

**TESIS**  
**PREDIKSI TINGKAT RISIKO KESEHATAN PADA JEMAAH**  
**HAJI DENGAN METODE RANDOM FOREST DAN**  
**EXTREME GRADIENT BOOSTING MENGGUNAKAN**  
**OPTIMASI BAYESIAN**  
**(STUDI KASUS : JEMAAH HAJI EMBARKASI BALIKPAPAN)**



disusun oleh:

**NAMA** : LUTHFI BHAKTIAWAN HUSAG  
**NIM** : 24.55.1576  
**KONSENTRASI** : Digital Transformation Intelligence

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

2026

**TESIS**

**PREDIKSI TINGKAT RISIKO KESEHATAN PADA JEMAAH  
HAJI DENGAN METODE RANDOM FOREST DAN  
EXTREME GRADIENT BOOSTING MENGGUNAKAN  
OPTIMASI BAYESIAN**

**(STUDI KASUS : JEMAAH HAJI EMBARKASI BALIKPAPAN)**

***PREDICTION OF HEALTH RISK LEVELS IN HAJJ PILGRIMS  
USING RANDOM FOREST AND EXTREME GRADIENT  
BOOSTING METHODS WITH BAYESIAN OPTIMIZATION  
(CASE STUDY: HAJJ PILGRIMS FROM BALIKPAPAN EMBARKATION)***

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana

Program Studi S2 PJJ INFORMATIKA



Disusun oleh:

**NAMA : LUTHFI BHAKTIWAN HUSAG**  
**NIM : 24.55.1576**  
**KONSENTRASI : Digital Transformation Intelligence**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

2026

**HALAMAN PERSETUJUAN**

**PREDIKSI TINGKAT RISIKO KESEHATAN PADA JEMAAH HAJI  
DENGAN METODE RANDOM FOREST DAN EXTREME GRADIENT  
BOOSTING MENGGUNAKAN OPTIMASI BAYESIAN  
(STUDI KASUS : JEMAAH HAJI EMBARKASI BALIKPAPAN)**

*PREDICTION OF HEALTH RISK LEVELS IN HAJJ PILGRIMS USING  
RANDOM FOREST AND EXTREME GRADIENT BOOSTING METHODS  
WITH BAYESIAN OPTIMIZATION  
(CASE STUDY: HAJJ PILGRIMS FROM BALIKPAPAN EMBARKATION)*

yang disusun dan diajukan oleh

**Luthfi Bhaktiawan Husag**

24.55.1576

telah disetujui oleh Dosen Pembimbing Tesis  
pada tanggal 02 Februari 2026

**Dosen Pembimbing,**



**Prof. Dr. Kusriani, M.Kom.**

**NIK. 190302106**

HALAMAN PENGESAHAN

PREDIKSI TINGKAT RISIKO KESEHATAN PADA JEMAAH HAJI  
DENGAN METODE RANDOM FOREST DAN EXTREME GRADIENT  
BOOSTING MENGGUNAKAN OPTIMASI BAYESIAN  
(STUDI KASUS : JEMAAH HAJI EMBARKASI BALIKPAPAN)

*PREDICTION OF HEALTH RISK LEVELS IN HAJJ PILGRIMS USING  
RANDOM FOREST AND EXTREME GRADIENT BOOSTING METHODS  
WITH BAYESIAN OPTIMIZATION  
(CASE STUDY: HAJJ PILGRIMS FROM BALIKPAPAN EMBARKATION)*

yang disusun dan diajukan oleh

**Luthfi Bhaktiawan Husag**

24.55.1576

Telah dipertahankan di depan Dewan Penguji

pada tanggal 02 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Hanafi, S.Kom., M.Eng., Ph.D  
NIK. 190302024

Dr. Sri Ngudi Wahyuni, ST., M.Kom  
NIK. 190302060

Prof. Dr. Kusriani, M.Kom  
NIK. 190302106

Tanda Tangan



Tesis ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Magister Komputer  
Tanggal 12 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.  
NIK. 190302106

## HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Luthfi Bhaktiawan Husag  
NIM : 24.55.1576

Menyatakan bahwa Tesis dengan judul berikut:

**PREDIKSI TINGKAT RISIKO KESEHATAN PADA JEMAAH HAJI  
DENGAN METODE RANDOM FOREST DAN EXTREME GRADIENT  
BOOSTING MENGGUNAKAN OPTIMASI BAYESIAN  
(STUDI KASUS : JEMAAH HAJI EMBARKASI BALIKPAPAN)**

Dosen Pembimbing : Prof. Dr. Kusriani, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUMPERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian 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 AMKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran 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, 02 Februari 2026

Yang Menyatakan,

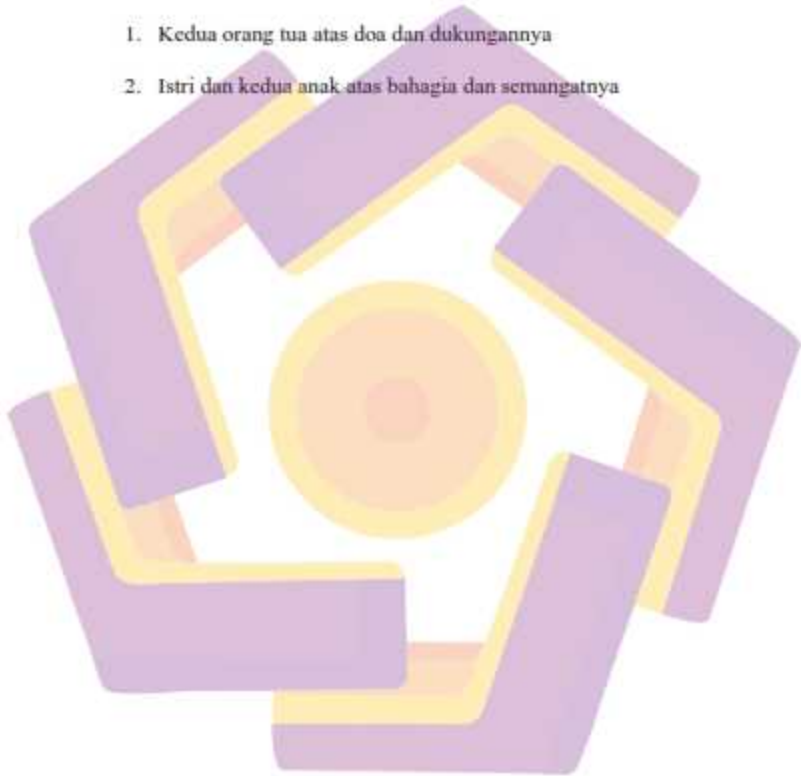


Luthfi Bhaktiawan Husag

## HALAMAN PERSEMBAHAN

Alhamdulillah, segala puji bagi Allah SWT yang telah melimpahkan rahmat serta hidayahNya sehingga tesis ini dapat terselesaikan. Tesis ini penulis persembahkan untuk :

1. Kedua orang tua atas doa dan dukungannya
2. Istri dan kedua anak atas bahagia dan semangatnya



## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya sehingga Tesis yang berjudul *Prediksi Tingkat Risiko Kesehatan Pada Jemaah Haji Dengan Metode Random Forest Dan Extreme Gradient Boosting Menggunakan Optimasi Bayesian (Studi Kasus : Jemaah Haji Embarkasi Balikpapan)* dapat diselesaikan dengan baik sebagai salah satu syarat untuk memperoleh gelar Magister pada Universitas Amikom Yogyakarta.

