

**SKRIPSI**



**PREDIKSI JENIS PERAWATAN PADA PASIEN DI RUMAH  
SAKIT MENGGUNAKAN METODE *ENSEMBLE LEARNING*  
BERDASARKAN HASIL TES LABORATORIUM PASIEN**

**ADRIAN DWI ADINATA**

**NIM. 1910511038**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**2023**

**SKRIPSI**  
**Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana**  
**Komputer**



**PREDIKSI JENIS PERAWATAN PADA PASIEN DI RUMAH**  
**SAKIT MENGGUNAKAN METODE *ENSEMBLE LEARNING***  
**BERDASARKAN HASIL TES LABORATORIUM PASIEN**

**ADRIAN DWI ADINATA**  
**NIM. 1910511038**

**PROGRAM STUDI INFORMATIKA**  
**FAKULTAS ILMU KOMPUTER**  
**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**  
**2023**

## PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Adrian Dwi Adinata

NIM : 1910511038

Tanggal : 7 Juli 2023

Judul Skripsi : **Prediksi Jenis Perawatan Pada Pasien Di Rumah Sakit  
Menggunakan Metode *Ensemble Learning* Berdasarkan Hasil Tes  
Laboratorium Pasien**

Bilamana pada kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 7 Juli 2023

Yang Menyatakan,



Adrian Dwi Adinata

**PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK  
KEPENTINGAN AKADEMIS**

Sebagai civitas akademika Universitas Pembangunan Nasional "Veteran" Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Adrian Dwi Adinata  
NIM : 1910511038  
Fakultas : Ilmu Komputer  
Program Studi : SI Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan karya ilmiah saya kepada Universitas Pembangunan Nasional "Veteran" Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exchange Royalty Free Right*) untuk dipublikasikan dengan judul:

**Prediksi Jenis Perawatan Pada Pasien Di Rumah Sakit Menggunakan  
Metode *Ensemble Learning* Berdasarkan Hasil Tes Laboratorium Pasien**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional "Veteran" Jakarta berhak menyimpan, mengalih media atau memformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 7 Juli 2023

Yang Menyatakan,



Adrian Dwi Adinata

## LEMBAR PENGESAHAN

Tugas Akhir ini diajukan oleh:


Nama : Adrian Dwi Adinata

NIM : 1910511038

Program Studi : SI Informatika

Judul Tugas Akhir : PREDIKSI JENIS PERAWATAN PADA PASIEN DI RUMAH SAKIT MENGGUNAKAN METODE ENSEMBLE LEARNING BERDASARKAN HASIL TES LABORATORIUM PASIEN

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

  
Bayu Hananto, S.Kom.,M.Kom.

Penguji I



Rio Wirawan, S.Kom., M.M.S.I.

Penguji II



Dr. Didit Widiyanto, S.Kom, M.Si.

Pembimbing

  
  
Dr. Ernata, M.Kom.  
Dekan

  
Dr. Widya Cholil, M.I.T.  
Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 7 Juli 2023



# **PREDIKSI JENIS PERAWATAN PADA PASIEN DI RUMAH SAKIT MENGGUNAKAN METODE *ENSEMBLE LEARNING* BERDASARKAN HASIL TES LABORATORIUM PASIEN**

