

KLASIFIKASI KESEHATAN KUDA BERBASIS DATA BIOMETRIK MENGGUNAKAN ENSEMBLE LEARNING

ANTONIO



**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Klasifikasi Kesehatan Kuda Berbasis Data Biometrik Menggunakan *Ensemble Learning*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apapun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, 31 Juli 2025

Antonio

G6401211111

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRAK

ANTONIO. Klasifikasi Kesehatan Kuda Berbasis Data Biometrik Menggunakan *Ensemble Learning*. Dibimbing oleh SONY HARTONO WIJAYA dan ENDANG PURNAMA GIRI.

Kuda merupakan aset berharga dalam industri peternakan. Kesehatan kuda seringkali menjadi perhatian karena potensi penyakit yang dapat menyebabkan kerugian ekonomi signifikan. Diagnosis dini dan akurat pada kesehatan kuda merupakan hal yang penting untuk mencegah dan mengurangi dampak penyakit. Penelitian ini mengimplementasikan metode *Ensemble Learning*, yang memanfaatkan metode *bagging*, *boosting*, dan *stacking*, untuk mengklasifikasikan kondisi kesehatan kuda ke dalam kelas sehat dan sakit. Data yang digunakan mencakup detak jantung, suhu tubuh, laju pernafasan, dan saturasi oksigen. Perbandingan performa tiga metode tersebut dilakukan untuk menentukan metode terbaik dalam klasifikasi kesehatan kuda. Hasil penelitian menunjukkan bahwa metode model *Random Forest* menunjukkan hasil baik dengan performa seimbang dengan akurasi 96.33% dan *recall* 94%. Namun, *Stacking Ensemble learning* tercatat sebagai model terbaik berkat kemampuannya menggabungkan kekuatan berbagai model dasar dengan skor tertinggi, dengan akurasi 96.89%, *recall* 94%, dan AUC 95.27%. Selain itu, *hyperparameter tuning* terbukti berkontribusi signifikan dalam meningkatkan performa model khususnya pada AUC dengan rata-rata 0.98, khususnya dalam mendeteksi kasus kuda sakit. Penelitian ini diharapkan dapat memberikan kontribusi pada bidang peternakan dan kesehatan hewan melalui peningkatan kesejahteraan kuda melalui deteksi dini penyakit.

Kata kunci: Kuda, Klasifikasi, *Ensemble Learning*, Kesehatan Hewan.

ABSTRACT

ANTONIO. *Horse Health Classification Based on Biometric Data Using Ensemble Learning*. Supervised by SONY HARTONO WIJAYA and ENDANG PURNAMA GIRI.

Horses are valuable assets in the livestock industry. Their health is often a major concern due to the potential for diseases that can lead to significant economic losses. Early and accurate diagnosis of horse health is crucial to prevent and mitigate the impact of such diseases. This study implements Ensemble Learning methods utilizing bagging, boosting, and stacking to classify horse health conditions into healthy and sick categories. The data used includes heart rate, body temperature, respiratory rate, and oxygen saturation. A performance comparison of these three methods was conducted to determine the most effective approach for horse health classification. The results show that the Random Forest model achieved good and balanced performance with 96.33% accuracy and 94% recall. However, Stacking Ensemble Learning was recorded as the

best-performing model due to its ability to combine the strengths of various base models, achieving the highest scores with 96.89% accuracy, 94% recall, and an AUC of 95.27%. Furthermore, hyperparameter tuning significantly contributed to improving model performance, particularly in increasing AUC by an average of 0.98%, especially for detecting sick horses. This research is expected to contribute to the livestock and animal health sectors by enhancing horse welfare through early disease detection.

Keywords: Horse, Classification, Ensemble Learning, Animal Health.

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB.

KLASIFIKASI KESEHATAN KUDA BERBASIS DATA BIOMETRIK MENGGUNAKAN ENSEMBLE LEARNING

ANTONIO

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
Dr. Hari Agung Adrianto, S.Kom., M.Si.