Penyusunan Tesis ini tidak terlepas dari bimbingan, arahan, dan bantuan dari berbagai pihak. Penulis menyadari bahwa tanpa dukungan tersebut, penyelesaian Tesis ini tidak akan berjalan lancar. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang tulus dan penghargaan setinggi-tingginya kepada:

1. Bapak H. Abdul Khaliq sebagai Kepala Kantor Kementerian Agama Provinsi Kalimantan Timur yang telah memberikan izin untuk melanjutkan studi.
2. Prof. Dr. Kusriani, M.Kom yang telah memberikan bimbingan, arahan, dan motivasi dengan penuh kesabaran sejak awal hingga selesainya tesis ini.
3. Ayahanda Husni Anwar dan Ibunda Agustinawati, serta seluruh keluarga tercinta yang senantiasa memberikan doa, dukungan moral, dan materiel yang tak terhingga.
4. Istri tercinta Oktaviana Tri Hapsani, anak pertama saya Baarakallah Muhammad Husag dan anak kedua saya Hayyindya Medina Husag yang selalu mendukung dan mencerahkan hari.

Penulis menyadari bahwa Tesis ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun dari pembaca akan diterima dengan terbuka. Akhir kata, semoga Tesis ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan pihak-pihak yang membutuhkan.

Yogyakarta , 02 Februari 2026

Luthfi Bhaktiawan Husag



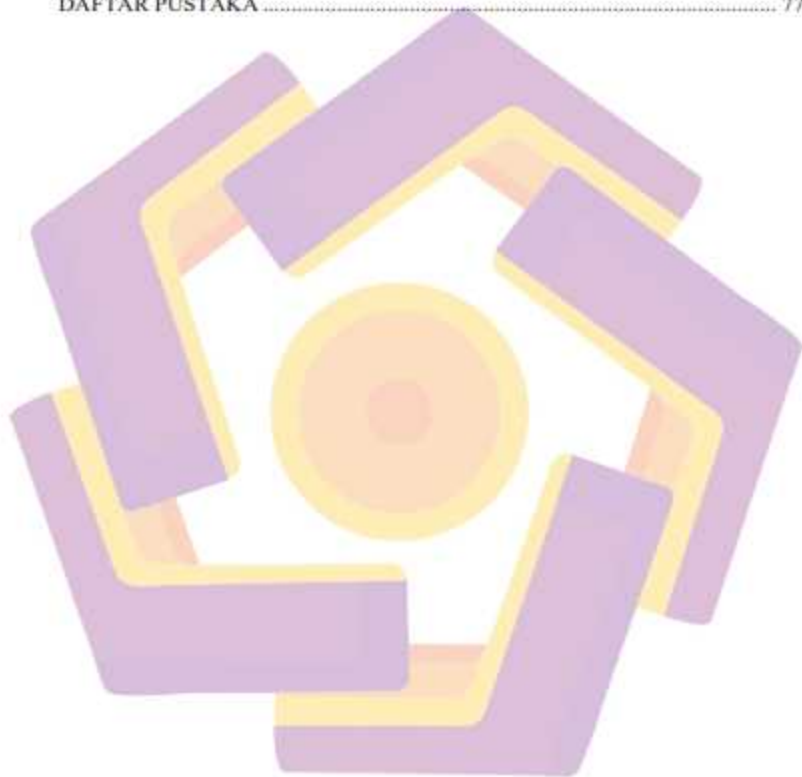
## DAFTAR ISI

COVER .....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN TESIS .....	v
HALAMAN PERSEMBAHAN .....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
INTISARI.....	xv
<i>ABSTRACT</i> .....	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah .....	1
1.2. Rumusan Masalah.....	5
1.3. Batasan Masalah .....	5
1.4. Tujuan Penelitian.....	6
1.5. Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1. Tinjauan Pustaka.....	8
2.2. Keaslian Penelitian.....	13
2.3. Landasan Teori.....	27

2.3.1 Ensemble Learning.....	27
2.3.2 Random forest (RF).....	27
2.3.3 Extreme Gradient Boosting (XGBoost).....	27
2.3.4 Bayesian Optimization.....	28
2.3.5 Hyperparameter tuning.....	29
2.3.6 Accuracy (Akurasi).....	29
2.3.7 F1-Score (Macro).....	29
2.3.8 AUC - ROC (Area Under the Receiver Operating Characteristic Curve).....	30
<b>BAB III METODE PENELITIAN.....</b>	<b>31</b>
3.1 Jenis, Sifat, dan Pendekatan Penelitian.....	31
3.2 Metode Pengumpulan Data.....	32
3.3 Metode Analisis Data.....	32
1. Pengumpulan Dataset.....	32
2. Pra-pemrosesan Data.....	33
3.4 Alur Penelitian.....	38
1. Identifikasi Masalah.....	38
2. Analisis dan Perancangan.....	39
3. Hasil dan Kesimpulan.....	40
<b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....</b>	<b>42</b>
4.1. Pengumpulan Data.....	42

1.	Data Fitur (Variabel Independen).....	44
2.	Data Label (Variabel Dependen).....	45
4.2	Skenario Penelitian.....	47
1.	Algoritma Klasifikasi.....	47
2.	Optimasi Hiperparameter.....	47
3.	Validasi Model.....	47
4.	Stratified K-Fold Cross-Validation.....	48
4.3	Pre-processing Data.....	50
1.	Label Encoding.....	50
2.	Standard Scaling.....	51
4.4	Hasil dan Evaluasi Model.....	53
	<b>4.4.1 Model 1 XGBoost dengan Bayesian Optimization.....</b>	<b>53</b>
	<b>4.4.2 Model 2 Random Forest dengan Bayesian Optimization.....</b>	<b>59</b>
4.5	Pembahasan.....	66
	4.5.1 Perbandingan Performa Model.....	66
	4.5.2 Analisis Efisiensi Komputasi (Computation Cost).....	67
	4.5.3 Implikasi Penelitian.....	68
	4.5.4 Keterbatasan dan Arah Penelitian Selanjutnya.....	69
4.6	Perbandingan dengan Penelitian Terdahulu.....	69
BAB V PENUTUP.....		73
5.1	Kesimpulan.....	73

5.2 Saran .....	74
5.2.1 Saran Akademis (Pengembangan Penelitian Lanjutan) .....	74
5.2.2 Saran Praktis (Implementasi Operasional) .....	76
DAFTAR PUSTAKA .....	77



## DAFTAR TABEL

Tabel 2.1 Matriks literatur review dan posisi penelitian.....	13
Tabel 4.1 Data Fitur .....	44
Tabel 4.2 Data Fitur (lanjutan).....	45
Tabel 4.3 Data Label .....	46
Tabel 4.5 Hasil setelah Label Encoding.....	51
Tabel 4.6 hasil setelah dilakukan scaler pada data umur .....	52
Tabel 4.8 Search Space XGBoost.....	54
Tabel 4.9 Rata-rata 10-Fold CV XGBoost + BO.....	55
Tabel 4.10 Rata-rata 10-Fold CV XGBoost + BO (Lanjutan).....	56
Tabel 4.11 Search Space Random Forest.....	60
Tabel 4.12 Rata-rata 10-Fold CV RF + BO .....	61
Tabel 4.13 Tabel Perbandingan metrik rata-rata kedua skenario.....	66

