

**PERBANDINGAN KINERJA ALGORITMA MACHINE  
LEARNING DALAM KLASIFIKASI JAMUR BERACUN**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**ARIC YOHANES**

**21.11.4381**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**PERBANDINGAN KINERJA ALGORITMA MACHINE  
LEARNING DALAM KLASIFIKASI JAMUR BERACUN**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh  
**ARIC YOHANES**  
**21.11.4381**

Kepada

**FAKULTAS ILMU KOMPUTER**  
**UNIVERSITAS AMIKOM YOGYAKARTA**  
**YOGYAKARTA**  
**2025**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**PERBANDINGAN KINERJA ALGORITMA MACHINE LEARNING  
DALAM KLASIFIKASI JAMUR BERACUN**

yang disusun dan diajukan oleh

**ARIC YOHANES**

**21.11.4381**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 17 April 2025

Dosen Pembimbing,



**Ferian Fauzi Abdullah, S.Kom., M.Kom.**  
**NIK. 190302276**

**HALAMAN PENGESAHAN**  
**SKRIPSI**  
**PERBANDINGAN KINERJA ALGORITMA MACHINE LEARNING**  
**DALAM KLASIFIKASI JAMUR BERACUN**

yang disusun dan diajukan oleh

**ARIC YOHANES**

**21.11.4381**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 17 April 2025

**Susunan Dewan Penguji**

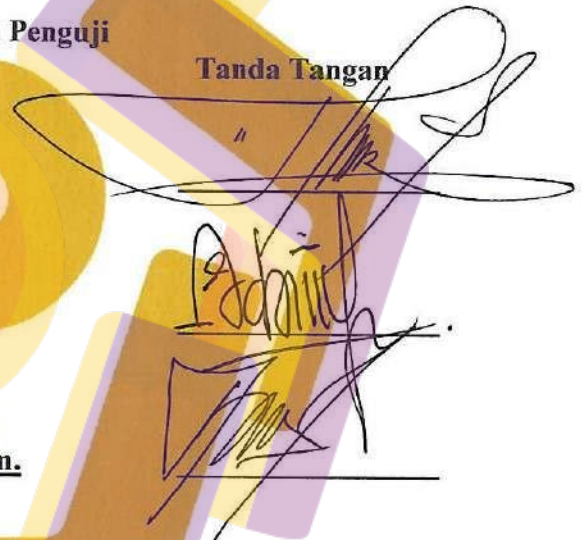
**Nama Penguji**

**Haryoko, S.Kom., M.Cs.**  
**NIK. 190302286**

**Irma Rofni Wulandari, S.Pd., M.Eng.**  
**NIK. 190302329**

**Ferian Fauzi Abdulloh, S.Kom., M.Kom.**  
**NIK. 190302276**

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 17 April 2025

**DEKAN FAKULTAS ILMU KOMPUTER**



**Prof. Dr. Kusrini, M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : **ARIC YOHANES**  
NIM : **21.11.4381**

Menyatakan bahwa Skripsi dengan judul berikut:

### **PERBANDINGAN KINERJA ALGORITMA MACHINE LEARNING DALAM KLASIFIKASI JAMUR BERACUN**

Dosen Pembimbing : Ferian Fauzi Abdulloh, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 17 April 2025

