

**ANALYSIS OF STACKING ENSEMBLE METHOD IN  
MACHINE LEARNING ALGORITHMS TO  
PREDICT STUDENT DEPRESSION**

**LAPORAN NON-REGULER**

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



Disusun oleh :

**NAOUTHLAASIA LEVINA**

**22.11.4585**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**ANALYSIS OF STACKING ENSEMBLE METHOD IN MACHINE  
LEARNING ALGORITHMS TO PREDICT STUDENT  
DEPRESSION**

**LAPORAN NON-REGULER**

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



Disusun oleh :

**NAOUTHLA ASIA LEVINA**

**22.11.4585**

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

**HALAMAN PERSETUJUAN**

**JALUR NON-REGULER**

**ANALYSIS OF STACKING ENSEMBLE METHOD IN MACHINE LEARNING  
ALGORITHMS TO PREDICT STUDENT DEPRESSION**

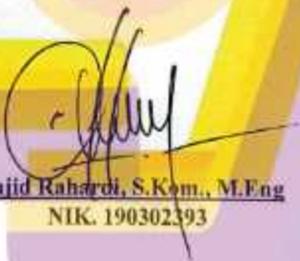
yang disusun dan diajukan oleh

**Naouthla Asia Levina**

**22.11.4585**

telah disetujui oleh Dosen Pembimbing  
pada tanggal 15 Desember 2025

**Dosen Pembimbing,**



**Majid Rahardi, S.Kom., M.Eng**  
**NIK. 190302393**

**HALAMAN PENGESAHAN**

**JALUR NON-REGULER**

**ANALYSIS OF STACKING ENSEMBLE METHOD IN MACHINE LEARNING  
ALGORITHMS TO PREDICT STUDENT DEPRESSION**

yang disusun dan diajukan oleh

**Naouthla Asia Levina**  
**22.11.4585**

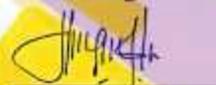
Telah dipertahankan di depan Dewan Penguji  
pada tanggal 15 Desember 2025

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Dr. Hartatik, S.T., M.Cs.**  
**NIK. 190302232**



**Ike Verawati, S.Kom., M.Kom.**  
**NIK. 190302237**



**Majid Rahardi, S.Kom., M.Eng.**  
**NIK. 190302393**



Laporan ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 15 Desember 2025

**DEKAN FAKULTAS ILMU KOMPUTER**



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

## HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Naouthla Asia Levina

NIM : 22.11.4585

Menyatakan bahwa Laporan dengan judul berikut:

### **Analysis of Stacking Ensemble Method in Machine Learning Algorithms to Predict Student Depression**

Dosen Pembimbing : Majid Rahardi, S.Kom., M.Eng

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 kegiatan 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 ketidak-benaran 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, 15 Desember 2025

Yang Menyatakan,



The image shows an official stamp of Universitas AMIKOM Yogyakarta. The stamp includes the university's logo, the text 'UNIVERSITAS AMIKOM YOGYAKARTA', and 'MEPERNI TEMPI'. Below the stamp is a handwritten signature in black ink.

Naouthla Asia Levina

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga laporan sidang ini dapat diselesaikan dengan baik. Laporan ini disusun sebagai salah satu syarat sidang akhir sekaligus bentuk pertanggungjawaban atas kegiatan penelitian yang telah dilakukan. Ucapan terima kasih penulis sampaikan kepada pihak-pihak yang telah memberikan dukungan, bantuan, dan doa selama proses penyusunan laporan ini, yaitu:

1. Orang tua dan adik-adik tercinta atas doa dan dukungan yang tiada henti.
2. Prof. Dr. Mohammad Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
3. Prof. Dr. Kusrini, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Eli Pujastuti, M.Kom, selaku Ketua Program Studi Informatika, Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
5. Majid Rahardi, S.Kom, M.Eng, selaku Dosen Pembimbing yang telah memberikan bimbingan dan arahan selama penyusunan laporan.
6. Tim Dosen Penguji, atas masukan kritis dan saran konstruktif yang memperdalam pemahaman penulis terhadap penelitian ini.
7. Teman-teman seperjuangan, khususnya TPA Sragen Pusat, COS, dan BWB yang selalu memberikan semangat dan dukungan.

