KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN *RANDOM FOREST*

TUGAS AKHIR

APLIKATIF

AKMAL FATURRAHMAN

4151911000

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2026

lib.mercubuana.ac.id

KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN *RANDOM FOREST*

TUGAS AKHIR

APLIKATIF

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

AKMAL FATURRAHMAN

4151911000

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2026

lib.mercubuana.ac.id

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : AKMAL FÅTURRAHMAN

NIM : 41519110001

Program Studi : Teknik Informatika

Judul Laporan Skripsi : KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN RANDOMFOREST

Menyatakan dengan seben -benarnya bahwa Tugas Akhir berjudul:

KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN RANDOM FOREST 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, I Agustus 2026



Akmal Faturrahman.

lib.mercubuana.ac.id

HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN

072.423.4.03.01

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	**: Akmal Faturrahman**
NIM /Student ID Number	**: 41519110001**
Program Studi /Study Program	**: Teknik Informatika**

Dengan Judul Tugas Akhir
The title:

"KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN RANDOM FOREST"

telah dilakukan pengecekan similarity dengan sistem Turnitin pada tanggal:
Similarity checks have been carried out with the Turnitin system on the date:

23 Agustus 2026

dengan nilai persentase sebesar :
with a percentage value of:

13%

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/17ks-Y2dW-02ebA9ddq-Ag6BNZ_upqre/view?usp=drive_link

Jakarta, 23 Agustus 2026
Admin Turnitin Fasilkom UMB



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

lib.mercubuana.ac.id

HALAMAN PENGESAHAN

Tugas Akhir İni diajukan oleh:

Nama : AKMAL FATURRAHMAN

NİM : 41519110001

Program Studi .: .Teknik Informatika

Judul Tugas Akhir .:.KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN RANDOM FOREST

Telah berhasil dipertahankan pada sidang tanggal II September 2026 dan diterima sebagai bagİan persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studİ Teknik Informatika, Fakültas limu Kompüter Universitas Mercu Buana.

Diğahkan oleh:

Pembimbing

Mohamad Yusuf, S.Kom., M.C.S

NIDN/NUPTK: 0307097606

Jakarta, I I September 2026

Mengetahui,

Dekan Fakültas limu Kompüter	Ketua Program Studi Teknik Informatika
Dr. Bambang Jokonowo, S. Si-, M.Ti	Dr.lr. Hadi Santoso, S. Kom., M. Kom
NIDN/NUJPTK: 0320037002	NIDN/NUJPTK•. 0225067701

lib.mercubuana.ac.id

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, karunia, serta pertolongan-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul **"Klasifikasi Rumah Sakit Berdasarkan Kelas Menggunakan Random Forest"** dengan baik.

Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan pendidikan dan memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana. Penelitian ini membahas penerapan metode *Machine Learning*, khususnya algoritma *Random Forest*, dalam melakukan klasifikasi kelas rumah sakit berdasarkan karakteristik data profil rumah sakit. Melalui penelitian ini, penulis berupaya memberikan gambaran mengenai pemanfaatan teknologi berbasis data sebagai salah satu pendekatan yang dapat digunakan untuk mendukung proses klasifikasi rumah sakit secara lebih objektif dan sistematis.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa penyelesaian penelitian tidak terlepas dari bantuan, dukungan, bimbingan, serta doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT yang telah memberikan kesehatan, kekuatan, kemudahan, dan kelancaran kepada penulis selama proses penyusunan skripsi.
2. Kedua orang tua serta keluarga tercinta yang senantiasa memberikan doa, dukungan, motivasi, perhatian, dan semangat kepada penulis selama menempuh pendidikan hingga penyelesaian skripsi ini.
3. **Bapak Prof. Dr. Andi Adriansyah, M.Eng.**, selaku Rektor Universitas Mercu Buana.
4. **Bapak Dr. Bambang Jokonowo, S.Si., MTI**, selaku Dekan Fakultas Ilmu Komputer Universitas Mercu Buana.
5. **Bapak Dr.Ir. Hadi Santoso, S.Kom., M.Kom.**, selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
6. **Bapak Mohamad Yusuf, S.Kom., M.C.S.**, selaku dosen pembimbing yang telah meluangkan waktu untuk memberikan bimbingan, arahan, masukan, serta saran kepada penulis selama proses penyusunan penelitian dan skripsi ini.

lib.mercubuana.ac.id

7. Seluruh dosen Program Studi Teknik Informatika Universitas Mercu Buana yang telah memberikan ifinu pengetahuan, wawasan, dan pengalaman kepada penulis selama menjalani masa perkuliahan.
8. Seluruh staf dan tenaga kependidikan Universitas Mercu Buana yang telah membantu penulis dalam berbagai proses administrasi selama masa perkuliahan.
9. Teman-teman mahasiswa Program Studi Teknik Informatika yang telah memberikan dukungan, semangat, bantuan, serta menjadi rekan berdiskusi selama proses perkuliahari dan penyusunan skripsi.
10. Seluruh pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan, dÜkungan, maupun doa sehingga skrİpsİ ini dapat diselesaikan.