Yang Menyatakan,

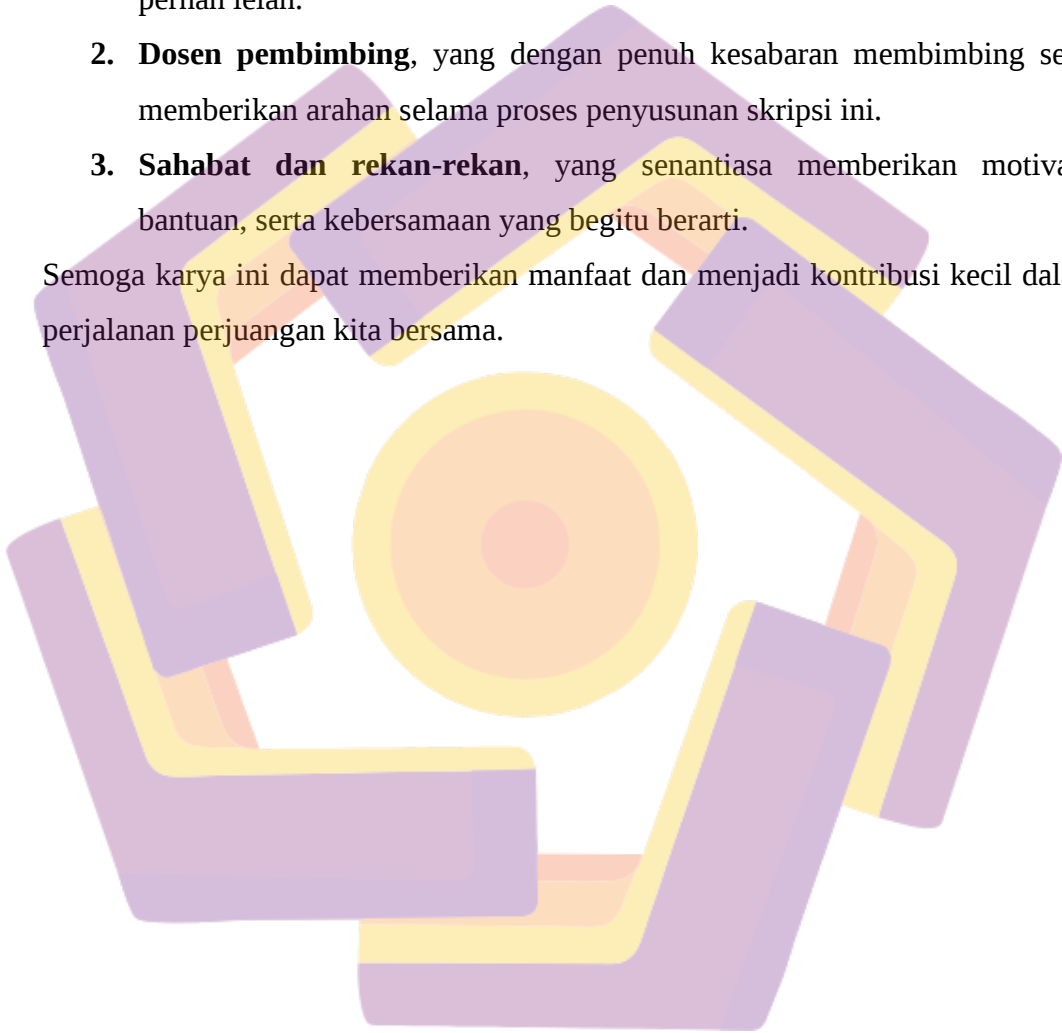
  
  
**ARIC YOHANES**

## HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan dengan rasa syukur dan penghargaan yang mendalam kepada:

1. **Orang tua saya**, yang selalu memberikan cinta, dukungan, dan doa tanpa pernah lelah.
2. **Dosen pembimbing**, yang dengan penuh kesabaran membimbing serta memberikan arahan selama proses penyusunan skripsi ini.
3. **Sahabat dan rekan-rekan**, yang senantiasa memberikan motivasi, bantuan, serta kebersamaan yang begitu berarti.

Semoga karya ini dapat memberikan manfaat dan menjadi kontribusi kecil dalam perjalanan perjuangan kita bersama.



## KATA PENGANTAR

Segala puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan skripsi ini yang berjudul “PERBANDINGAN KINERJA ALGORITMA MACHINE LEARNING DALAM KLASIFIKASI JAMUR BERACUN” sebagai bagian dari pemenuhan syarat untuk meraih gelar S.Kom pada program studi Informatika, Universitas Amikom Yogyakarta.

Pada kesempatan ini, saya ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. **Prof. Dr. M. Suyanto, M.M.**, selaku Rektor Universitas Amikom Yogyakarta, atas segala dukungan dan fasilitas yang telah diberikan selama masa studi.
2. **Bapak Ferian Fauzi Abdulloh, S.Kom., M.Kom.**, yang dengan penuh kesabaran memberikan bimbingan, arahan, serta dukungan selama proses penyusunan skripsi ini.
3. **Keluarga tercinta**, yang senantiasa mendukung dengan doa dan semangat tanpa henti.
4. **Semua pihak yang telah berkontribusi**, baik secara langsung maupun tidak langsung.

Saya menyadari bahwa skripsi ini masih memiliki keterbatasan, sehingga dengan rendah hati saya mengharapkan kritik dan saran yang membangun untuk penyempurnaan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan dan masyarakat luas.