## DAFTAR GAMBAR

Gambar 3.1 Alur Model Yang Diusulkan.....	37
Gambar 3.2 Alur Penelitian.....	38
Gambar 4.1 Alur Skema Pengumpulan Data.....	42
Gambar 4.2 Distribusi data label.....	46
Gambar 4.3 Proses Iteratif Bayesian Optimization XGBoost (Cumulative Best F1-Macro Score).....	55
Gambar 4.4 Hasil Metrik Evaluasi pada Setiap Fold Stratified K-Fold CV XGBoost.....	56
Gambar 4.5 Confusion Matrix XGBoost (Single Label, Prediksi vs Aktual).....	58
Gambar 4.6 Distribusi Hasil Prediksi Model 1.....	59
Gambar 4.7 Proses Iteratif Bayesian Optimization (Cumulative Best F1-Macro Score Random Forest).....	61
Gambar 4.8 Hasil Metrik Evaluasi pada Setiap Fold Stratified K-Fold CV Random Forest.....	62
Gambar 4.9 Confusion Matrix Random Forest (Single Label, Prediksi vs Aktual).....	64
Gambar 4.10 Distribusi Hasil Prediksi Model 2.....	65

## INTISARI

Kondisi kesehatan fisik jemaah haji merupakan aspek vital dalam kelancaran ibadah, terutama mengingat tingginya kerentanan pada kelompok lansia dan jemaah dengan komorbiditas, sehingga diperlukan metode deteksi dini yang presisi untuk memitigasi risiko medis. Penelitian ini bertujuan untuk membangun model prediksi tingkat risiko kesehatan jemaah haji dengan menerapkan algoritma ensemble learning, yakni Random Forest dan Extreme Gradient Boosting (XGBoost), yang kinerjanya ditingkatkan melalui penalaan hyperparameter otomatis menggunakan Bayesian Optimization. Berbasis data pemeriksaan kesehatan 6.037 jemaah haji Embarkasi Balikpapan tahun 2024 yang meliputi atribut demografis dan riwayat penyakit (ICD-10), penelitian ini melakukan evaluasi model menggunakan metode Stratified 10-Fold Cross-Validation untuk menangani ketidakseimbangan data secara valid. Hasil eksperimen menunjukkan bahwa integrasi Bayesian Optimization mampu mencapai konvergensi performa optimal secara efisien dalam kurang dari 25 iterasi, di mana model Random Forest berbasis optimasi (BO-RF) menghasilkan kinerja paling seimbang dengan rata-rata Akurasi 89,30% dan F1-Score 89,11%, sementara model XGBoost (BOXGB) mencatat Akurasi 89,15% dengan keunggulan diskriminatif pada nilai AUC 0,9597. Penelitian ini menyimpulkan bahwa pendekatan komputasi yang diusulkan terbukti efektif dalam meningkatkan akurasi prediksi dibandingkan metode konvensional dan layak diimplementasikan sebagai sistem pendukung keputusan strategis bagi tim kesehatan haji dalam menentukan prioritas penanganan medis sebelum keberangkatan.

Kata kunci: Jemaah Haji, Risiko Kesehatan, Random Forest, XGBoost, Bayesian Optimization, Machine Learning.

## ABSTRACT

Physical health condition is a vital aspect of the Hajj pilgrimage, particularly given the high vulnerability among elderly groups and pilgrims with comorbidities, necessitating a precise early detection method to mitigate medical risks. This study aims to build a predictive model for Hajj pilgrim health risk levels by applying ensemble learning algorithms, namely Random Forest and Extreme Gradient Boosting (XGBoost), with performance enhanced through automated hyperparameter tuning using Bayesian Optimization. Based on health examination data from 6,037 pilgrims of the Balikpapan Embarkation in 2024, which includes demographic attributes and disease history (ICD-10), this study evaluated the models using Stratified 10-Fold Cross-Validation to validly handle data imbalance. Experimental results demonstrated that the integration of Bayesian Optimization successfully achieved optimal performance convergence efficiently in fewer than 25 iterations, where the optimized Random Forest model (BO-RF) produced the most balanced performance with an average Accuracy of 89.30% and F1-Score of 89.11%, while the XGBoost model (BOXGB) recorded an Accuracy of 89.15% with superior discriminative capability indicated by an AUC value of 0.9597. This study concludes that the proposed computational approach effectively improves prediction accuracy compared to conventional methods and is viable for implementation as a strategic decision support system for Hajj health teams in determining medical handling priorities prior to departure.

Keywords: Hajj Pilgrims, Health Risk, Random Forest, XGBoost, Bayesian Optimization, Machine Learning.

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Ibadah haji merupakan salah satu ritual keagamaan terbesar di dunia yang melibatkan jutaan jemaah dari berbagai negara. Kondisi fisik dan kesehatan jemaah haji menjadi faktor krusial dalam menjamin kelancaran pelaksanaan ibadah haji. Data Kementerian Kesehatan menunjukkan bahwa risiko kematian jemaah haji cenderung meningkat setiap tahunnya, terutama pada kelompok usia lanjut dan jemaah dengan riwayat penyakit tertentu, seperti hipertensi, diabetes, dan penyakit jantung. Pada tahun 2024 pemerintah kembali mengusung tema "Haji Ramah Lansia". Hal ini dikarenakan jumlah jemaah haji lansia masih mendominasi dalam daftar tunggu. Indonesia pada tahun 2024 mendapat kuota haji sebanyak 221.000 dan mendapat kuota tambahan sebesar 20.000, dari kuota tersebut terdapat lebih dari 40.000 jemaah haji berstatus lansia. Dalam laporan PPIH Bidang Kesehatan Tahun 2024 Embarkasi Balikpapan menunjukkan distribusi 10 Penyakit Risti terbanyak, sebagian besar penyakit risti yang diderita jemaah merupakan penyakit degeneratif dengan jumlah terbanyak yaitu penyakit Hipertensi dan Senility. Kondisi risti yang tergolong penyakit degeneratif seperti Diabetes mellitus (DM), dan hipertensi perlu mendapatkan perhatian dan pembinaan sejak pemeriksaan di daerah. Penyelenggaraan Kesehatan Haji Indonesia khususnya di Embarkasi Haji Balikpapan merupakan pemeriksaan haji tahap ketiga. Pemeriksaan haji tahap

ketiga dilaksanakan dalam rangka menetapkan status kesehatan jamaah haji laik atau tidak laik terbang berdasarkan pemeriksaan kesehatan yang dilakukan pada tahap sebelumnya, serta penilaian kembali terhadap istithaah kesehatan jamaah haji. Penyelenggaraan Kesehatan Haji Indonesia, khususnya di Embarkasi Haji Balikpapan, melibatkan pemeriksaan kesehatan tahap ketiga yang krusial. Tahap ini bertujuan untuk menetapkan status kelayakan terbang jamaah berdasarkan pemeriksaan sebelumnya dan penilaian ulang istithaah kesehatan, sesuai standar keselamatan dan kesehatan penerbangan internasional. Prinsip utamanya adalah memastikan tidak ada jamaah yang mengidap penyakit menular atau potensial wabah, tidak membahayakan penumpang lain dan penerbangan, serta tidak membahayakan diri sendiri. Oleh karena itu, pemeriksaan yang cermat terhadap kondisi yang dapat membahayakan kesehatan dan keselamatan penerbangan, diikuti dengan pembinaan dan pelayanan kesehatan yang tepat, menjadi sangat esensial.

Dalam upaya memitigasi risiko kesehatan ini, pemanfaatan teknologi *machine learning* (ML) telah menunjukkan potensi besar dalam memprediksi tingkat risiko kesehatan. Berbagai penelitian telah mengaplikasikan algoritma ML untuk deteksi dini dan klasifikasi penyakit. Misalnya, penelitian oleh (Alhazmi, 2024) mengembangkan sistem prediksi triase untuk unit gawat darurat selama periode haji menggunakan model ML, yang menunjukkan kinerja lebih baik dibandingkan metode konvensional. (Soelistijadi et al., n.d.) berhasil meningkatkan akurasi klasifikasi penyakit ginjal menggunakan XGBoost yang dioptimalkan, mencapai akurasi 99,16%. Demikian pula,

(Abdurrahman et al., 2022) menunjukkan peningkatan akurasi signifikan hingga 95% pada klasifikasi diabetes menggunakan XGBoost dengan hyperparameter tuning Grid Search dan Random Search.

