>      | <b>20</b> |
| <b>Tabel III. 4 Proses Filter Stopword.....</b>              | <b>22</b> |
| <b>Tabel III. 5 Proses Stem.....</b>                         | <b>23</b> |
| <b>Tabel III. 6 Tweets Data Uji.....</b>                     | <b>24</b> |
| <b>Tabel III. 7 Tabel Pembobotan Kata dengan TF-IDF.....</b> | <b>25</b> |
| <b>Tabel III. 8 Tweet Data Testing.....</b>                  | <b>26</b> |
| <b>Tabel III. 9 Matrix <math>xjT</math> .....</b>            | <b>30</b> |
| <b>Tabel III. 10 Kemunculan Term Pada Kalimat Baru .....</b> | <b>31</b> |
| <b>Tabel III. 11 Bobot Term Baru Pada Data .....</b>         | <b>31</b> |
| <b>Tabel III. 12 Matrix K .....</b>                          | <b>31</b> |
| <b>Tabel III. 13 Matrix y .....</b>                          | <b>31</b> |
| <b>Tabel III. 14 Matrix <math>x,xi</math> .....</b>          | <b>31</b> |

## DAFTAR GAMBAR

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| Gambar II. 1 Pembobotan TF-IDF .....                                    | 7  |
| Gambar II. 2 Support Vector Machine .....                               | 10 |
| Gambar III. 1 Kerangka Berpikir.....                                    | 15 |
| Gambar III. 2 Pengumpulan Data.....                                     | 16 |
| Gambar III. 3 Flowchart Pengumpulan Data .....                          | 17 |
| Gambar III. 4 Flowchart cleansing .....                                 | 18 |
| Gambar III. 5 Flowchart transform.....                                  | 19 |
| Gambar III. 6 Flowchart tokenizing .....                                | 20 |
| Gambar III. 7 Flowchart filter stopword.....                            | 21 |
| Gambar III. 8 Flowchart stem .....                                      | 23 |
| Gambar III. 9 Filter Token .....                                        | 24 |
| Gambar III. 10 Hasil Akurasi .....                                      | 26 |
| Gambar III. 11 Proses Crawling Sebagai Dataset .....                    | 33 |
| Gambar III. 12 Proses Cleansing Data.....                               | 33 |
| Gambar III. 13 Proses Pre-processing .....                              | 34 |
| Gambar III. 14 Model Pembobotan TF-IDF dan K-Fold Cross Validation..... | 34 |
| Gambar III. 15 Model Klasifikasi NB .....                               | 35 |
| Gambar III. 16 Model Klasifikasi SVM .....                              | 35 |
| Gambar IV. 14 Tampilan Menu Profil Pengguna.....                        | 37 |
| Gambar IV. 15 Tampilan Halaman Beranda .....                            | 38 |
| Gambar IV. 16 Tampilan Menu Pembayaran .....                            | 39 |
| Gambar IV. 17 Tampilan Menu Produk .....                                | 40 |



## DAFTAR PUSTAKA

- Administrator. (2019). Online Shop, Market Place, E-Commerce. Apa bedanya?. Berita Kedinasan. Diakses pada Rabu 8 Februari 2023 dari <https://diskominfo.kedirikab.go.id/baca/online-shop-market-place-ecommerce-apa-bedanya>.
- Auditya Capriani, V. (2016). Strategi Komunikasi Pemasaran E- Commerce dengan sistem Marketplace *Shopee* Dalam Meningkatkan Jumlah Konsumen. Diakses pada Rabu 8 Februari 2023 dari <http://eprints.upnyk.ac.id/10673>.
- Admin2Lpm. (2022). Analisis Sentimen (Sentiment Analysis) : Definisi, Tipe dan Cara Kerjanya. LP2M. Diakses pada Rabu 8 Februari 2023 dari <https://lp2m.uma.ac.id/2022/02/21/analisis-sentimen-sentiment-analysis-definisi-tipe-dan-cara-kerjanya>
- Mona Cindo, D. P. (2019). Literatur Review : Metode Klasifikasi Pada Sentimen Analisis . *SAINTEKS* , 66-70.
- S. Gusriani, K. D. K. Wardhani, and M. I. Zul, “Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naïve Bayes (Studi Kasus: Facebook Page BerryBenka),” 4th Appl. Bus. Eng. Conf., vol. 1, no. 1, pp. 1–7, 2016.
- “Analisis Sentimen Menggunakan *Support Vector Machine* dan Maximum Entropy Sentiment Analysis Using *Support Vector Machine* and Maximum Entropy Method,” vol. 4, no. 2, pp. 2389–2395, 2017.
- Universitas Esa Unggul.( 2023). Pengertian *Machine Learning*, Apa itu?. Diakses pada Selasa 14 Februari 2023 dari <https://fasilkom.esaunggul.ac.id/pengertian-machine-learning-apa-itu>.
- Rita Apriani, D. G. (2019). Analisis sentimen dengan *naïve bayes* terhadap komentar aplikasi *Shopee*. *Rekayasa Teknologi Nusa Putra*, Hal : 54-62.
- Anggreini, N. M., Nasir, B. M., & Noor, I. L. (2016). Pemanfaatan Media Sosial *Twitter* di Kalangan Pelajar SMK Negeri 5 Samarinda. *Sosiologi*, 4(2), 239- 51.

- Lutfia Afifah. (Tanpa tahun) . Apa itu Confusion Matrix di Machine Learning?. I lmutatapy. Diakses pada Selasa 14 Februari 2023 dari <https://ilmudatapy.com/apa-itu-confusion-matrix>.
- Suyanto. 2017. Data mining Untuk Klasifikasi Dan Klasterisasi Data. Bandung: Informatika Bandung.
- Mukminin, Amirul, Dwiza Riana. 2017. Komparasi Algoritma C4.5, Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah, ISSN: 2355-6579 E-ISSN: 2528-2247. Bandung : JURNAL INFORMATIKA, Vol.4 No.1 April 2017, pp. 21-31
- Nengah W. U, Joshael J.S. (2021). USING SVM (SUPPORT VECTOR MACHINE) METHOD TO PREDICT BEST SELLER BOOK IN AMAZON, eISSN 2541-0679. SMART TECHNO, Vol. 3 No. 1, Februari 2021, hlm. 1–4,
- Samsudiney. (19). Penjelasan Sederhana tentang Apa Itu SVM?. Diakses pada Selasa 14 Februari 2023 dari <https://medium.com/@samsudiney/penjelasan-sederhana-tentang-apa-itu-svm-149fec72bd02>
- Rony Setiawan. (2021). Flowchart Adalah: Fungsi, Jenis, Simbol, dan Contohnya. Diakses pada Selasa 14 Februari 2023 dari <https://www.dicoding.com/blog/flowchart-adalah>.
- Advernesia. (2020). Pengertian Microsoft Excel dan Fungsinya. Diakses pada Selasa 14 Februari 2023 dari <https://www.advernesia.com/blog/microsoft-excel/microsoft-excel-adalah>.
- Dian Ardiansyah, W. W. (2018). ALGORITMA C4.5 UNTUK KLASIFIKASI CALON PESERTA LOMBA CERDAS CERMAT SISWA SMP DENGAN MENGGUNAKAN APLIKASI RAPIDMINER. *Inkofar*, 2581-2920.