Yogyakarta, 17 April 2025

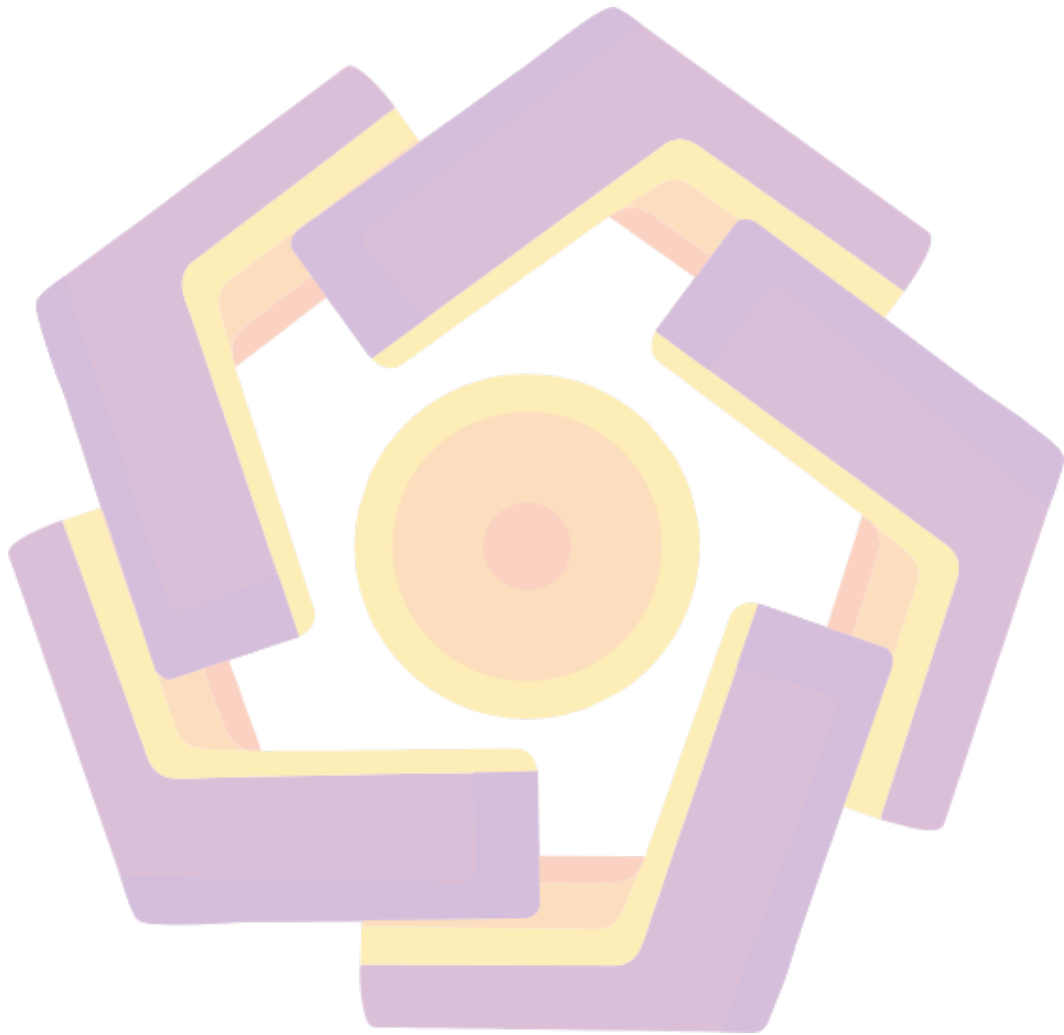
Penulis

## DAFTAR ISI

|                                           |      |
|-------------------------------------------|------|
| HALAMAN JUDUL .....                       | i    |
| HALAMAN PERSETUJUAN.....                  | ii   |
| HALAMAN PENGESAHAN .....                  | iii  |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI ..... | iii  |
| HALAMAN PERSEMBAHAN .....                 | v    |
| KATA PENGANTAR .....                      | vi   |
| DAFTAR ISI.....                           | vii  |
| DAFTAR TABEL .....                        | x    |
| DAFTAR GAMBAR .....                       | xi   |
| DAFTAR LAMPIRAN.....                      | xii  |
| DAFTAR LAMBANG DAN SINGKATAN .....        | xiii |
| DAFTAR ISTILAH .....                      | xiv  |
| INTISARI .....                            | xv   |
| ABSTRACT.....                             | xvi  |
| BAB I PENDAHULUAN .....                   | 1    |
| 1.1 Latar Belakang .....                  | 1    |
| 1.2 Rumusan Masalah .....                 | 3    |
| 1.3 Batasan Masalah.....                  | 3    |
| 1.4 Tujuan Penelitian.....                | 4    |
| 1.5 Manfaat Penelitian.....               | 4    |
| 1.6 Sistematika Penulisan.....            | 5    |
| BAB II TINJAUAN PUSTAKA .....             | 6    |
| 2.1 Studi Literatur .....                 | 6    |