**ADRIAN DWI ADINATA**

## **ABSTRAK**

Jenis perawatan di rumah sakit dapat dibagi menjadi rawat inap dan rawat jalan. Sebagai bagian dari sistem otomasi rumah sakit, dibutuhkan sebuah model klasifikasi untuk memprediksi jenis perawatan pasien untuk mempermudah dan mempercepat pengambilan keputusan. *Dataset* yang digunakan diambil dari Kaggle yang merujuk kepada *Electronic Health Record Predicting* yang dikumpulkan dari Rumah Sakit swasta di Indonesia. *Dataset* berisi hasil uji laboratorium darah pasien. Penelitian ini bertujuan untuk membangun model pembelajaran mesin untuk memprediksi apakah pasien harus diklasifikasikan dalam jenis perawatan rawat inap atau rawat jalan. Metode yang digunakan pada penelitian ini adalah *Ensemble Learning* yakni menggabungkan algoritma *logistic regression*, *k-nearest neighbours*, *support vector machine*, *decision tree*, dan *naïve bayes*. Pemodelan yang akan diterapkan menggunakan *Ensemble Vote* dan *Ensemble Stacking* untuk mendapatkan model prediksi dengan akurasi terbaik. Model klasifikasi yang diperoleh akan dievaluasi menggunakan nilai *accuracy*, *precision*, *recall*, dan *F1-Score*. Hasil akurasi terbaik didapatkan pada saat menggunakan algoritma Ensemble Stacking dengan akurasi sebesar 0.786. Dapat disimpulkan bahwa dengan menerapkan Algoritma Ensemble Stacking terbukti dapat meningkatkan akurasi untuk melakukan prediksi jenis perawatan pasien di rumah sakit berdasarkan tes laboratorium darah pasien dibanding algoritma lain yang digunakan.

**Kata Kunci** : Rumah Sakit, Prediksi, Jenis Perawatan, *Ensemble Learning*.

# **PREDICTION OF PATIENTS TREATMENT IN THE HOSPITAL USING THE ENSEMBLE LEARNING METHOD BASED ON THE RESULTS OF PATIENT LABORATORY TESTS**

**ADRIAN DWI ADINATA**

## **ABSTRACT**

Types of treatment at the hospital can be divided into inpatient and outpatient care. As part of a hospital automation system, a classification model is needed to predict the type of patient care to simplify and speed up decision making. The dataset used is taken from Kaggle which refers to Electronic Health Record Predicting collected from private hospitals in Indonesia. The dataset contains the results of the patient's blood laboratory tests. This study aims to build a machine learning model to predict whether patients should be classified under inpatient or outpatient treatment types. The method used in this study is Ensemble Learning, which combines logistic regression algorithms, k-nearest neighbors, support vector machines, decision trees, and naïve bayes. The modeling that will be applied uses Ensemble Vote and Ensemble Stacking to get the prediction model with the best accuracy. The classification model obtained will be evaluated using accuracy, precision, recall, and F1-Score values. The best accuracy results are obtained when using the Ensemble Stacking algorithm with an accuracy of 0.786. It can be concluded that applying the Ensemble Stacking Algorithm is proven to increase the accuracy for predicting the type of patient care in a hospital based on laboratory tests of the patient's blood compared to other algorithms used.

**Keywords** : Hospitals, Predictions, Types of Treatment, Ensemble Learning

## KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas limpahan berkat, rahmat, dan karunia-Nya sehingga penulis dapat menyelesaikan Skripsi berjudul “Prediksi Jenis Perawatan Pada Pasien Di Rumah Sakit Menggunakan Metode *Ensemble Learning* Berdasarkan Hasil Tes Laboratorium Pasien” dengan baik dan lancar.

Skripsi ini disusun untuk memperoleh gelar Sarjana pada program studi Informatika Universitas Pembangunan Nasional Veteran Jakarta. Selama proses penelitian dan penyusunan Skripsi ini penulis banyak mendapat bimbingan, arahan, masukan, dan dukungan dari berbagai pihak. Untuk itu, pada kesempatan ini penulis mengucapkan terimakasih setulus-tulusnya kepada yang terhormat:

