



**Perbandingan Optimasi *Hyperparameter Gaussian Naïve Bayes*
dan *Bernoulli Naïve Bayes* Menggunakan *Tree-Structured Parzen*
Estimator (TPE) dalam Memprediksi Risiko Penyakit Jantung**

**TUGAS AKHIR
SKRIPSI**

MARCELLINUS RYAN DANUARTA

41822010048

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**



**Perbandingan Optimasi *Hyperparameter Gaussian Naïve Bayes*
dan *Bernoulli Naïve Bayes* Menggunakan *Tree-Structured Parzen*
Estimator (TPE) dalam Memprediksi Risiko Penyakit Jantung**

**TUGAS AKHIR
SKRIPSI**

**MARCELLINUS RYAN DANUARTA
41822010048**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar
sarjana**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2026

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Marcellinus Ryan Danuarta

NIM : 41822010048

Program Studi : Sistem Informasi

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:
“ Perbandingan Optimasi Hyperparameter Gaussian Naïve Bayes dan Bernoulli Naïve Bayes Menggunakan Tree-structured Parzen Estimator (TPE) dalam Memprediksi Risiko Penyakit Jantung” adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.



Jakarta, 30 Juli 2026



Marcellinus Ryan Danuarta

UNIVERSITAS
MERCU BUANA

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name : Marcellinus Ryan Danuarta
NIM /Student ID Number : 41822010048
Program Studi /Study Program : Sistem Informasi

Dengan Judul Tugas Akhir
/The title:

“Perbandingan Optimasi Hyperparameter Gaussian Naïve Bayes dan Bernoulli Naïve Bayes Menggunakan Tree-structured Parzen Estimator (TPE) dalam Memprediksi Risiko Penyakit Jantung”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

16 Juli 2026

dengan nilai persentase sebesar :
/with a percentage value of:

27%

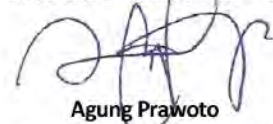
dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana. */declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:
/Turnitin similarity report file

https://drive.google.com/file/d/1HVbmsJmN6AGiUelHOnNPbeXgsDd2pGeh/view?usp=drive_link



Jakarta, 16 Juli 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto

Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

HALAMAN PERSETUJUAN

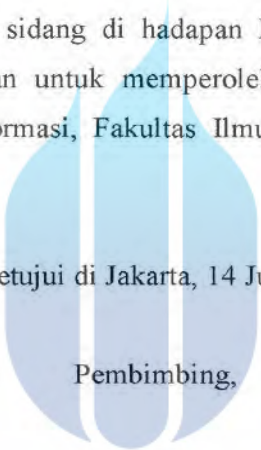
Laporan Skripsi ini diajukan oleh:

Nama : Marcellinus Ryan Danuarta
NIM : 41822010048
Program Studi : Sistem Informasi
Judul Proposal Penelitian : Perbandingan Optimasi Hyperparameter
Gaussian Naïve Bayes dan Bernoulli Naïve
Bayes Menggunakan Tree-structured Parzen
Estimator (TPE) dalam Memprediksi Risiko
Penyakit Jantung

Untuk dipertahankan pada sidang di hadapan Dewan Penguji sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disetujui di Jakarta, 14 Juli 2026

Pembimbing,



UNIVERSITAS
MERCU BUANA

(Nia Rahma Kurnianda, S.Kom., M.Kom)

NIDN/NUPTK: 0323098803

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Marcellinus Ryan Danuarta

NIM : 41822010048

Program Studi : Sistem Informasi

Judul Laporan Skripsi : Perbandingan Optimasi Hyperparameter Gaussian Naïve Bayes dan Bernoulli Naïve Bayes Menggunakan Tree-structured Parzen Estimator (TPE) dalam Memprediksi Risiko Penyakit Jantung

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:
Pembimbing,

(Nia Rahma Kurnianda, S.kom., M.kom.)
NIDN/NUPTK: 0323098803

Jakarta, 30 Juli 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer

Ketua Program Studi
Sistem Informasi

Dr. Bambang Jokonowo, S.Si., M.T.I.
NIDN: 0320037002

Wawan Gunawan, S.Kom., M.T., M.Kom.
NIDN: 0424108104

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof. Dr. Andi Adriansyah, M.Eng., selaku Rektor Universitas Mercu Buana
2. Prof. Dr. Bambang Jokonowo, S.Si, M.T.I selaku Dekan Fakultas Ilmu Komputer
3. Wawan Gunawan, S.Kom, MT., M.Kom selaku Ketua Program Studi Sistem Informasi
4. Nia Rahma Kurnianda, S.Kom., M.Kom selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini
5. Kedua orang tua dan adik penulis yang senantiasa memberikan doa, dukungan, serta motivasi selama proses perkuliahan hingga penyusunan skripsi ini
6. Pasangan penulis yang telah memberikan perhatian, dukungan moral, serta semangat dalam membantu penulis menyelesaikan skripsi ini