Penulis menyadari bahwa laporan ini masih jauh dari sempurna. Semoga laporan ini dapat menjadi sarana pembelajaran dan menumbuhkan semangat untuk terus berkembang dalam bidang ilmu pengetahuan.

Yogyakarta, 12 Desember 2025

  
Naouthia Asia Levina

## DAFTAR ISI

Halaman Judul.....	i
Halaman Persetujuan.....	ii
Halaman Pengesahan.....	iii
Halaman Pernyataan Keaslian Karya.....	iv
Kata Pengantar.....	v
Daftar Isi.....	vi
Daftar Tabel.....	viii
Daftar Gambar.....	ix
Daftar Lampiran.....	x
Daftar Lambang dan Singkatan.....	xi
Daftar Istilah.....	xii
Intisari.....	xiv
<i>Abstract</i> .....	xv
<b>Bab I</b> Pendahuluan.....	1
1.1. Gambaran Umum.....	1
1.2. Rumusan Masalah.....	1
1.3. Batasan Masalah.....	2
1.4. Tujuan.....	2
<b>Bab II</b> Tinjauan Pustaka.....	4
2.1. Studi Literatur.....	4
2.2. Landasan Teori.....	5
<b>BAB III</b> Metode Penelitian.....	11
3.1. Jenis Penelitian.....	11
3.2. Objek Penelitian.....	11
3.3. Alur Penelitian.....	13
3.4. Implementasi Teknik.....	14
3.5. Sumber Data.....	15
3.6. Pre-processing Data.....	15
3.7. Split Data.....	16
3.8. Komponen Stacking.....	17

3.9. Base Learners (level-0) .....	17
3.10. Meta Learner (level-1) .....	18
3.11. Mekanisme Cross-Validation Stacking .....	18
3.12. Evaluasi Model .....	19
3.13. Analisis Overfitting .....	19
3.14. Analisis Feature Importance (khusus Random Forest) .....	20
BAB IV Pembahasan .....	21
4.1. Hasil Pre-processing Data .....	21
4.2. Hasil Split Data .....	24
4.3. Hasil Evaluasi Base Learners (Random Forest) .....	25
4.4. Hasil Evaluasi Base Learners (Gradient Boosting) .....	27
4.5. Hasil Evaluasi Base Learners (XGBoost) .....	29
4.6. Hasil Evaluasi Model Stacking Ensemble .....	31
4.7. Perbandingan Kinerja Model Stacking Ensemble dengan Base Learners .....	33
4.8. Perbandingan Kinerja Model Base Learners, Stacking Ensemble, dan Bagging .....	35
4.9. Analisis Overfitting .....	36
4.10. Analisis Feature Importance .....	37
BAB V Kesimpulan .....	39
5.1. Kesimpulan .....	39
5.2. Saran .....	40
Referensi .....	41
Lampiran .....	45
Curriculum Vitae .....	45
Lampiran dan Bukti Pendukung .....	46
a. Letter of Acceptance (LOA) .....	46
b. Lembar Review .....	47
c. Bukti Terbit/Terindex .....	51
d. Bukti pembayaran .....	52

## DAFTAR TABEL

Tabel 3.1 Deskripsi Dataset.....	11
Tabel 4.1 Hasil sebagian label encoding.....	23
Tabel 4.2 Hasil pembagian data.....	24
Tabel 4.3 Hasil Evaluasi Random Forest.....	26
Tabel 4.4 Hasil Evaluasi Gradient Boosting.....	28
Tabel 4.5 Hasil Evaluasi XGBoost.....	30
Tabel 4.6 Hasil Perbandingan Performa SE dengan Base Learners.....	34
Tabel 4.7 Hasil Perbandingan Performa Base Learners, SE,.....	35
Tabel 4.8 Hasil Perbandingan Performa Training dan Testing.....	37