1. Bapak Dr. Anter Venus, MA, Comm, selaku Rektor Universitas Pembangunan Nasional Veteran Jakarta.
2. Ibu Dr. Ermatita, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
3. Ibu Dr. Widya Cholil, M.I.T, selaku kepala Program Studi S1 Informatika Universitas Pembangunan Nasional Veteran Jakarta.
4. Ibu Neny Rosmawarni, S.Kom., M.Kom, selaku dosen pembimbing akademik yang telah memberikan banyak saran dan arahan selama masa perkuliahan.
5. Bapak Dr. Didit Widiyanto, S.KOM., M.Si, selaku dosen pembimbing Skripsi yang telah memberikan dukungan, arahan, dan bimbingan selama penelitian dan penyusunan laporan akhir.
6. Bapak dan Ibu dosen serta staff Program Studi S1 Informatika Universitas Pembangunan Nasional Veteran Jakarta yang telah memberikan ilmu dan arahan selama masa perkuliahan.
7. Kedua orang tua, kakak, dan seluruh keluarga yang telah memberikan dukungan, semangat, dan doa.



8. Teman-teman S1 Informatika Angkatan 2019 yang selalu memberikan semangat dan dukungan.
9. Semua pihak yang tidak dapat disebutkan satu persatu, yang telah membantu pelaksanaan penelitian dan penyusunan Skripsi ini.

Meski demikian, penulis menyadari bahwa dalam penulisan Skripsi ini masih terdapat banyak kekurangan, baik dari segi penyusunan, bahasa, maupun penulisannya. Oleh karena itu, penulis secara terbuka menerima segala masukan baik saran maupun kritik demi perbaikan Skripsi ini dan dapat menjadi acuan agar penulis bisa menjadi lebih baik lagi di masa mendatang.

Demikian yang dapat penulis sampaikan. Semoga Skripsi ini dapat bermanfaat bagi semua pihak dan dapat menjadi referensi yang baik bagi pembaca khususnya dalam bidang informasi dan teknologi.

Jakarta, 5 Juni 2023

Penulis

## DAFTAR ISI

HALAMAN JUDUL.....	ii
PERNYATAAN ORISINALITAS .....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS .....	iv
LEMBAR PENGESAHAN .....	v
ABSTRAK .....	vi
<i>ABSTRACT</i> .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR .....	xv
DAFTAR LAMPIRAN.....	xvii
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Ruang Lingkup.....	4
1.6 Luaran Yang Diharapkan .....	4
1.7 Sistematika Penulisan.....	4
BAB 2 TINJAUAN PUSTAKA .....	6
2.1 Darah .....	6
2.2 Prediksi.....	7

2.3	Klasifikasi.....	8
2.4	<i>Machine learning</i> .....	8
2.5	<i>Supervised learning</i> .....	9
2.6	<i>Ensemble Learning</i> .....	9
2.7	Algoritma <i>Logistic Regression</i> .....	10
2.8	Algoritma <i>Decision Tree</i> .....	11
2.9	Algoritma <i>Support Vector Machine</i> .....	11
2.10	Algoritma <i>K-Nearest Neighbors</i> .....	12
2.11	Algoritma <i>Naïve Bayes</i> .....	13
2.12	<i>Ensemble Stacking</i> .....	13
2.13	<i>Ensemble Vote</i> .....	14
2.14	<i>Preprocessing</i> .....	14
2.15	<i>Encoding</i> .....	16
2.16	<i>Splitting Data</i> .....	16
2.17	Visualisasi Data.....	17
2.18	<i>Scikit-Learn</i> .....	17
2.19	<i>Confusion Matrix</i> .....	18
2.20	<i>Classification Report</i> .....	19
2.21	<i>Correlation Matrix (Pearson)</i> .....	19
2.22	<i>Hyperparameter</i> .....	20
2.23	Normalitas Data.....	21
2.24	Review Penelitian Terdahulu .....	21
BAB 3 METODOLOGI PENELITIAN.....		26
3.1	Sumber Data.....	26
3.2	Tahapan Penelitian .....	27
3.3	Desain Eksperimen.....	28
3.3.1	Identifikasi Masalah.....	30
3.3.2	Studi Literatur .....	30
3.3.3	Akuisisi Data .....	30