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 17 Desember 2025



Marcellinus Ryan Danuarta

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Marcellinus Ryan Danuarta
NIM : 41822010048
Fakultas/Program Studi : Fakultas Ilmu Komputer / Sistem Informasi
Judul Tugas Akhir : Perbandingan Optimasi Hyperparameter Gaussian Naïve Bayes dan Bernoulli Naïve Bayes Menggunakan Tree-structured Parzen Estimator (TPE) dalam Memprediksi Risiko Penyakit Jantung

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, 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.

Jakarta, 30 Juli 2026
Yang menyatakan,

UNIVERSITAS
MERCU BUANA



(Marcellinus Ryan Danuarta)

Perbandingan Optimasi *Hyperparameter Gaussian Naïve Bayes* dan *Bernoulli Naïve Bayes* Menggunakan *Tree-Structured Parzen Estimator (TPE)* dalam Memprediksi Risiko Penyakit Jantung
MARCELLINUS RYAN DANUARTA

ABSTRAK

Prediksi risiko penyakit jantung memerlukan model klasifikasi yang mampu mengolah karakteristik data klinis secara akurat. Perbedaan tipe atribut pada dataset kesehatan, mulai dari data numerik hingga data biner, menyebabkan setiap varian algoritma *Naïve Bayes* memiliki kemampuan prediksi yang berbeda. Selain karakteristik algoritma, konfigurasi *Hyperparameter* juga berpengaruh terhadap kualitas model yang dihasilkan. Atas dasar tersebut, penelitian ini mengevaluasi efektivitas optimasi *Hyperparameter* sekaligus membandingkan performa *Gaussian Naïve Bayes (GaussianNB)* dan *Bernoulli Naïve Bayes (BernoulliNB)* dalam memprediksi risiko penyakit jantung.

Penelitian ini menggunakan *Heart Failure Prediction Dataset* yang diperoleh melalui platform *Kaggle* sebagai sumber data eksperimen. Sebelum proses pemodelan dilakukan, data terlebih dahulu melewati tahapan prapemrosesan yang meliputi *Data Cleaning*, transformasi variabel kategorikal menggunakan teknik *One-Hot Encoding*, normalisasi nilai fitur, serta pembagian dataset ke dalam data pelatihan dan data pengujian. Selanjutnya, optimasi *Hyperparameter* diterapkan dengan memanfaatkan metode *Tree-Structured Parzen Estimator (TPE)*. Kinerja model kemudian dianalisis menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk menilai kemampuan klasifikasinya.

Performa terbaik diperoleh oleh *Bernoulli Naïve Bayes* setelah proses optimasi diterapkan. Model tersebut menghasilkan *accuracy* sebesar 89%, sedangkan *Gaussian Naïve Bayes* memperoleh 87%. Keunggulan *Bernoulli Naïve Bayes* juga terlihat pada nilai *recall* dan *F1-score* untuk kelas *Disease*, yang menunjukkan kemampuan lebih baik dalam mengenali pasien dengan risiko penyakit jantung. Perbandingan yang dilakukan memperlihatkan bahwa optimasi *Hyperparameter* menggunakan *TPE* memberikan kontribusi terhadap peningkatan performa

klasifikasi pada kedua model. Berdasarkan hasil evaluasi, *Bernoulli Naïve Bayes* menjadi algoritma yang paling sesuai untuk dataset yang digunakan dan berpotensi diterapkan sebagai pendekatan dalam pengembangan sistem prediksi risiko penyakit jantung berbasis *Machine learning*.

Kata kunci: *Naïve Bayes, GaussianNB, BernoulliNB, Tree-Structured Parzen Estimator (TPE), optimasi Hyperparameter, penyakit jantung, Machine learning.*



Comparison of Hyperparameter Optimization of Gaussian Naïve Bayes and Bernoulli Naïve Bayes Using Tree-Structured Parzen Estimator (TPE) in Predicting Heart Disease Risk

MARCELLINUS RYAN DANUARTA

ABSTRACT

The effectiveness of a Machine learning model in predicting heart Disease risk is influenced by the suitability of the classification algorithm and the selected Hyperparameter configuration. Since medical datasets generally contain both numerical and binary attributes, each Naïve Bayes variant may exhibit different predictive capabilities. Therefore, evaluating the performance of different model variants and examining the contribution of Hyperparameter optimization are important to obtain a more reliable classification model.