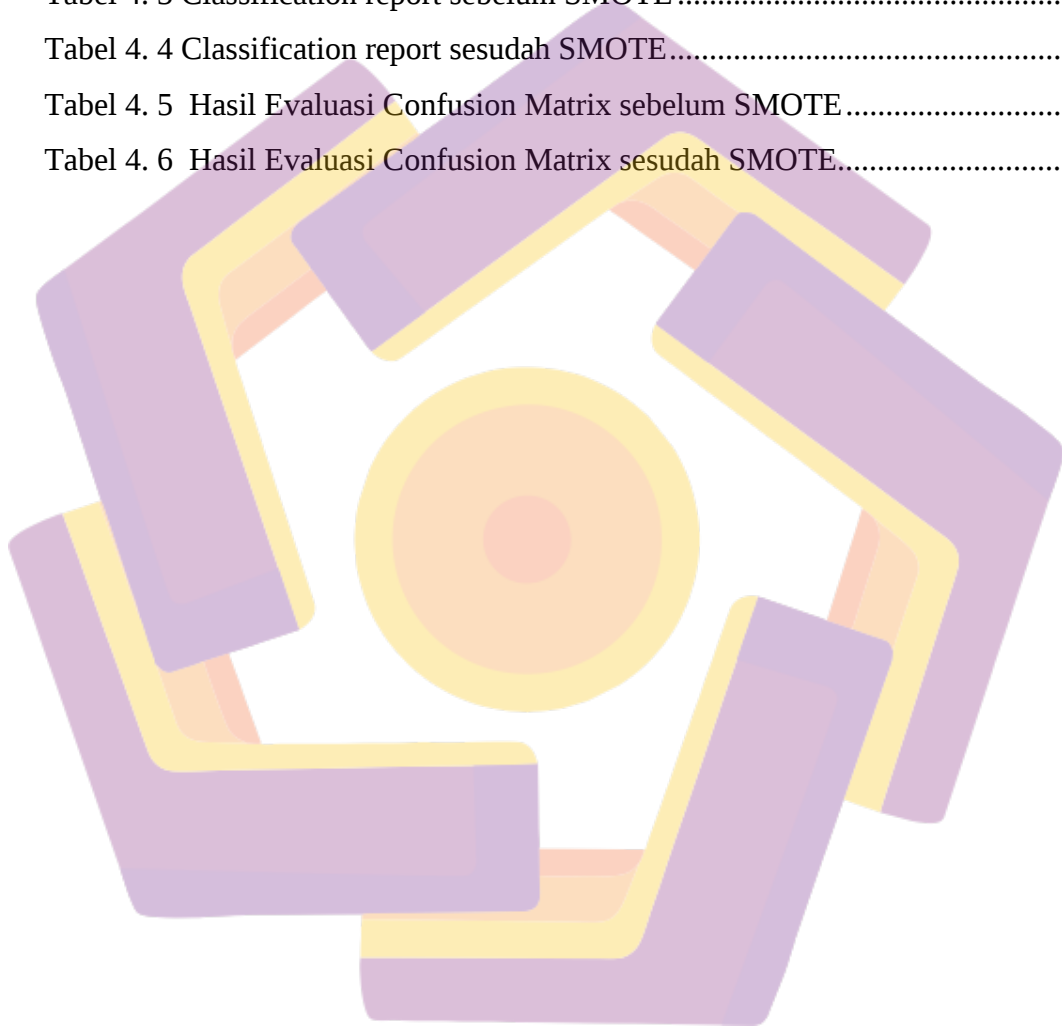
|                                   |                                        |    |
|-----------------------------------|----------------------------------------|----|
| 2.2                               | Dasar Teori .....                      | 11 |
| 2.2.1.                            | <i>Python</i> .....                    | 11 |
| 2.2.2.                            | <i>Google Colab</i> .....              | 11 |
| 2.2.3.                            | <i>Machine Learning</i> .....          | 11 |
| 2.2.4.                            | Jamur .....                            | 12 |
| 2.2.5.                            | <i>Kaggle</i> .....                    | 12 |
| 2.2.6.                            | <i>Preprocessing</i> .....             | 13 |
| 2.2.7.                            | <i>Naïve Bayes</i> .....               | 13 |
| 2.2.8.                            | <i>Logistic Regression</i> .....       | 14 |
| 2.2.9.                            | <i>Support Vector Machine</i> .....    | 15 |
| 2.2.10.                           | Model Evaluasi .....                   | 16 |
| BAB III METODE PENELITIAN .....   |                                        | 19 |
| 3.1                               | Objek Penelitian .....                 | 19 |
| 3.2                               | Alur Penelitian .....                  | 19 |
| 3.3                               | Alat dan Bahan .....                   | 24 |
| BAB IV HASIL DAN PEMBAHASAN ..... |                                        | 26 |
| 4.1.                              | Pengumpulan Data .....                 | 26 |
| 4.2.                              | <i>Exploratory Data Analysis</i> ..... | 26 |
| 4.3.                              | <i>Preprocessing</i> .....             | 28 |
| 4.4.                              | <i>Modeling</i> .....                  | 32 |
| 4.5.                              | Evaluasi Model .....                   | 33 |
| BAB V PENUTUP .....               |                                        | 38 |
| 5.1                               | Kesimpulan .....                       | 38 |
| 5.2                               | Saran .....                            | 39 |
| REFERENSI .....                   |                                        | 40 |