Judul Skripsi : Klasifikasi Kesehatan Kuda Berbasis Data Biometrik
Menggunakan *Ensemble Learning*

Nama : Antonio
NIM : G6401211111

Disetujui oleh

Pembimbing 1:

Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
19810809 200812 1 002

Pembimbing 2:

Endang Purnama Giri, S.Kom., M.Kom.
19821010 200604 1 000

Diketahui oleh

Ketua Program Sarjana Ilmu Komputer:

Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
19810809 200812 1 002

Tanggal Ujian:

17 Juni 2025

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan yang Maha Esa atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Penelitian yang dilaksanakan sejak bulan September 2024 sampai bulan Juni 2025 ini berjudul “Klasifikasi Kesehatan Kuda Berbasis Data Biometrik Menggunakan *Ensemble Learning*”.

Terima kasih penulis ucapkan kepada:

- a Kedua orang tua saya, Bapak Yusup Ali dan Ibu Merry Rosa yang telah memberikan dukungan dan doa kepada penulis,
- b Bapak Dr. Sony Hartono Wijaya, S.Kom., M.Kom. dan Endang Purnama Giri, S.Kom., M.Kom. selaku dosen pembimbing yang telah membimbing dan banyak memberi saran dalam penelitian dan penulisan skripsi saya,
- c Alexandra Victoria Gunawan yang merupakan pasangan penulis, yang sudah memberi bantuan, dukungan, dan semangat kepada penulis.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, 2 Juni 2025

Antonio

DAFTAR ISI

DAFTAR TABEL	XII
DAFTAR GAMBAR	XIII
I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	2
I.3 Tujuan	2
I.4 Manfaat	2
I.5 Ruang Lingkup	2
II TINJAUAN PUSTAKA	3
II.1 Kuda	3
II.2 Status Kesehatan Kuda	3
II.3 Klasifikasi	5
II.4 Ensemble Learning	6
II.5 Random Forest	7
II.6 Gradient Boosting	8
II.7 XGBoost	10
II.8 Stacking Ensemble Learning	11
II.9 Evaluasi	12
III METODE	16
III.1 Data Penelitian	16
III.2 Tahapan Penelitian	16
III.3 Pengumpulan Data	16
III.4 Praproses Data	17
III.5 Pembuatan Model	17
III.6 Hyperparameter Tuning	18
III.7 Evaluasi	18
III.8 Peralatan Penelitian	19
IV HASIL DAN PEMBAHASAN	20
IV.1 Hasil Pengumpulan Data	20
IV.2 Hasil Exploratory Data Analysis	21
IV.3 Hasil Praproses Data	24
IV.4 Hasil Pembuatan Model <i>Ensemble Learning</i>	26

IV.5	Hasil Hyperparameter Tuning	28
IV.6	Evaluasi	30
V	SIMPULAN DAN SARAN	32
V.1	Simpulan	32
V.2	Saran	32
	DAFTAR PUSTAKA	33
	LAMPIRAN	35
	RIWAYAT HIDUP	39

@Hak cipta milik IPB University

IPB University



DAFTAR TABEL

1	Detail fitur dan sebarannya yang digunakan	21
2	Inisiasi parameter <i>random forest</i>	26
3	Inisiasi parameter <i>gradient boosting</i>	27
4	Inisiasi parameter <i>xgboost</i>	27
5	Hasil inisiasi model dengan parameter awal	28
6	Kombinasi <i>hyperparameter random forest</i>	29
7	Kombinasi <i>hyperparameter gradient boosting</i>	29
8	Kombinasi <i>hyperparameter xgboost</i>	29
9	Perbandingan performa model sebelum dan sesudah <i>tuning</i>	30

DAFTAR GAMBAR

1	Ilustrasi klasifikasi	5
2	<i>Framework Ensemble Learning</i>	6
3	<i>Framework Random Forest</i>	8
4	<i>Framework Gradient Boosting</i>	9
5	<i>Framework Stacking XGBoost</i>	10
6	<i>Framework Stacking Ensemble Learning</i>	11
7	<i>Confusion Matrix</i>	13
8	Ilustrasi kurva AUC-ROC	15
9	Tahapan Penelitian	16
10	Pemasangan IPB Smart Halter pada Kuda Sehat	20
11	Pie chart perbandingan jumlah kelas sehat dan sakit	22
12	Distribusi fitur biometrik berdasarkan kelas kesehatan	23
13	Korelasi fitur biometrik dataset	24
14	Kondisi awal dataset	25
15	Perbandingan distribusi setelah penyesuaian <i>outlier</i>	25
16	Perbandingan distribusi dataset setelah dilakukan normalisasi dataset	26
17	<i>Confusion matrix model stacking ensemble learning</i>	31