3.3.4	<i>Preprocessing Data</i> .....	30
3.3.5	<i>Exploratory Data Analysis</i> .....	31
3.3.6	<i>Training dan Testing</i> .....	31
3.3.7	Pemodelan .....	31
3.3.8	Evaluasi .....	32
3.3.9	Model Inference .....	32
3.3.10	Kesimpulan .....	32
3.4	Waktu dan Tempat Penelitian .....	33
3.5	Instrumen Penelitian.....	33
3.6	Jadwal Rencana Penelitian .....	34
BAB 4 HASIL DAN PEMBAHASAN.....		35
4.1	<i>Data Preprocessing</i> .....	35
4.2	Visualisasi Data .....	36
4.2.1	Perbandingan komposisi data antar jenis kelamin pasien berdasarkan jumlah kadar hemoglobin .....	36
4.2.2	Perbandingan Komposisi data antar Status Penanganan berdasarkan jumlah kadar hematocrit.....	37
4.2.3	Perbandingan data antar status final berdasarkan jumlah kadar kandungan sel darah putih (leukosit) .....	38
4.2.4	Perbandingan data antar status final berdasarkan jumlah kadar kandungan sel darah putih (leukosit) dan jenis kelamin .....	39
4.2.5	Perbandingan data antar status final berdasarkan jumlah kadar kandungan sel darah merah (eritrosit) .....	40
4.2.6	Korelasi/Keterkaitan Data di Masing-Masing Kolom .....	41
4.2.7	Pola Hubungan Korelasi Tertinggi .....	43
4.2.8	Pola Hubungan Korelasi Terendah .....	44

4.3	<i>Data Splitting</i> .....	45
4.4	Normalisasi Data .....	47
4.5	<i>Data Modeling</i> Menggunakan Algoritma Dasar .....	49
	4.5.1 Hasil <i>Tuning Hyperparameter</i> Algoritma Dasar .....	49
	4.5.2 Hasil Pengujian Prediksi Menggunakan Algoritma Dasar .....	52
4.6	<i>Data Modelling</i> Menggunakan <i>Ensemble Learning</i> .....	53
	4.6.1 Hasil <i>Tuning Hyperparameter</i> Algoritma <i>Ensemble Learning</i> ..	53
	4.6.2 Hasil Pengujian Prediksi Menggunakan Algoritma <i>Ensemble Learning</i> .....	55
4.7	Perbandingan Akurasi .....	56
	4.7.1 Perbandingan Akurasi Algoritma Dasar .....	56
	4.7.2 Perbandingan Akurasi Algoritma Ensemble .....	57
	4.7.3 Perbandingan Akurasi Algoritma Dasar dan Ensemble .....	57
4.8	<i>Model Inference</i> .....	58
BAB V PENUTUP .....		61
5.1	Kesimpulan .....	61
5.2	Saran .....	61
DAFTAR PUSTAKA .....		63
RIWAYAT HIDUP .....		67
LAMPIRAN .....		69

## DAFTAR TABEL

Tabel 2. 1 Nilai Acuan Kadar Normal .....	7
Tabel 2. 2 Confusion Matrix .....	18
Tabel 2. 3 Ringksaan Penelitian Terdahulu .....	23
Tabel 3. 1 Variabel Data Penelitian .....	26
Tabel 3. 2 Kelompok Umur .....	27
Tabel 3. 3 Jadwal Rencana Penelitian .....	34
Tabel 4. 1 Data Asli sebelum dilakukan tahapan Preprocessing .....	35
Tabel 4. 2 Data setelah dilakukan tahapan Preprocessing .....	36
Tabel 4. 3 Kumpulan Data Training sebelum Normalisasi .....	47
Tabel 4. 4 Kumpulan Data Testing sebelum Normalisasi .....	48
Tabel 4. 5 Kumpulan Data Training setelah Normalisasi .....	48
Tabel 4. 6 Kumpulan Data Testing setelah Normalisasi .....	49
Tabel 4. 7 Hasil Tuning Parameter Algoritma Dasar .....	49
Tabel 4. 8 Hasil Pengujian Prediksi Menggunakan Algoritma Dasar .....	53
Tabel 4. 9 Hasil Tuning Parameter Algoritma Ensemble Learning .....	54
Tabel 4. 10 Hasil Pengujian Prediksi Menggunakan Algoritma Ensemble .....	56
Tabel 4. 11 Hasil Percobaan dengan 10 data .....	59