|               |    |
|---------------|----|
| LAMPIRAN..... | 46 |
|---------------|----|



## DAFTAR TABEL

|                                                                   |    |
|-------------------------------------------------------------------|----|
| Tabel 2. 1 Keaslian Penelitian .....                              | 9  |
| Tabel 4. 1 Pembagian Data Train dan Data Test sebelum SMOTE ..... | 33 |
| Tabel 4. 2 Pembagian Data Train dan Data Test sesudah SMOTE ..... | 33 |
| Tabel 4. 3 Classification report sebelum SMOTE .....              | 34 |
| Tabel 4. 4 Classification report sesudah SMOTE.....               | 34 |
| Tabel 4. 5 Hasil Evaluasi Confusion Matrix sebelum SMOTE.....     | 35 |
| Tabel 4. 6 Hasil Evaluasi Confusion Matrix sesudah SMOTE.....     | 35 |

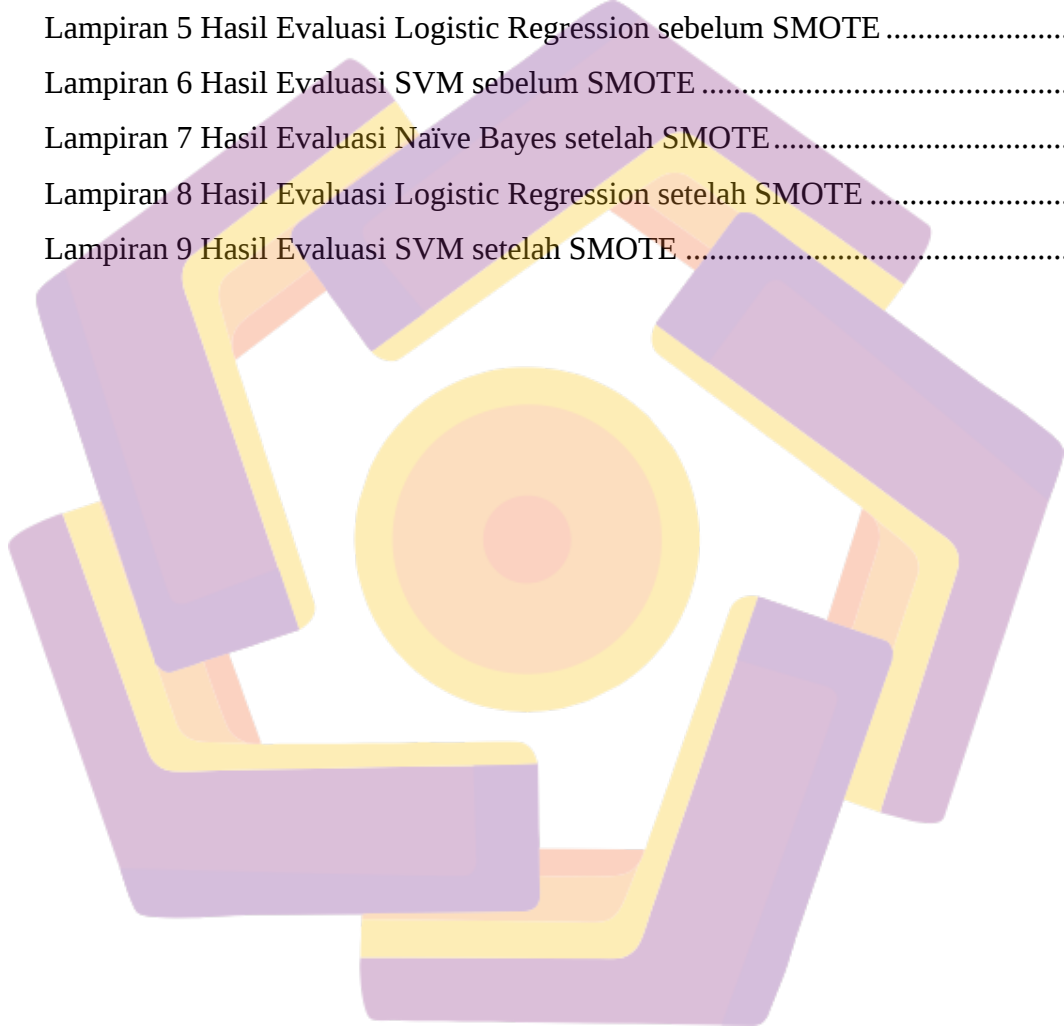


## DAFTAR GAMBAR

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| Gambar 3. 1 Alur Penelitian .....                                         | 20 |
| Gambar 4. 1 Dataset Mushroom Classification .....                         | 26 |
| Gambar 4. 2 Informasi Data.....                                           | 27 |
| Gambar 4. 3 menghitung jumlah ‘?’ pada fitur “stalk-root” .....           | 27 |
| Gambar 4. 4 Distribusi data Class.....                                    | 28 |
| Gambar 4. 5 jumlah data ‘?’ sebelum diganti dengan modus .....            | 29 |
| Gambar 4. 6 jumlah data ‘?’ setelah diganti dengan modus .....            | 29 |
| Gambar 4. 7 data setelah dilakukan label encoding .....                   | 29 |
| Gambar 4. 8 10 fitur terbaik berdasarkan MI Score .....                   | 30 |
| Gambar 4. 9 hasil dari normalisasi data .....                             | 30 |
| Gambar 4. 10 distribusi data kelas sebelum dilakukan balancing data.....  | 31 |
| Gambar 4. 11 distribusi data kelas setelah dilakukan balancing data ..... | 32 |
| Gambar 4. 12 Perbandingan accuracy dari semua model sebelum SMOTE .....   | 36 |
| Gambar 4. 13 Perbandingan accuracy dari semua model sesudah SMOTE .....   | 37 |