This study compares the performance of Gaussian Naïve Bayes (GaussianNB) and Bernoulli Naïve Bayes (BernoulliNB) by applying Hyperparameter optimization using the Tree-Structured Parzen Estimator (TPE) . The experiment utilized the Heart Failure Prediction Dataset obtained from Kaggle, which contains patients' clinical information related to heart Disease risk. Prior to model development, the dataset underwent Data Cleaning, categorical feature transformation using One-Hot Encoding, normalization, and train-test splitting. Hyperparameter optimization was then performed to identify the most appropriate configuration before evaluating both models using accuracy, precision, recall, and F1-score.

The experimental findings demonstrate that the TPE optimization approach improves the predictive performance of both classifiers. GaussianNB achieved an accuracy of 87%, whereas BernoulliNB reached 89%. In addition to achieving higher accuracy, BernoulliNB also produced superior recall and F1-score values for the Disease class, indicating a stronger capability in identifying patients with a higher risk of heart Disease.

The results suggest that Hyperparameter optimization contributes positively to improving the predictive capability of Naïve Bayes classifiers. Among the evaluated

models, Bernoulli Naïve Bayes with the optimized Hyperparameter configuration provided the best overall performance for heart Disease risk prediction using the dataset employed in this study.

Keywords: Naïve Bayes, GaussianNB, BernoulliNB, Tree-Structured Parzen Estimator (TPE), Hyperparameter optimization, heart Disease, Machine learning.



DAFTAR ISI

HALAMAN SAMPUL.....	0
HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN SURAT KETERANGAN UJI TURNITIN	iii
HALAMAN PENGESAHAN.....	v
KATA PENGANTAR.....	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB	vii
ABSTRAK	viii
ABSTRACT	x
DAFTAR ISI.....	xii
DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Teori Utama	7
2.1.1 Analisis Data dan <i>Machine learning</i> dalam Pengambilan Keputusan Medis.....	7
2.1.2 Algoritma <i>Naïve Bayes</i>	9
2.1.3 Jenis-jenis <i>Naïve Bayes</i>	10
2.1.4 Penyakit Jantung.....	14
2.2 Penelitian Terdahulu.....	16
2.3 Teori Pendukung	33
2.3.1 <i>Data mining</i> dan <i>Machine learning</i>	33

2.3.2 Evaluasi Model Klasifikasi.....	34
BAB III METODE PENELITIAN	36
3.1 Jenis Penelitian	36
3.2 Tahapan Penelitian.....	37
3.3 Instrumen Penelitian	40
3.4 Subjek Penelitian	42
3.5 <i>Framework</i> Algoritma Yang Digunakan	44
3.5.1 Perbedaan <i>Framework</i> Penelitian dengan Implementasi <i>Naïve Bayes</i> Konvensional	49
3.5.2 Tahapan yang Mempengaruhi Performa Model	51
BAB IV PEMBAHASAN.....	55
4.1 Akuisisi dan Persiapan Dataset.....	55
4.2 <i>Preprocessing Data</i>	57
4.2.1 Pemeriksaan Kualitas Data	58
4.2.2 <i>Data Cleaning</i>	60
4.2.3 Transformasi Data.....	62
4.3 <i>Split Data</i>	64
4.4 Implementasi Model <i>Gaussian Naïve Bayes</i>	66
4.4.1 Persiapan Data	67
4.4.2 <i>Split & Normalize Cleaned Dataset</i>	68
4.4.3 <i>Model Training with TPE</i>	69
4.4.4 <i>Model Evaluation</i>	69
4.5 Implementasi Model <i>Bernoulli Naïve Bayes</i>	70
4.5.1 Persiapan Data	71
4.5.2 <i>Split & Normalize Cleaned Dataset</i>	72
4.5.3 <i>Model Training with TPE</i>	72
4.5.4 <i>Model Evaluation</i>	73
4.6 Penerapan Metode <i>Tree-Structured Parzen Estimator (TPE)</i> Untuk Optimasi <i>Hyperparameter</i>	74
4.6.1 Optimasi <i>Hyperparameter Gaussian Naïve Bayes</i>	75
4.6.2 Optimasi <i>Hyperparameter Bernoulli Naïve Bayes</i> Menggunakan <i>Tree-Structured Parzen Estimator (TPE)</i>	77