Meskipun telah banyak penelitian yang menggunakan algoritma ML untuk prediksi kesehatan, masih terdapat tantangan dalam mencapai kinerja deteksi yang optimal. Banyak model ML, termasuk yang berbasis ensemble seringkali menghadapi masalah dalam penentuan *hyperparameter* yang tepat dan rentan terhadap *overfitting* atau *underfitting* jika tidak diatur dengan baik. Beberapa penelitian menunjukkan bahwa akurasi deteksi masih berada di kisaran 80-90%, yang berarti masih ada ruang untuk peningkatan signifikan, terutama untuk aplikasi kritis seperti prediksi risiko kesehatan jemaah haji di mana *false negatives* dapat berakibat fatal. Misalnya, (Kusuma et al., 2025) mencapai akurasi 87% dalam penentuan risiko ibu hamil, dan (Budholiya et al., 2022) mencapai akurasi 91,8% dalam prediksi penyakit jantung. Kinerja ini, meskipun baik, menunjukkan bahwa ada kebutuhan untuk eksplorasi lebih lanjut dalam metode optimasi untuk mendorong batas akurasi dan *robustness model*.

Penelitian ini mengusulkan pengembangan model prediksi tingkat risiko kesehatan pada jemaah haji dengan memanfaatkan kekuatan model algoritma Random Forest dan Extreme Gradient Boosting (XGBoost) yang keduanya dikenal karena kemampuan prediktifnya yang tinggi. Untuk mengatasi tantangan optimasi hyperparameter dan meningkatkan performa model secara lebih sistematis, penelitian ini akan mengintegrasikan *Bayesian Optimization*. *Bayesian Optimization* adalah metode yang efisien untuk mencari

hyperparameter optimal pada fungsi objektif yang mahal untuk dievaluasi, seperti pelatihan model ML, dengan secara cerdas memilih kombinasi hyperparameter berikutnya berdasarkan hasil percobaan yang dilakukan oleh (Frazier, 2018; Garnett, 2023). Pendekatan ini diharapkan dapat menghasilkan model dengan Accuracy, AUC, Precision dan Recall (F1 Score) yang lebih tinggi dan lebih stabil dibandingkan dengan optimasi yang kurang efisien.

Dalam perkembangannya, metode ini mengalami berbagai macam kemajuan dalam memecahkan suatu masalah, pemanfaatan metode optimasi pada hyperparameter dilakukan untuk meningkatkan akurasi dalam memprediksi suatu masalah. Penelitian yang dilakukan oleh (C. Yang et al., 2023) menunjukkan bahwa model BO-RF mengungguli metode lain dalam akurasi dan kemampuan anti-noise untuk estimasi parameter saluran dan penggabungan validasi silang *K-fold* membantu dalam mendapatkan partisi model terbaik dan mengurangi *overfitting*. Penelitian lain (Wang et al., 2021) menunjukkan bahwa optimasi hyperparameter Bayesian secara signifikan meningkatkan kinerja prediktif model Random Forest (RF) dan XGBoost untuk pemetaan kerentanan longsor, dengan nilai AUC masing-masing 0,88 dan 0,86, menunjukkan peningkatan 4% dan 3% dibandingkan model sebelumnya. Penelitian yang dilakukan oleh (Tomita et al., 2023) dalam pengklasifikasi XGBoost menunjukkan kinerja yang unggul dengan akurasi 81% dan AUC 85%, menunjukkan keandalannya dalam mendiagnosis asma.

Penelitian ini bertujuan untuk menguji metode dalam memprediksi tingkat risiko kesehatan pada jemaah haji menggunakan metode random forest dan

extreme gradient boosting menggunakan optimasi bayesian berdasarkan data jemaah haji embarkasi Balikpapan tahun 2024 sehingga menjadi sebuah deteksi dini dalam mengantisipasi kesehatan pada jemaah haji.

### **1.2. Rumusan Masalah**

Berdasarkan penjelasan pada latar belakang masalah, maka rumusan pada penelitian ini adalah :

- a. Bagaimana pengembangan model Random Forest dan Extreme Gradient Boosting menggunakan optimasi Bayesian dapat meningkatkan performa dalam memprediksi tingkat risiko kesehatan pada jemaah haji Embarkasi Balikpapan berdasarkan data tahun keberangkatan 2024?
- b. Seberapa efektif peningkatan performa model Extreme Gradient Boosting (BOXGB) dan Random Forest (BO-RF) yang diusulkan dalam mengklasifikasikan tingkat risiko kesehatan jemaah haji?

### **1.3. Batasan Masalah**

Batasan masalah dalam penelitian ini adalah sebagai berikut :

- a. Penelitian ini dibatasi pada ruang lingkup Jemaah haji di Embarkasi Haji Balikpapan.
- b. Dataset yang digunakan adalah hasil pemeriksaan kesehatan jemaah haji pada tahun keberangkatan 2024, yang mungkin tidak mencerminkan dinamika risiko kesehatan pada tahun-tahun lain atau perubahan tren kesehatan jemaah di masa depan.
- c. Penggunaan metode Random Forest dan Extreme Gradient Boosting dengan optimasi Bayesian dibatasi pada kemampuan dan karakteristik intrinsik dari

algoritma tersebut, serta efektivitas Bayesian Optimization dalam mengeksplorasi ruang hyperparameter yang kompleks. Model ini tidak mencakup integrasi dengan metode ensemble atau optimasi lain di luar cakupan yang telah ditentukan.

- d. Hasil penelitian adalah pengujian pengembangan model untuk prediksi tingkat risiko kesehatan pada jemaah haji menggunakan optimasi Extreme Gradient Boosting berbasis Bayesian Optimization (BOXGB) dan Random Forest berbasis Bayesian Optimization (BO-RF).

#### **1.4. Tujuan Penelitian**

Tujuan penelitian adalah hal-hal yang diharapkan dapat dicapai dalam penelitian ini. Berikut adalah pemaparan terkait tujuan penelitian:

- a. Mengembangkan model prediktif tingkat risiko kesehatan jemaah haji Embarkasi Balikpapan menggunakan pendekatan ensemble learning (Random Forest dan Extreme Gradient Boosting) yang dioptimasi dengan Bayesian Optimization berdasarkan data tahun keberangkatan 2024.
- b. Mengevaluasi peningkatan performa model yang dikembangkan (BOXGB dan BO-RF) berdasarkan Accuracy, AUC, Precision dan Recall (F1 Score) dalam menganalisis dataset pemeriksaan kesehatan jemaah haji berdasarkan data tahun keberangkatan 2024.