## DAFTAR LAMPIRAN

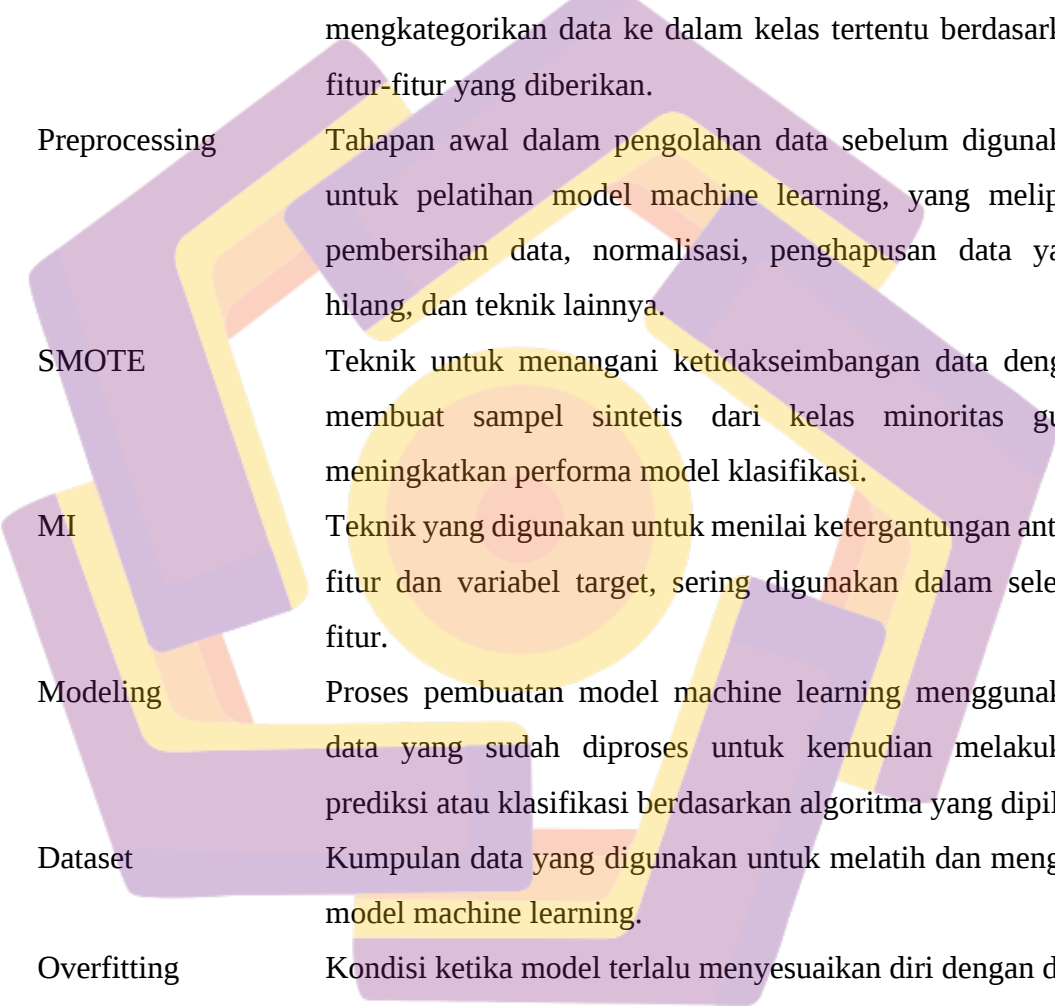
|                                                                   |    |
|-------------------------------------------------------------------|----|
| Lampiran 1 Tampilan website Kaggle .....                          | 46 |
| Lampiran 2 Tampilan website UCI Machine Learning Repository.....  | 46 |
| Lampiran 3 Tampilan Dataset Klasifikasi Jamur.....                | 47 |
| Lampiran 4 Hasil Evaluasi Naïve Bayes sebelum SMOTE.....          | 47 |
| Lampiran 5 Hasil Evaluasi Logistic Regression sebelum SMOTE ..... | 48 |
| Lampiran 6 Hasil Evaluasi SVM sebelum SMOTE .....                 | 48 |
| Lampiran 7 Hasil Evaluasi Naïve Bayes setelah SMOTE.....          | 49 |
| Lampiran 8 Hasil Evaluasi Logistic Regression setelah SMOTE ..... | 49 |
| Lampiran 9 Hasil Evaluasi SVM setelah SMOTE .....                 | 50 |



## DAFTAR LAMBANG DAN SINGKATAN

|       |                                            |
|-------|--------------------------------------------|
| AI    | Artificial Intelligence                    |
| ML    | Machine Learning                           |
| NB    | Naïve Bayes                                |
| LR    | Logistic Regression                        |
| SVM   | Support Vector Machine                     |
| SMOTE | Synthetic Minority Over-sampling Technique |
| MI    | Mutual Information                         |
| TP    | True Positive                              |
| TN    | True Negative                              |
| FP    | False Positive                             |
| FN    | False Negative                             |
| EDA   | Exploratory Data Analysis                  |
| CSV   | Comma-Separated Values                     |