Penulis menyadari bahwa skriisi ini masih memiliki keterbatasan dan belum sepenuhnya sempuma. Hal tersebut tidak terlepas dari keterbatasan pengetahuan, pengalaman, serta kemampuan yang dimiliki penulis. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun sebagai bahan perbaikan dan pengembangan di masa mendatang.

Akhir kata, penulis berharap skrİpsi ini dapat memberikan manfaat bagİ pembaca, khususnya dalam menambah wawasan mengenai penerapan Machine Learning dan algoritma Random Forest dalam proses klasifikasi data di bidang kesehatan. Penulis juga berharap hasil penelitian ini dapat menjadi referensi bagi penelitian selanjumya serta memberikan kontribusi dalam pengembangan sistem pendukung keputusan berbasis data pada bidang pelayanan kesehatan.

UNIVERSITAS MERCU BUANA

Jakarta 20 Agustus 2026



Akmal Faturrahman

lib.mercubuana.ac.id

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR Dl REPOSITORI UMB

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

Nama	: AKMAL FATURRAHMAN
Nim	: 41519110001
Program Studİ	: Teknik Informatika
Judul Tugas Akhir	: KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MEGGUNAKAN RANDOM FOREST

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

Dengan Hak Bebas Royaltİ Non-Eksklusif İni 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 İni saya buat dengan sebenarnya.

UNIVERSITAS MERCU BUANA

Jakarta, 20 Agustus 2026

Yang menyatakan,



AKMAL FATUR RAHMAN

lib.mercubuana.ac.id

KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN *RANDOM FOREST*

AKMAL FATURRAHMAN

ABSTRAK

Penelitian ini menguji Random Forest untuk mengenali kelas A, B, C, dan D pada data profil rumah sakit. Dataset memuat 400 rumah sakit dan enam atribut awal. Pemodelan hanya menggunakan jenis_layanan dan jumlah_kasur; kode_rs, nama_rs, serta alamat dipertahankan sebagai identitas, sedangkan kelas menjadi target. Kategori jenis_layanan diubah dengan One-Hot Encoding, sementara median imputation ditempatkan pada pipeline numerik sebagai pengaman apabila data baru mengandung nilai kosong. Pembagian secara stratified menghasilkan 320 data pada training pool dan 80 data pada final test set. Pengembangan model meliputi evaluasi baseline, Repeated Stratified 5-Fold Cross-Validation sebanyak 10 pengulangan, dan pencarian parameter melalui GridSearchCV dengan Macro-F1 sebagai skor pemilihan. Pada final test set, 54 dari 80 rumah sakit diprediksi secara tepat. Model menghasilkan accuracy 67,50%, balanced accuracy 58,55%, Macro-F1 61,53%, dan ROC-AUC One-vs-Rest Macro 83,04%. Analisis feature importance menunjukkan nilai kepentingan relatif sebesar 71,99% pada jumlah_kasur dan 28,01% secara agregat pada seluruh fitur turunan jenis_layanan. Temuan ini memperlihatkan bahwa kedua karakteristik tersebut memuat informasi pembeda kelas, tetapi kemampuan pengenalan kelas A masih terbatas dan distribusi target belum seimbang. Model ditujukan sebagai eksperimen analitis dan bukan sebagai pengganti penetapan kelas rumah sakit oleh instansi berwenang.

Kata kunci: Rumah Sakit, Klasifikasi, *Random Forest*, *Machine Learning*, *Feature Importance*

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

KLASIFIKASI RUMAH SAKIT BERDASARKAN KELAS MENGGUNAKAN *RANDOM FOREST*

AKMAL FATURRAHMAN

ABSTRACT

This study evaluates Random Forest for recognizing hospital classes A, B, C, and D from profile data. The dataset contains 400 hospitals and six original attributes. Only jenis_layanan and jumlah_kasur are used as predictors; kode_rs, nama_rs, and alamat remain identifiers, while kelas is the target. One-Hot Encoding is applied to jenis_layanan, and median imputation is retained in the numeric pipeline to handle possible missing values in future data. A stratified 80:20 split produces a 320-record training pool and an untouched 80-record final test set. Model development includes baseline assessment, 10 repetitions of stratified five-fold cross-validation, and GridSearchCV using Macro-F1 for selection. The final model correctly predicts 54 of 80 test records, with 67.50% accuracy, 58.55% balanced accuracy, 61.53% Macro-F1, and 83.04% macro One-vs-Rest ROC-AUC. Feature importance analysis shows a relative importance of 71.99% for jumlah_kasur and 28.01% in aggregate for the encoded features derived from jenis_layanan. These findings indicate that both characteristics contain class-discriminating information, although class A remains difficult to identify and the target distribution is imbalanced. The model is an analytical experiment rather than a substitute for official hospital-class designation.