#### **1.5. Manfaat Penelitian**

Penelitian ini diharapkan dapat memberi manfaat dalam meningkatkan pengetahuan ilmiah di bidangnya dan dapat menjadi landasan bagi penelitian lanjutan. Berikut adalah manfaat yang dapat diperoleh dari penelitian ini:

- a. **Kontribusi Ilmiah:** Menjadi kontribusi terhadap pengembangan model machine learning prediktif yang dioptimasi (khususnya Random Forest dan XGBoost dengan Bayesian Optimization) dalam domain kesehatan, memperkaya khazanah ilmu pengetahuan di bidang Digital Transformation Intelligence.
- b. **Peningkatan Kualitas Layanan Kesehatan Haji:** Mengurangi angka kejadian medis darurat dan komplikasi kesehatan yang dialami jemaah haji dengan menerapkan tindakan pencegahan dini yang lebih terarah dan berbasis data.
- c. **Pengambilan Keputusan yang Lebih Akurat:** Mempermudah pengambilan keputusan bagi Kementerian Agama dan petugas kesehatan haji dalam penyusunan strategi mitigasi risiko kesehatan. Dengan data prediksi yang lebih akurat dan berbasis teknologi, alokasi sumber daya dan intervensi kesehatan dapat dilakukan secara lebih efisien dan tepat sasaran.
- d. **Identifikasi Jemaah Berisiko Tinggi:** Membantu petugas kesehatan haji dalam mengidentifikasi jemaah dengan risiko kesehatan tinggi secara proaktif, sehingga mereka dapat diberikan perhatian khusus, pembinaan, dan pelayanan medis yang disesuaikan sebelum dan selama perjalanan haji.
- e. **Dasar untuk Penelitian Lanjutan:** Hasil penelitian ini dapat menjadi landasan dan referensi bagi penelitian-penelitian selanjutnya yang berfokus pada pengembangan model prediktif kesehatan yang lebih kompleks, integrasi dengan teknologi lain, atau studi kasus di embarkasi haji lainnya.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Tinjauan Pustaka**

Dalam penelitian ini, penulis mengumpulkan data dari penelitian-penelitian sebelumnya untuk analisis perbandingan, dengan fokus mengidentifikasi kelebihan dan kekurangan dari penelitian yang ada. Pendekatan ini dimaksudkan untuk membangun dasar dalam memahami keterkaitan antara penelitian terdahulu dan penelitian yang dilakukan oleh penulis saat ini, serta mencegah terjadinya redundansi. Oleh karena itu, penelitian sebelumnya memainkan peran penting dalam penelitian ini. Tinjauan pustaka juga bertujuan untuk menunjukkan bahwa penelitian yang dilakukan oleh penulis memiliki nilai dan makna penting sebagai kontribusi terhadap pemahaman ilmiah. Berikut ini adalah ulasan dari berbagai jurnal terkait penelitian terdahulu yang relevan dengan data dan metodologi yang digunakan oleh penulis sebagai referensi.

Penelitian yang telah dilakukan oleh (Alhazmi, 2024) yang membahas tentang sistem prediksi triase untuk unit gawat darurat selama periode haji menggunakan model pembelajaran mesin ditemukan bahwa prediktor risiko kesehatan berperan penting dalam menganalisis data pribadi dan hasil pemeriksaan kesehatan. Prediktor ini membantu mengidentifikasi kondisi kesehatan jamaah sehingga petugas kesehatan dapat memberikan prioritas kepada mereka yang memerlukan bantuan segera. Sistem prediksi yang dikembangkan ini menggunakan 11.584 data catatan pasien, yang secara efektif mengelompokkan pasien ke tingkat triase yang

sesuai dengan kinerja lebih baik dibandingkan metode konvensional. Model XGBoost dan DNN menunjukkan akurasi tinggi dalam membedakan lima tingkat triase, sehingga mengurangi kejadian *under-triage* dan *over-triage*.

Penelitian lainnya (Soelistijadi et al., n.d.) tentang peningkatan akurasi klasifikasi penyakit ginjal menggunakan metode XGBoost menunjukkan bahwa ketika dikombinasikan dengan teknik hyperparameter, hal tersebut ternyata efektif untuk peningkatan dalam pengklasifikasian Penyakit Ginjal Kronis (Chronic Kidney Disease/CKD) dan Non-CKD, dengan mencapai akurasi sebesar 99,16%, presisi 98,17%, recall 99,16%, dan skor F1 sebesar 99,16%.

Penelitian selanjutnya (Abdurrahman et al., 2022) yang membahas tentang optimasi algoritma klasifikasi xgboost menggunakan hyperparameter *gridsearch* dan *random search* pada klasifikasi penyakit diabetes juga menunjukkan hasil yang baik dalam pengujian metode ini yaitu peningkatan ketika menggunakan hyperparameter tuning melalui *gridsearch* dan *random search*, akurasi meningkat secara signifikan hingga mencapai 95%. Waktu eksekusi berdasarkan penelitian untuk *gridsearch* lebih lama karena metode ini melakukan pencarian menyeluruh untuk kombinasi hyperparameter terbaik, yang menjadi kelemahan dibandingkan dengan *random search*. Penelitian ini menyoroti efektivitas hyperparameter tuning dalam meningkatkan kinerja algoritma XGBoost untuk klasifikasi diabetes.

Penelitian lanjutan (Afiatuddin et al., 2024) menunjukkan bahwa penggunaan SMOTE bersama dengan optimasi parameter melalui *grid search* secara signifikan meningkatkan efektivitas algoritma data mining saat bekerja dengan dataset yang tidak seimbang. Hal ini menekankan pentingnya prapemrosesan data yang tepat dan

penyesuaian parameter dalam meningkatkan akurasi prediksi untuk analisis data medis. Selain itu, temuan ini konsisten dengan penelitian lain yang menyarankan bahwa metode optimasi dan praprosesan yang berhasil dapat mengurangi kesulitan dalam memprediksi penyakit jantung.

Dalam analisis eksperimental yang ditulis oleh (J. Yang & Guan, 2022), peneliti mengevaluasi kinerja XGBoost dibandingkan dengan lima algoritma dasar (baseline). Hasilnya menunjukkan bahwa model yang diusulkan unggul dalam keempat metrik penilaian, dengan mencapai akurasi prediksi sebesar 93,44%. Selain itu, peneliti menganalisis pentingnya fitur-fitur yang diidentifikasi oleh algoritma terpilih, yang memiliki relevansi signifikan untuk prediksi penyakit jantung.

Penelitian tentang model *machine learning* XGBoost yang dikembangkan dalam penelitian (Murdiansyah, 2024) menghasilkan hasil yang baik, dengan tingkat presisi sebesar 94,3%, recall sebesar 96,6%, F-Measure sebesar 95,4%, AUC sebesar 95,4%, dan akurasi sebesar 95,4%. Meskipun hasil ini positif, XGBoost belum mampu melampaui performa model *Stacking* dan RF yang digunakan oleh peneliti lain dalam studi sebelumnya. Dalam penelitian ini, SMOTE digunakan untuk mengatasi ketidakseimbangan data, sementara *Bayesian Optimization* diterapkan untuk penyetelan hyperparameter. Para peneliti menemukan bahwa XGBoost masih memiliki potensi untuk meningkatkan kinerjanya. Meskipun telah menggunakan *Bayesian Optimization* untuk penyesuaian hyperparameter, mereka menunjukkan bahwa peningkatan lebih lanjut

dalam optimasi hyperparameter dapat menghasilkan hasil yang lebih baik, mengingat luasnya ruang pencarian untuk nilai hyperparameter.

Penelitian yang dilakukan (Jaya Kusuma et al., 2025) menerapkan algoritma Bayesian Optimization (BO) pada hyper-parameter model pembelajaran mesin untuk menilai tingkat risiko kehamilan pada wanita hamil. Peneliti memperoleh dan memproses data menggunakan metode normalisasi, yang menghasilkan pembentukan dataset pelatihan dan pengujian. Penggunaan metode Bayesian Optimization (BO) dilakukan pada hyper-parameter model Decision Tree (DT) dan Extreme Gradient Boosting (XGB), model BOXGB ini mencapai tingkat akurasi 87%, sedangkan rata-rata precision, recall, dan F1-score masing-masing tercatat sebesar 88%, 87%, dan 88%. Hasil evaluasi menunjukkan pentingnya nilai iterasi dalam Bayesian Optimization, di mana kinerja XGB meningkat setelah penyesuaian hyper-parameter melalui Bayesian Optimization, dibandingkan dengan model penelitian sebelumnya.