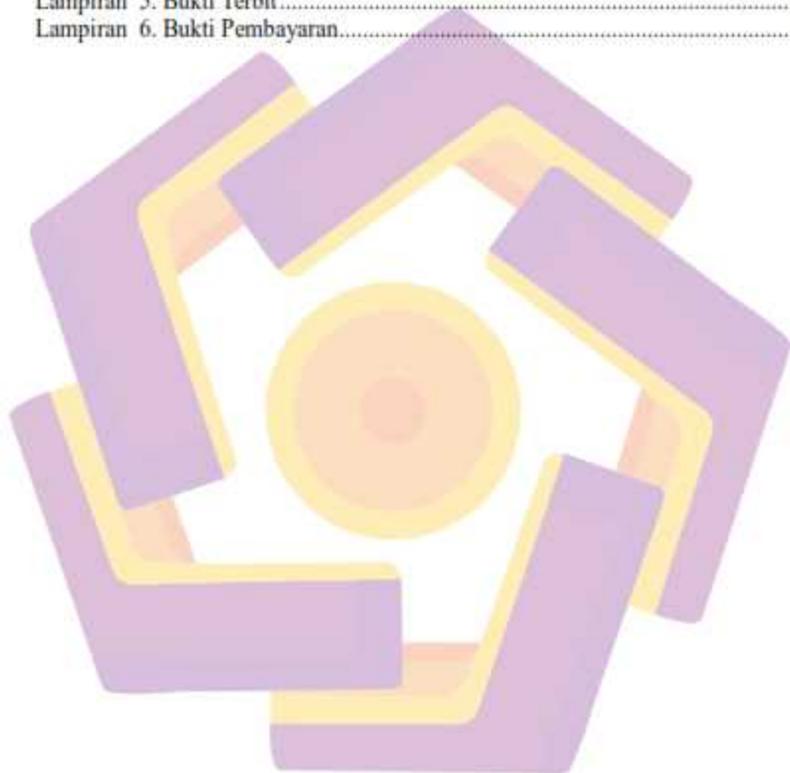


## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	13
Gambar 4.1 Hasil Missing Value .....	21
Gambar 4.2 Hasil Encoding kolom kategorikal.....	22
Gambar 4.3 Hasil normalisasi fitur numerik.....	23
Gambar 4.4 Distribusi kelas target 'Depression'.....	25
Gambar 4.5 Confusion Matrix Random Forest.....	25
Gambar 4.6 Kurva ROC Random Forest.....	26
Gambar 4.7 Kurva Precision-Recall Random Forest.....	27
Gambar 4.8 Confusion Matrix Gradient Boosting.....	28
Gambar 4.9 Kurva ROC AUC Gradient Boosting.....	29
Gambar 4.10 Kurva Precision-Recall Gradient Boosting.....	29
Gambar 4.11 Confusion Matrix XGBoost.....	30
Gambar 4.12 Kurva ROC XGBoost.....	31
Gambar 4.13 Kurva Precision-Recall XGBoost.....	31
Gambar 4.14 Confusion Matrix Model Stacking Ensemble.....	32
Gambar 4.15 Hasil Evaluasi Stacking Ensemble.....	33
Gambar 4.16 Kurva ROC.....	33
Gambar 4.17 Kurva Precision-Recall.....	33
Gambar 4.18 Perbandingan Akurasi SE dengan Base Learners.....	35
Gambar 4.19 Hasil Perbandingan Akurasi Base Learners, SE, dan BC.....	36
Gambar 4.20 Feature Importance.....	38

## DAFTAR LAMPIRAN

Lampiran 1. Curriculum Vitae .....	45
Lampiran 2. Letter of Acceptance (LOA).....	46
Lampiran 3. Lembar Review.....	47
Lampiran 4. Bukti Terindex .....	51
Lampiran 5. Bukti Terbit.....	52
Lampiran 6. Bukti Pembayaran.....	52