Keywords: *Hospital, Classification, Random Forest, Machine Learning, Feature Importance*

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR ISI

HALAMAN SAMPUL 0
HALAMAN JUDUL i
HALAMAN PERNYATAAN KARYA SENDIRI ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN iii
HALAMAN PENGESAHAN iv
KATA PENGANTAR v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI vii
ABSTRAK viii
ABSTRACT ix
DAFTAR ISI x
DAFTAR TABEL v
DAFTAR GAMBAR vii
DAFTAR LAMPIRAN viii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 4
1.3 Tujuan Penelitian 4
1.4 Manfaat Penelitian 5
1.4.1 Manfaat Teoritis 5
1.4.2 Manfaat Praktis 5
1.5 Batasan Masalah 5
1.6 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA 7
2.1 Teori Utama 7
2.2 Teori Pendukung 8
2.2.1 Rumah Sakit 8

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

2.2.2 Data Mining 9
2.2.3 Klasifikasi 13
2.2.4 Random Forest 13
2.2.5 Preprocessing Data 17
2.2.6 Evaluasi Model Klasifikasi 19
2.3 Penelitian Terdahulu 21
2.4 Gap Penelitian 28
2.5 Kerangka Pemikiran 29
BAB III METODOLOGI PENELITIAN 31
3.1 Desain Penelitian 31
3.2 Diagram Alur Penelitian (Flowchart) 31
3.3 Sumber Data 33
3.4 Teknik Pengumpulan Data 34
3.5 Preprocessing Data 35
3.5.1 Normalisasi Data 35
3.5.2 Data Cleaning 35
3.5.3 Encoding Data 35
3.6 Pembentukan Fitur dan Label 36
3.7 Pembagian Data 36
3.8 Pemodelan Machine Learning 36
3.8.1 Random Forest 36
3.8.2 Cross-Validation 37
3.8.3 Optimasi Hyperparameter 37
3.9 Evaluasi Model 37
3.10 Visualisasi Model 38
3.11 Perangkat dan Tools Penelitian 38
BAB IV HASIL DAN PEMBAHASAN 40

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

4.1 Gambaran Umum Dataset 40
4.1.1 Karakteristik Dataset 41
4.1.2 Audit Kualitas Data Awal 42
4.1.3 Distribusi Kelas Rumah Sakit 43
4.2 Hasil Data Cleaning dan Preprocessing 45
4.2.1 Normalisasi dan Pembersihan Data 45
4.2.2 Penanganan Missing Value 46
4.2.3 Pemilihan Fitur 47
4.2.4 Transformasi Data 48
4.3 Pembagian Data dan Strategi Validasi 50
4.3.1 Pembentukan Training Pool dan Final Test Set 50
4.3.2 Stratified 5-Fold Cross-Validation 52
4.3.3 Repeated Cross-Validation 53
4.4 Hasil Model Baseline Random Forest 55
4.4.1 Performa Baseline pada Repeated Cross-Validation 55
4.4.2 Analisis Model Baseline 57
4.5 Hyperparameter Tuning Random Forest 58
4.5.1 Proses Hyperparameter Tuning 59
4.5.2 Hyperparameter Terbaik 59
4.5.3 Hasil GridSearchCV 60
4.5.4 Perbandingan Sepuluh Konfigurasi Terbaik 61
4.5.5 Analisis Hasil Hyperparameter Tuning 62
4.6 Evaluasi Model Final dengan Repeated Cross Validation 63
4.6.1 Hasil Validasi Model Final 64
4.6.2 Analisis Generalization Gap 65
4.6.3 Perbandingan Model Baseline dan Model Final 66
4.6.4 Interpretasi Kestabilan Model Final 67