Penelitian lanjutan yang ditulis oleh (Budholiya et al., 2022) juga menunjukkan hasil yang baik, peneliti memperkenalkan alat diagnostik untuk penyakit kardiovaskular yang menggunakan One-Hot Encoding dan XGBoost yang ditingkatkan untuk keperluan klasifikasi. Teknik ensemble berbasis pohon melampaui tiga metode lain yang telah dikembangkan sebelumnya, menunjukkan peningkatan akurasi. Optimasi hyper-parameter dilakukan melalui *Bayesian Optimization*, yang terbukti efektif dalam mengidentifikasi hyper-parameter optimal. Model dalam penelitian ini mencapai peningkatan akurasi sebesar 3,28% dibandingkan dengan model penelitian sebelumnya. Hasil eksperimen

menunjukkan bahwa pendekatan diagnostik ini dapat meningkatkan proses pengambilan keputusan dalam mendiagnosis penyakit jantung.

Penelitian selanjutnya (Gao & Ding, 2020) mengevaluasi efektivitas klasifikasi dari berbagai classifier (LightGBM, GBDT, LR, RF, BPNN, dan DT) dibandingkan dengan XGBoost menggunakan dua dataset yaitu dataset penyakit kanker payudara dan dataset penyakit kardiovaskular. Dalam penelitian ini terdapat 1000 sampel ulangan yang dilakukan untuk memastikan keandalan hasil. Dalam temuannya pada dataset Kanker Payudara menyimpulkan hasil bahwa XGBoost melampaui metode klasifikasi lainnya dalam dataset yang terbatas, meskipun menunjukkan stabilitas yang lebih rendah dibandingkan GBDT. Sedangkan pada hasil dataset Penyakit Kardiovaskular XGBoost menunjukkan hasil klasifikasi yang cukup konsisten dalam dataset yang lebih besar. Kesimpulan dalam penelitian ini XGBoost umumnya menunjukkan kinerja yang kuat di berbagai dataset, meskipun stabilitasnya bervariasi jika dibandingkan dengan classifier lain seperti GBDT.

Penelitian oleh (Tomita et al., 2023) menunjukkan bahwa Pengklasifikasi XGBoost menunjukkan kinerja yang unggul dengan akurasi 81% dan AUC 85%, menunjukkan keandalannya dalam mendiagnosis asma. Penelitian ini menyajikan algoritma diagnostik untuk asma dewasa yang secara efektif memanfaatkan random forest (RF) dan pengklasifikasi eXtreme Gradient Boosting (XGBoost) yang dioptimalkan.

## 2.2. Keastian Penelitian

Peneliti mengembangkan ide baru dan menginovasi yang sudah ada, dengan menganalisis literatur ilmiah untuk menemukan perbedaan pengetahuan. Fokusnya kebaruan dan kontribusi pada pengetahuan dalam penelitian ini ditunjukkan dalam Tabel 2.1.

Tabel 2.1 Matriks literatur review dan posisi penelitian

Prediksi Tingkat Risiko Kesehatan Pada Jemaah Haji Dengan Metode Random Forest Dan Extreme Gradient Boosting Menggunakan Optimasi Bayesian

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	OPTIMASI MODEL EXTREME GRADIENT BOOSTING DALAM UPAYA PENENTUAN TINGKAT RISIKO PADA IBU HAMIL BERBASIS BAYESIAN OPTIMIZATION (BOXGB)	Edi Jaya Kusuma, Ririn Nurmandhani, Lenci Aryani, Ika Pantiawati, Garuh Fajar Shidik  Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK) DOI: 10.25126/jtik.2025129001  Vol. 12, No. 1, Februari 2025, hlm. 111-120	a. Studi ini bertujuan untuk mengoptimalkan hiper-parameter Decision Tree (DT) dan Extreme Gradient Boosting (XGB) menggunakan Bayesian Optimization (BO)  b. Studi ini berusaha untuk meningkatkan prediksi risiko kehamilan berdasarkan data pengukuran klinis	a. Implementasi optimasi Bayesian meningkatkan penyetalan hiper-parameter untuk model pembelajaran mesin yang memprediksi risiko kehamilan;  b. Model Decision Tree dan Extreme Gradient Boosting menunjukkan peningkatan kinerja setelah pengoptimalan Bayesian;  c. Model BOXGB mencapai akurasi 87%, dengan presisi, ingatan,	a. Implementasi Bayesian Optimization (BO) direkomendasikan untuk mengoptimalkan hiper-parameter dalam model pembelajaran mesin untuk prediksi risiko kehamilan  b. Penelitian masa depan harus mengeksplorasi ruang hiper-parameter yang lebih luas untuk meningkatkan kinerja model	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			<p>c. Studi ini membahas tantangan menemukan konfigurasi optimal untuk model pembelajaran mesin untuk meningkatkan akurasi prediksi</p> <p>d. Secara keseluruhan, tujuannya adalah untuk meningkatkan kinerja model pembelajaran mesin dalam memprediksi risiko kehamilan</p>	<p>dan skor F1 rata-rata masing-masing 88%, 87%, dan 88%.</p> <p>d. Secara keseluruhan, pengoptimalan Bayesian secara efektif meningkatkan model pembelajaran mesin untuk memprediksi risiko kehamilan berdasarkan pengukuran klinis;</p> <p>e. Penelitian di masa depan dapat mengeksplorasi teknik pengoptimalan yang lebih efisien dan ruang hiper-parameter yang lebih luas.</p>	<p>c. Teknik optimasi yang efisien dapat meningkatkan pengembangan sistem diagnosis dini untuk risiko kehamilan</p> <p>d. Studi ini menunjukkan bahwa Optimasi Bayesian dapat secara signifikan meningkatkan akurasi model dan kemampuan prediksi</p>	<p>gradient boosting menggunakan optimasi bayesian</p>
2	<p>Analisis Pengaruh Bayesian Optimization Terhadap Kinerja SVM Dalam Prediksi Penyakit Diabetes</p>	<p>Dwi Utami, Fathoni Dwiatmoko, Nuari Artisa Sivi</p> <p>Infotek : Jurnal Informatika dan Teknologi Vol. 8 No. 1, Januari 2025</p>	<p>a. Studi ini bertujuan untuk mengeksplorasi penggunaan Support Vector Machine (SVM) yang dioptimalkan dengan Bayesian Optimization untuk prediksi diabetes dini</p>	<p>a. Optimasi Bayesian efektif dalam meningkatkan kinerja Support Vector Machine (SVM) untuk prediksi diabetes dini</p> <p>b. Studi ini menunjukkan bahwa SVM dioptimalkan dengan</p>	<p>a. Implementasi optimasi Bayesian meningkatkan penyeteralan hiper-parameter untuk model pembelajaran mesin yang memprediksi risiko kehamilan</p>	<p>Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Hal.140-150	<p>b. Studi ini berfokus pada peningkatan kinerja SVM melalui optimasi hyperparameter yang efektif menggunakan metode Bayesian</p> <p>c. Studi ini membandingkan keakuratan SVM dengan dan tanpa Optimasi Bayesian dalam memprediksi diabetes</p>	<p>Bayesian Optimization mencapai akurasi yang lebih tinggi dibandingkan dengan SVM tanpa pengoptimalan</p> <p>c. Model prediktif yang akurat sangat penting untuk mencegah dan mengurangi peningkatan kasus diabetes</p>	<p>b. Model Decision Tree dan Extreme Gradient Boosting menunjukkan peningkatan kinerja setelah pengoptimalan Bayesian</p> <p>c. Model BOXGB mencapai akurasi 87%, dengan presisi, ingatan, dan skor F1 rata-rata masing-masing 88%, 87%, dan 88%</p> <p>d. Secara keseluruhan, pengoptimalan Bayesian secara efektif meningkatkan model pembelajaran mesin untuk memprediksi risiko kehamilan berdasarkan pengukuran klinis</p> <p>e. Studi di masa depan dapat mengeksplorasi teknik pengoptimalan yang lebih efisien dan</p>	<p>embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradient boosting menggunakan optimasi bayesian</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
					ruang hiper-parameter yang lebih luas.	
3	An optimized XGBoost based diagnostic system for effective prediction of heart disease	Kartik Budhniya, Shailendra Kumar Shrivastava, Vivek Sharma <a href="https://doi.org/10.1016/j.ijbi.2020.100113">https://doi.org/10.1016/j.ijbi.2020.100113</a> 1319-1578/© 2020 The Authors. Published by Elsevier B.V. on behalf of King Saud University	<p>a. Studi ini bertujuan untuk mengembangkan sistem diagnostik untuk memprediksi penyakit jantung menggunakan pengklasifikasi XGBoost yang dioptimalkan</p> <p>b. Studi ini menekankan pentingnya penyetelan hyperparameter untuk aplikasi pengklasifikasi yang efektif</p> <p>c. Studi ini mengevaluasi kemandirian model yang diusulkan pada dataset penyakit jantung Cleveland</p> <p>d. Kinerja model dibandingkan dengan pengklasifikasi Random Forest dan</p>	<p>a. Sistem diagnostik yang diusulkan secara efektif memprediksi penyakit jantung menggunakan teknik pengkodean XGBoost dan One-Hot yang dioptimalkan</p> <p>b. Model mencapai akurasi prediksi tinggi 91,8% pada dataset penyakit jantung Cleveland</p> <p>c. Metode yang diusulkan mengungguli model sebelumnya dengan akurasi 3,28%</p> <p>d. Optimasi Bayesian digunakan untuk penyetelan hiper-parameter yang efisien</p> <p>e. Studi ini menunjukkan bahwa metode ini dapat meningkatkan pengambilan keputusan dalam diagnosis penyakit jantung</p>	<p>a. Studi ini membahas tantangan memprediksi penyakit jantung secara efektif menggunakan teknik pembelajaran mesin</p> <p>b. Studi ini menyoroti perlunya pengklasifikasi yang dioptimalkan untuk meningkatkan akurasi diagnostik dalam perawatan Kesehatan</p> <p>c. Studi ini menekankan pentingnya penyetelan hyperparameter untuk kinerja pengklasifikasi</p> <p>d. Studi Ini mengidentifikasi kesenjangan dalam penelitian sebelumnya mengenai penggunaan algoritma berbasis pohon dan pengkodean fitur kategoris</p>	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradient boosting menggunakan optimasi bayesian