## DAFTAR ISTILAH



|                  |                                                                                                                                                                                             |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Machine Learning | Cabang dari kecerdasan buatan (AI) yang memungkinkan komputer belajar dari data untuk membuat prediksi atau keputusan tanpa diprogram secara eksplisit.                                     |
| Klasifikasi      | Metode dalam machine learning yang bertujuan untuk mengkategorikan data ke dalam kelas tertentu berdasarkan fitur-fitur yang diberikan.                                                     |
| Preprocessing    | Tahapan awal dalam pengolahan data sebelum digunakan untuk pelatihan model machine learning, yang meliputi pembersihan data, normalisasi, penghapusan data yang hilang, dan teknik lainnya. |
| SMOTE            | Teknik untuk menangani ketidakseimbangan data dengan membuat sampel sintetis dari kelas minoritas guna meningkatkan performa model klasifikasi.                                             |
| MI               | Teknik yang digunakan untuk menilai ketergantungan antara fitur dan variabel target, sering digunakan dalam seleksi fitur.                                                                  |
| Modeling         | Proses pembuatan model machine learning menggunakan data yang sudah diproses untuk kemudian melakukan prediksi atau klasifikasi berdasarkan algoritma yang dipilih.                         |
| Dataset          | Kumpulan data yang digunakan untuk melatih dan menguji model machine learning.                                                                                                              |
| Overfitting      | Kondisi ketika model terlalu menyesuaikan diri dengan data latih sehingga performanya menurun pada data baru.                                                                               |
| Fitur Diagnostik | Fitur yang memiliki pengaruh signifikan dalam membedakan kelas dalam suatu klasifikasi, misalnya tipe cincin pada jamur.                                                                    |

## INTISARI

Kesalahan dalam mengidentifikasi jamur beracun dapat menyebabkan dampak serius bagi kesehatan manusia. Oleh karena itu, penelitian ini membandingkan kinerja algoritma Naïve Bayes, Logistic Regression, dan Support Vector Machine (SVM) dalam klasifikasi jamur beracun dan tidak beracun. Data mengalami prapemrosesan, termasuk penanganan missing value, konversi data kategorikal dengan Label Encoder, serta normalisasi menggunakan MinMaxScaler. Untuk mengatasi ketidakseimbangan data, diterapkan metode SMOTE, sementara Mutual Information digunakan dalam seleksi fitur guna meningkatkan relevansi atribut yang diproses oleh model. Evaluasi menggunakan confusion matrix menunjukkan bahwa SVM memiliki akurasi terbaik sebesar 97%, diikuti oleh Logistic Regression dengan 91%, dan Naïve Bayes dengan 89% sebelum SMOTE serta 90% setelahnya. Hasil ini menunjukkan bahwa SVM merupakan algoritma paling efektif untuk klasifikasi jamur beracun dibandingkan algoritma lainnya. Penelitian ini dapat dimanfaatkan dalam pengembangan sistem klasifikasi otomatis guna meningkatkan keamanan konsumsi jamur dan membantu industri pangan dalam mengurangi risiko kesalahan identifikasi jamur yang berbahaya bagi kesehatan manusia. Selain itu, hasil penelitian ini dapat menjadi dasar bagi penelitian selanjutnya yang berfokus pada peningkatan akurasi model dengan teknik tuning hyperparameter atau penerapan model ensemble.

**Kata kunci:** Klasifikasi, Jamur Beracun, Naïve Bayes, Logistic Regression, Support Vector Machine, SMOTE.

## ***ABSTRACT***

Errors in identifying poisonous mushrooms can have serious consequences for human health. Therefore, this study compares the performance of the Naïve Bayes, Logistic Regression, and Support Vector Machine (SVM) algorithms in classifying poisonous and non-poisonous mushrooms. The data underwent preprocessing, including handling missing values, converting categorical data using Label Encoding, and normalization using MinMaxScaler. To address data imbalance, the SMOTE method was applied, while Mutual Information was used for feature selection to enhance the relevance of attributes processed by the model. Evaluation using the confusion matrix showed that SVM achieved the highest accuracy of 97%, followed by Logistic Regression with 91%, and Naïve Bayes with 89% before SMOTE and 90% after applying SMOTE. These results indicate that SVM is the most effective algorithm for classifying poisonous mushrooms compared to the others. This research can be utilized in developing automated classification systems to improve mushroom consumption safety and help the food industry reduce the risk of misidentifying toxic mushrooms. Additionally, these findings can serve as a foundation for future research focused on improving model accuracy through hyperparameter tuning techniques or the application of ensemble models.

**Keyword:** *Classification, Poisonous Mushrooms, Naïve Bayes, Logistic Regression, Support Vector Machine, SMOTE.*