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

4.7 Analisis Learning Curve 69
4.7.1 Hasil Learning Curve 69
4.7.2 Interpretasi Learning Curve 71
4.8 Hasil Pengujian pada Final Test Set 72
4.8.1 Hasil Prediksi Umum 72
4.8.2 Metrik Evaluasi Final Test Set 74
4.9 Analisis Classification Report 75
4.9.1 Hasil Classification Report per Kelas 75
4.9.2 Pembahasan Performa per Kelas 75
4.10 Analisis Confusion Matrix 76
4.10.1 Hasil Confusion Matrix 76
4.10.2 Interpretasi Kesalahan Klasifikasi 77
4.11 ROC Curve dan Nilai AUC 78
4.11.1 ROC Curve Multiclass One-vs-Rest 78
4.11.2 Nilai AUC per Kelas 79
4.11.3 Interpretasi ROC-AUC 79
4.12 Analisis Feature Importance 79
4.12.1 Feature Importance Individual 79
4.12.2 Feature Importance Berdasarkan Kelompok Variabel 81
4.12.3 Interpretasi Feature Importance 82
4.13 Pembahasan Hasil Akhir 82
BAB V PENUTUP 84
5.1 Kesimpulan 84
5.2 Saran 85
DAFTAR PUSTAKA 87
LAMPIRAN 89
Lampiran 1. Kartu Asistensi 89

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

Lampiran 2. Curiculum Vitae 90

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu 21
Tabel 4. 1 Atribut Dataset Rumah Sakit 41
Tabel 4. 2 Cuplikan Dataset 41
Tabel 4. 3 Hasil Pemeriksaan Missing Value 42
Tabel 4. 4 Distribusi Kelas Rumah Sakit 44
Tabel 4. 5 Hasil Normalisasi dan Pembersihan Data 45
Tabel 4. 6 Penanganan Missing Value 46
Tabel 4. 7 Hasil Pemilihan Fitur Model 47
Tabel 4. 8 Hasil Transformasi Data untuk Pemodelan 49
Tabel 4. 9 Pembagian Dataset Penelitian 51
Tabel 4. 10 Distribusi Kelas pada Training Pool dan Final Test Set 51
Tabel 4. 11 Ukuran Data pada Stratified 5-Fold Cross-Validation 53
Tabel 4. 12 Konfigurasi Repeated Cross-Validation 53
Tabel 4. 13 Ringkasan Pembagian Dataset dan Strategi Validasi 54
Tabel 4. 14 Hasil Evaluasi Model Baseline Random Forest 55
Tabel 4. 15 Perbandingan Train dan Validation Macro-F1 Model Baseline 57
Tabel 4. 16 Ringkasan Analisis Model Baseline 57
Tabel 4. 17 Konfigurasi Proses Hyperparameter Tuning 59
Tabel 4. 18 Hyperparameter Terbaik Model Random Forest 59
Tabel 4. 19 Hasil Kandidat Terbaik pada Proses GridSearchCV 61
Tabel 4. 20 Sepuluh Konfigurasi Terbaik Hasil GridSearchCV 61
Tabel 4. 21 Perbandingan Model Baseline dan Hasil Hyperparameter Tuning 63
Tabel 4. 22 Hasil Evaluasi Model Final pada Repeated Cross-Validation 64
Tabel 4. 23 Perbandingan Training dan Validation Macro-F1 Model Final 65
Tabel 4. 24 Perbandingan Model Baseline dan Model Final 66
Tabel 4. 25 Hasil Learning Curve Model Random Forest 70

lib.mercubuana.ac.id

Tabel 4. 26 Ringkasan Hasil Prediksi pada Final Test Set 73
Tabel 4. 27 Metrik Evaluasi Model pada Final Test Set..... 74
Tabel 4. 28 Hasil Classification Report pada Final Test Set..... 75
Tabel 4. 29 Confusion Matrix pada Final Test Set 76
Tabel 4. 30 Nilai AUC pada Masing-Masing Kelas..... 79
Tabel 4. 31 Feature Importance Individual Model Random Forest 80

DAFTAR GAMBAR

Gambar 2. 1 Kerangka Pemikiran 30
Gambar 3. 1 Flowchart Tahapan Penelitian 32
Gambar 4. 1 Pembagian Training Pool dan Final Test Set 52
Gambar 4. 2 Distribusi Training dan Validation Macro-F1 68
Gambar 4. 3 Distribusi Generalization Gap..... 69
Gambar 4. 4 Learning Curve Model Random Forest 71
Gambar 4. 5 Hasil Prediksi pada Final Test Set 73
Gambar 4. 6 Confusion Matrix Final Test Set..... 77
Gambar 4. 7 ROC Curve Multiclass One-vs-Rest..... 78
Gambar 4. 8 Feature Importance Individual Random Forest 80
Gambar 4. 9 Feature Importance Berdasarkan Kelompok Variabel 81

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi 89
Lampiran 2. Curiculum Vitae 89
Lampiran 3. Sertifikat BNSP 90
Lampiran 4. Surat Pernyataan & Pengalihan HAKI..... 91

lib.mercubuana.ac.id