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			<p>Extra Tree menggunakan berbagai metrik evaluasi</p> <p>e. Metode yang diusulkan mencapai akurasi prediksi tinggi 91,8%.</p>		<p>e. Model yang diusulkan bertujuan untuk meningkatkan akurasi prediksi dibandingkan dengan metode yang ada</p>	
4	Disease prediction via Bayesian hyperparameter optimization and ensemble learning	Gao and Ding <i>BMC Res Notes</i> (2020) 13:205 <a href="https://doi.org/10.1186/s13104-020-05050-0">https://doi.org/10.1186/s13104-020-05050-0</a>	<p>a. Studi ini bertujuan untuk menganalisis dan membandingkan strategi pembelajaran mesin untuk memprediksi Kanker Payudara (BC) dan Penyakit Kardiovaskular (CVD)</p> <p>b. Studi ini menekankan mengidentifikasi fitur prediktif awal penyakit untuk meningkatkan kelangsungan hidup pasien</p> <p>c. Studi ini menyajikan analisis komparatif komprehensif dari</p>	<p>a. Studi ini menekankan pentingnya skrining dan diagnosis penyakit dini untuk meningkatkan kelangsungan hidup pasien</p> <p>b. Tekanan darah sistolik sangat penting untuk memprediksi penyakit kardiovaskular</p> <p>c. Model XGBoost menunjukkan akurasi dan sensitivitas tinggi untuk diagnosis kanker dada</p> <p>d. Optimasi hyperparameter Bayesian terbukti lebih</p>	<p>a. Penelitian lebih lanjut diperlukan pada metode pemilihan fitur untuk meningkatkan kinerja model</p> <p>b. Menjelajahi data multiclass dapat memberikan wawasan tentang perbandingan model</p>	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradient boosting menggunakan optimasi bayesian