4.6.3 Perbandingan Hasil Optimasi <i>Hyperparameter Gaussian Naïve Bayes</i> dan <i>Bernoulli Naïve Bayes</i>	80
4.7 Hasil Evaluasi Dan Perbandingan Model	82
4.7.1 Evaluasi Model <i>Gaussian Naïve Bayes</i>	82
4.7.2 Evaluasi Model <i>Bernoulli Naïve Bayes</i>	84
4.7.3 Perbandingan Performa Model.....	85
4.7.4 Perbandingan Hasil Penelitian dengan Penelitian Sebelumnya.....	87
BAB V KESIMPULAN DAN SARAN	90
5.1 Kesimpulan	90
5.2 Saran.....	91
DAFTAR PUSTAKA.....	93
LAMPIRAN.....	96



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	16
Tabel 3. 1 Spesifikasi <i>Hardware</i>	40
Tabel 3. 2 Versi <i>Software</i>	40
Tabel 3. 3 Variabel Data.....	43
Tabel 3. 4 Perbedaan Penelitian Ini dengan <i>Naïve Bayes</i> Biasa	50
Tabel 4. 1 Menyajikan beberapa baris data sebagai representasi dari dataset yang digunakan dalam penelitian ini.	55
Tabel 4. 2 Hasil Pemeriksaan Awal Dataset	59
Tabel 4. 3 Hasil Pembersihan Data	62
Tabel 4. 4 Transformasi Atribut Kategorikal	63
Tabel 4. 5 Hasil Pembagian Dataset	65
Tabel 4. 6 Skenario Optimasi <i>Hyperparameter Gaussian Naïve Bayes</i>	76
Tabel 4. 7 Skenario Optimasi <i>Hyperparameter Bernoulli Naïve Bayes</i>	79
Tabel 4. 8 Perbandingan Hasil Optimasi <i>Hyperparameter</i>	80
Tabel 4. 9 Hasil Evaluasi <i>Gaussian Naïve Bayes</i>	83
Tabel 4. 10 Hasil Evaluasi <i>Bernoulli Naïve Bayes</i>	85
Tabel 4. 11 Perbandingan Performa Model	86
Tabel 4. 12 Perbandingan Hasil Penelitian dengan Penelitian Sebelumnya .	87

DAFTAR GAMBAR

Gambar 3. 1 <i>Framework</i> Penelitian.....	37
Gambar 3. 2 Gambar Dataset	42
Gambar 3. 3 <i>Framework</i> Algoritma Penelitian	45
Gambar 4. 1 Mengecek tipe data	58
Gambar 4. 2 Mengecek Data Duplikat.....	58
Gambar 4. 3 Mengecek <i>Missing Value</i>	59
Gambar 4. 4 <i>Change Data Type</i>	61
Gambar 4. 5 <i>Remove Anomali</i>	61
Gambar 4. 6 Hasil <i>One-Hot Encoding</i>	63
Gambar 4. 7 Proses Pembagian Data Latih dan Data Uji.....	65
Gambar 4. 8 Persiapan Data Hasil <i>Preprocessing</i>	68
Gambar 4. 9 <i>Split & Normalize Cleaned Dataset</i>	69
Gambar 4. 10 <i>Model Training with TPE</i>	69
Gambar 4. 11 <i>Model Evaluation</i>	70
Gambar 4. 12 Persiapan Data Hasil <i>Preprocessing</i>	71
Gambar 4. 13 <i>Split & Normalize Cleaned Dataset</i>	72
Gambar 4. 14 <i>Model Training with TPE</i>	73
Gambar 4. 15 <i>Model Evaluation</i>	74
Gambar 4. 16 Optimasi <i>Hyperparameter Gaussian Naïve Bayes</i> Menggunakan TPE	76
Gambar 4. 17 Optimasi <i>Hyperparameter Bernoulli Naïve Bayes</i>	78
Gambar 4. 18 Hasil Evaluasi Model <i>Gaussian Naïve Bayes</i>	82
Gambar 4. 19 Hasil Evaluasi Model <i>Bernoulli Naïve Bayes</i>	84

MERCU BUANA

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi.....	96
Lampiran 2 Curriculum Vitae.....	97
Lampiran 3 Surat Keterangan Telah Mengikuti BNSP	98
Lampiran 4 Hasil Cek Turnitin.....	99
Lampiran 5 Surat Keterangan Bebas Pustaka.....	100