## DAFTAR GAMBAR

Gambar 2. 1 Fungsi Algoritma Logistic Regression.....	10
Gambar 2. 2 Rumus Gain .....	11
Gambar 2. 3 Rumus Entropy .....	11
Gambar 2. 4 Hyperplane pada SVM .....	12
Gambar 2. 5 Rumus Algoritma K-Nearest Neighbors .....	12
Gambar 2. 6 Rumus Algoritma Naïve Bayes .....	13
Gambar 2. 7 Metode Stacking.....	14
Gambar 2. 8 Proses Preprocessing .....	15
Gambar 2. 9 Rumus Perhitungan Akurasi.....	19
Gambar 2. 10 Rumus Perhitungan Recall .....	19
Gambar 2. 11 Rumus Perhitungan Precision .....	19
Gambar 2. 12 Contoh <i>Correlation Matrix</i> with <i>Heatmap</i> .....	20
Gambar 2. 13 Rumus Pearson Correlation.....	20
Gambar 2. 14 Rumus StandardScaler .....	21
Gambar 3. 1 Tahapan Penelitian .....	28
Gambar 3. 2 Desain Eksperimen.....	29
Gambar 4. 1 Nilai Persentase dari Jumlah Total Hemoglobin per masing-masing Jenis Kelamin Pasien.....	37
Gambar 4. 2 Rata-Rata Persentase terhadap Data Hematocrit per Status Penanganan Pasien .....	38
Gambar 4. 3 Perbandingan Status Final Decision untuk Pasien berdasarkan Jumlah Leukosit.....	39
Gambar 4. 4 Perbandingan Jumlah Kandungan Leukosit berdasarkan Status Penanganan .....	40
Gambar 4. 5 Perbandingan Data antar Jenis Kelamin Pasien berdasarkan Eritrosit dan Status Penanganan.....	41
Gambar 4. 6 Korelasi atau Keterkaitan Data pada masing-masing kolom .....	42

Gambar 4. 7 Hubungan Korelasi antara Kolom Mch dan Mcv berdasarkan Kategori Status Perawatan .....	43
Gambar 4. 8 Hubungan Korelasi antara Kolom Mcv dan Erythrocyte berdasarkan Perbedaan Kategori Status Perawatan.....	44
Gambar 4. 9 Perbandingan Tingkat Akurasi Algoritma Dasar .....	56
Gambar 4. 10 Perbandingan Tingkat Akurasi Algoritma Ensemble.....	57
Gambar 4. 11 Perbandingan Tingkat Akurasi Algoritma Dasar dan Ensemble .....	58
Gambar 4. 12 Hasil Percobaan dengan Output Rawat Jalan.....	59
Gambar 4. 13 Hasil Percobaan dengan Output Rawat Inap .....	59



## DAFTAR LAMPIRAN

Lampiran 1. Dataset .....	69
Lampiran 2. Source Code .....	270
Lampiran 3. Grafik Confusion Matrix Algoritma Logistic Regression .....	271
Lampiran 4. Grafik Confusion Matrix Algoritma Decision Tree .....	272
Lampiran 5. Grafik Confusion Matrix Algoritma Support Vector Machine (SVM) .....	273
Lampiran 6. Grafik Confusion Matrix Algoritma K-Naïve Bayes .....	274
Lampiran 7. Grafik Confusion Matrix Algoritma Ensemble Stacking .....	275
Lampiran 8. Grafik Confusion Matrix Algoritma Ensemble Voting .....	276
Lampiran 9. Hasil Turnitin .....	277