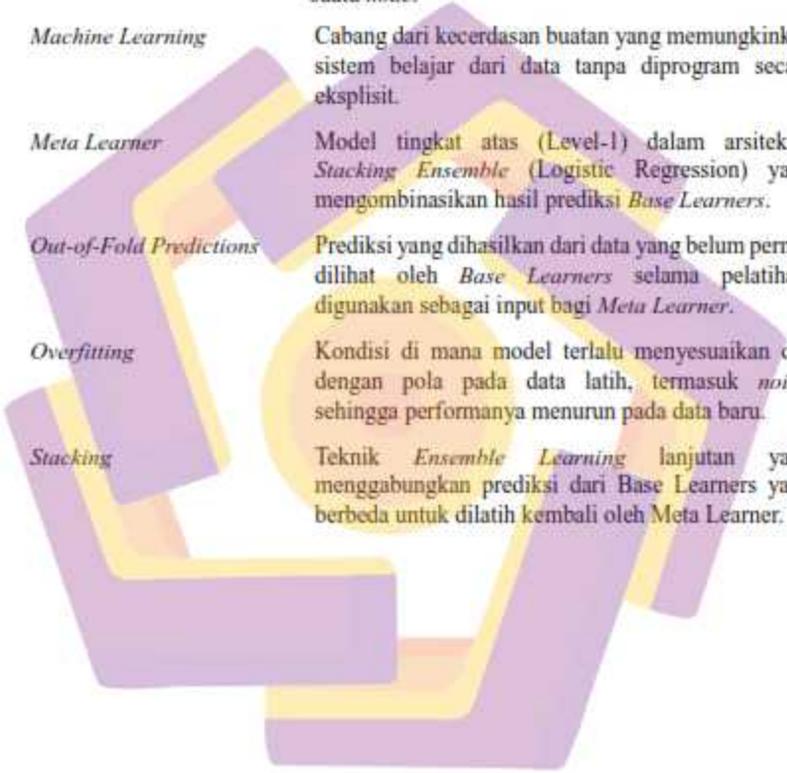
## DAFTAR LAMBANG DAN SINGKATAN

ML	Machine Learning
RF	Random Forest
GB	Gradient Boosting
XGB	XGBoost
BC	Bagging Classifier
CVS	Cross-Validation Stacking
OOF	Out-of-Fold
AP	Average Precision
DSM-5	Diagnostic and Statistical Manual of Mental Disorders
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve



## DAFTAR ISTILAH

<i>Bagging</i>	Teknik <i>Ensemble Learning</i> yang bertujuan mengurangi varian model dengan melatih beberapa model pada subset data yang diambil secara acak, lalu menggabungkan hasilnya melalui <i>voting</i> mayoritas.
<i>Base Learners</i>	Model tingkat pertama (Level-0) dalam arsitektur <i>Stacking Ensemble</i> (RF, GB, XGBoost) yang dilatih secara independen
<i>Boosting</i>	Teknik <i>Ensemble Learning</i> yang bertujuan mengurangi bias dengan melatih model secara berurutan, di mana setiap model baru berupaya memperbaiki kesalahan prediksi model sebelumnya.
<i>Data Leakage</i>	Kondisi ketika <i>Meta Learner</i> menerima informasi dari sampel yang telah digunakan oleh <i>Base Learners</i> saat pelatihan, yang menyebabkan peningkatan performa secara tidak valid.
<i>Deep Learning</i>	Sub-bidang dari <i>Machine Learning</i> yang menggunakan arsitektur <i>Neural Network</i> dengan banyak lapisan ( <i>deep layers</i> ).
<i>Depresi</i>	Gangguan suasana hati ( <i>mood disorder</i> ) yang ditandai dengan perasaan sedih berkepanjangan, kehilangan minat, perubahan pola tidur, dan penurunan energi.
<i>Ensemble Learning</i>	Pendekatan ML yang menggabungkan beberapa model prediksi untuk kinerja yang lebih akurat dan stabil.
<i>Explainable AI</i>	Disiplin ilmu yang berfokus pada pengembangan metode untuk membuat model <i>Machine Learning</i> lebih transparan dan mudah dipahami oleh manusia.



<i>Feature Importance</i>	Analisis kuantitatif yang mengukur kontribusi setiap fitur terhadap hasil prediksi model, berdasarkan akumulasi penurunan <i>Gini Impurity</i> .
<i>Gini Impurity</i>	Metrik yang digunakan oleh algoritma pohon keputusan untuk mengukur tingkat ketidakmurnian suatu <i>node</i> .
<i>Machine Learning</i>	Cabang dari kecerdasan buatan yang memungkinkan sistem belajar dari data tanpa diprogram secara eksplisit.
<i>Meta Learner</i>	Model tingkat atas (Level-1) dalam arsitektur <i>Stacking Ensemble</i> (Logistic Regression) yang menggabungkan hasil prediksi <i>Base Learners</i> .
<i>Out-of-Fold Predictions</i>	Prediksi yang dihasilkan dari data yang belum pernah dilihat oleh <i>Base Learners</i> selama pelatihan, digunakan sebagai input bagi <i>Meta Learner</i> .
<i>Overfitting</i>	Kondisi di mana model terlalu menyesuaikan diri dengan pola pada data latih, termasuk <i>noise</i> , sehingga performanya menurun pada data baru.
<i>Stacking</i>	Teknik <i>Ensemble Learning</i> lanjutan yang menggabungkan prediksi dari <i>Base Learners</i> yang berbeda untuk dilatih kembali oleh <i>Meta Learner</i> .

## INTISARI

Permasalahan kesehatan mental, khususnya depresi pada mahasiswa, memerlukan deteksi dini karena berdampak besar terhadap performa akademik dan kesejahteraan secara keseluruhan. Meskipun *machine learning* telah banyak diterapkan untuk memprediksi depresi, sebagian besar penelitian masih berfokus pada model tunggal dan belum mengeksplorasi pendekatan *stacking ensemble* dengan strategi *cross-validation stacking* serta analisis *feature importance* untuk meningkatkan interpretabilitas model. Penelitian ini bertujuan mengembangkan model prediksi depresi pada mahasiswa menggunakan metode *stacking ensemble* yang mengintegrasikan *Random Forest*, *Gradient Boosting*, dan *XGBoost* sebagai base learner serta *Logistic Regression* sebagai meta-learner. Dataset yang digunakan berasal dari Kaggle dan terdiri dari 502 data mahasiswa dengan 11 atribut. Tahapan *preprocessing* meliputi pengecekan nilai kosong, *feature encoding*, dan normalisasi. Kinerja model dievaluasi menggunakan *accuracy*, *precision*, *recall*, *F1-score*, *ROC-AUC*, dan *precision-recall curve*. Model yang dihasilkan memiliki performa sangat baik dengan *accuracy* sebesar 98.02%, *precision* 96%, *recall* 100%, dan *F1-score* 98% untuk kelas depresi, dan *ROC-AUC* sebesar 99.8%. Selain itu, hasil analisis *feature importance* menunjukkan bahwa *Academic Pressure* dan *Suicidal Thoughts* merupakan faktor yang paling berpengaruh dalam memprediksi depresi pada mahasiswa. Penelitian ini berkontribusi pada pengembangan model prediksi kesehatan mental berbasis *machine learning* dan dapat diterapkan sebagai sistem pendukung keputusan dalam deteksi dini depresi pada mahasiswa.

**Kata kunci:** Depresi, mahasiswa, machine learning, stacking ensemble, prediksi.

## **ABSTRACT**

*Mental health issues, particularly depression among university students, require early detection due to their significant impact on academic performance and overall well-being. Although machine learning has been widely applied to predict depression, most existing studies still rely on single-model approaches and have not yet explored stacking ensemble combined with cross-validation stacking and feature importance analysis to enhance model interpretability. This study aims to develop a depression prediction model for university students using stacking ensemble that integrates Random Forest, Gradient Boosting, and XGBoost as base learners, and Logistic Regression as the meta-learner. The dataset used in this study was obtained from Kaggle and consists of 502 student records with 11 attributes. The preprocessing stages included handling missing values, feature encoding, and normalization. The model's performance was evaluated using accuracy, precision, recall, F1-score, ROC-AUC, and the precision-recall curve. The resulting model demonstrated excellent performance, achieving 98.02% accuracy, 96% precision, 100% recall, and a 98% F1-score for the depression class, with a ROC-AUC of 99.8%. Additionally, the feature importance analysis indicated that Academic Pressure and Suicidal Thoughts are the most influential factors in predicting depression among university students. This study contributes to the development of machine learning-based mental health prediction models and can serve as a decision support system for the early detection of depression in university students.*

**Keyword:** *Depression, university students, machine learning, ensemble, prediction.*