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			sistem pembelajaran mesin yang berbeda.	<p>stabil daripada metode tradisional</p> <p>e. Penelitian ini memberikan analisis komparatif strategi pembelajaran mesin untuk memprediksi kanker dada dan penyakit kardiovaskular</p> <p>f. Temuan ini menyoroti pentingnya peringkat kepentingan fitur dalam mengidentifikasi fitur berisiko tinggi.</p>		
5	Optimizing health data analytics in fog computing using hyperparameter tuning and grid search		<p>a. Studi ini bertujuan untuk menguji integrasi komputasi kabut dengan analisis data kesehatan untuk meningkatkan analisis kesejahteraan pasien</p> <p>b. Studi ini menekankan penggunaan penyetelan hyperparameter dan teknik pencarian grid</p>	<p>a. Penyetelan hyperparameter melalui grid search mengoptimalkan model pembelajaran mesin untuk analitik data kesehatan dalam komputasi kabut, meningkatkan kinerja dalam batasan perangkat tepi</p>	<p>a. Penyetelan hyperparameter menggunakan pencarian grid direkomendasikan untuk mengoptimalkan model pembelajaran mesin dalam komputasi kabut untuk analisis data kesehatan</p>	<p>Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			<p>untuk meningkatkan pendekatan pengoptimalan dalam analisis data kesehatan</p> <p>c. Tujuannya termasuk meningkatkan akurasi, mengurangi latensi, dan meningkatkan efisiensi sumber daya di lingkungan komputasi kabut.</p>	<p>b. Grid search memungkinkan penyesuaian model untuk perangkat tepi yang beragam, memastikan ketahanan terhadap sumber data heterogen dan kemampuan beradaptasi terhadap perubahan kondisi</p>	<p>b. Sangat penting untuk menyesuaikan model dengan tantangan unik perangkat tepi yang terbatas sumber daya</p> <p>c. Integrasi komputasi kabut dengan analitik data kesehatan harus fokus pada peningkatan akurasi dan pengurangan latensi</p> <p>d. Menyesuaikan hyperparameter untuk beragam perangkat edge sangat penting untuk meningkatkan efisiensi sumber daya</p>	<p>metode yang digunakan yaitu metode random forest dan extreme gradient boosting menggunakan optimasi bayesian</p>
6	Comparative Analysis of Logistic Regression, SVM, Xgboost, and Random Forest Algorithms for	Rahmat Hidayat, Deni Mahdiana, Anggun Fergina Jurnal Teknologi Sistem Informasi dan Aplikasi ISSN: 2654-3788 Penerbit: Program Studi Teknik Informatika	<p>a. Studi ini bertujuan untuk menyederhanakan peramalan diabetes menggunakan berbagai metode pembelajaran mesin, termasuk algoritma KNN dan Naive Bayes</p>	<p>a. Regresi logistik mengungguli SVC dalam presisi, akurasi, recall, dan F1 skor, menunjukkan keefektifannya untuk tugas klasifikasi</p> <p>b. Random forest, meskipun dapat</p>	<p>a. Studi lebih lanjut dapat fokus pada peningkatan akurasi algoritma yang ada seperti KNN dan Naive Bayes dalam prediksi diabetes</p> <p>b. Studi juga dapat memeriksa penerapan</p>	<p>Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Diabetes Classification	Universitas Pamulang-e-ISSN: 2654-4229 Vol. 7, No. 1, Januari 2024 (281-291)	<p>b. Studi Ini berusaha untuk mengklasifikasikan diabetes secara efektif menggunakan dataset baru yang belum banyak digunakan</p> <p>c. Studi ini bermaksud untuk memberikan tindakan pencegahan dini terhadap diabetes melalui klasifikasi yang akurat</p> <p>d. Studi Ini mengeksplorasi efektivitas algoritma yang berbeda, termasuk Random Forest, Support Vector Classification, dan XGBoost, dalam klasifikasi diabetes</p> <p>e. Tujuan utamanya adalah untuk meningkatkan keakuratan diagnosis dan pengobatan diabetes</p>	<p>diandalkan, tidak melampaui regresi logistik dalam masalah klasifikasi khusus ini</p> <p>c. Studi ini menyoroti pentingnya menggunakan teknik pembelajaran mesin yang tepat untuk klasifikasi diabetes</p> <p>d. Penelitian ini menggunakan kumpulan data diabetes baru, meningkatkan relevansi temuan</p> <p>e. Secara keseluruhan, regresi logistik direkomendasikan untuk kinerja klasifikasi yang lebih baik dalam diagnosis diabetes</p>	<p>metode pembelajaran mesin lainnya untuk membandingkan kinerjanya dengan metode statistik tradisional</p> <p>c. Menjelajahi penggunaan kumpulan data yang lebih besar dan lebih beragam dapat menghasilkan wawasan yang lebih baik tentang klasifikasi diabetes.</p>	<p>embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradient boosting menggunakan optimasi bayesian</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
7	Prediction and analysis of train arrival delay based on XGBoost and Bayesian optimization	Rui Shi, Xinyue Xu +, Jianmin Li, Yanqiu Li  <a href="https://doi.org/10.1016/j.asoc.2021.107538">https://doi.org/10.1016/j.asoc.2021.107538</a> 1568-4946/© 2021 Elsevier B.V. All rights reserved.	<p>a. Studi ini bertujuan untuk meningkatkan prediksi keterlambatan kedatangan kereta menggunakan metode berbasis data yang menggabungkan XGBoost dan pengoptimalan Bayesian</p> <p>b. Ini mengidentifikasi sebelas karakteristik yang mempengaruhi waktu kedatangan kereta sebagai variabel independen untuk model prediksi</p> <p>c. Studi ini mengevaluasi efisiensi prediksi dan akurasi model yang diusulkan pada dua jalur kereta api berkecepatan tinggi di China</p> <p>d. Tujuannya adalah untuk meningkatkan akurasi prediksi melalui optimasi</p>	<p>a. Metode yang diusulkan secara efektif memprediksi pemundatan kedatangan kereta, mengungguli model tolak ukur dalam metrik akurasi</p> <p>b. Model mencapai koefisien penentuan 0,9889/0,9905 dan tingkat kesalahan rendah</p> <p>c. Kebenaran perkiraan melebihi 96% di sebagian besar stasiun ketika kesalahan yang diijinkan dalam waktu 3 menit</p> <p>d. Metode ini menunjukkan karakteristik generalisasi yang baik dan kebaikan yang diinginkan</p> <p>e. Hasil menunjukkan bahwa model dapat secara akurat memperkirakan pemundatan yang</p>	<p>a. Penelitian masa depan dapat fokus pada pengumpulan data sampel yang cukup tentang gangguan kereta api untuk meningkatkan model prediksi pemundatan</p> <p>b. Menyelidiki dampak dari berbagai jenis gangguan pada pemundatan kereta dapat menghasilkan wawasan yang berharga</p> <p>c. Studi lebih lanjut dapat mengeksplorasi optimalisasi hiperparameter dalam model XGBoost untuk meningkatkan akurasi</p> <p>d. Penelitian juga dapat menilai kinerja metode yang diusulkan di berbagai peristiwa abnormal</p> <p>e. Memperluas analisis untuk memasukkan</p>	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi pada Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradien boosting menggunakan optimasi bayesian.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			hyperparameter model XGBoost	disebabkan oleh peristiwa abnormal tertentu	lebih banyak jalur kereta api berkecepatan tinggi dapat memvalidasi penerapan model	
8	Optimasi Algoritma XGBoost Classifier Menggunakan Hyperparameter Gridsearch dan Random Search Pada Klasifikasi Penyakit Diabetes	<p>Gunanjar Abdurrahman, Hardian Oktavianto, Mukti Sintawati.</p> <p><i>Informatics Journal</i> Vol. 7 No. 3 (2022)</p>	<p>a. Studi ini bertujuan untuk mengklasifikasikan apakah seseorang menderita diabetes atau tidak menggunakan XGBoost Classifier</p> <p>b. Studi ini menggunakan dataset diabetes dari repositori Pembelajaran Mesin UCI</p> <p>c. Studi ini menyelidiki dampak penyetelan hyperparameter pada keakuratan algoritma XGBoost</p> <p>d. Studi ini juga membahas penanganan nilai yang hilang dalam</p>	<p>a. Klasifikasi menggunakan XGBoost tanpa penyetelan hyperparameter mencapai akurasi 75%</p> <p>b. Dengan penyetelan hyperparameter, pencarian kisi dan pencarian acak menghasilkan akurasi 95%</p> <p>c. Waktu eksekusi untuk gridsearch lebih lama karena sifat pencariannya yang lengkap</p> <p>d. Studi ini secara efektif menangani nilai yang hilang dalam kumpulan data, yang sangat penting untuk kinerja XGBoost</p>	<p>a. Makalah ini menyebutkan bahwa konfigurasi yang diamati terbaik mungkin bukan yang terbaik rata-rata karena ketidakpastian dalam pemilihan hyperparameter</p> <p>b. Studi ini menyoroti bahwa benih acak yang berbeda dapat menyebabkan pengamatan yang bervariasi, yang mempengaruhi koandalan konfigurasi terbaik</p> <p>c. Keterbatasan metode penyetelan hyperparameter seperti gridsearch dan</p>	<p>Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradient boosting menggunakan optimasi bayesian</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			kumpulan data sebelum klasifikasi	e. Penelitian menunjukkan bahwa penyetelan hyperparameter secara signifikan meningkatkan akurasi model	penelitian acak tidak secara eksplisit dibahas secara rinci d. Makalah ini tidak membahas potensi masalah overfitting yang mungkin timbul dari penyetelan hyperparameter	
9	Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang	Putri Septiana Rizky, Risti Haiban Hirzi, Umam Hidayatullohman  J Statistika Vol. 15, No. 2, (2022), Hal. 228 - 236	a. Studi ini bertujuan untuk mengatasi tantangan ketidakseimbangan kelas dalam algoritma klasifikasi b. Studi ini mengaulkan pendekatan ansambel menggunakan LightGBM dan XGBoost untuk kinerja yang lebih baik c. Studi ini mengevaluasi efektivitas metode ini pada kumpulan data yang tidak seimbang dari Repositori UCI	a. Studi ini membahas ketidakseimbangan kelas sebagai tantangan signifikan dalam algoritma klasifikasi b. Pendekatan ansambel menggunakan LightGBM dan XGBoost diusulkan untuk mengatasi masalah ini c. XGBoost mengungguli LightGBM dalam akurasi dan sensitivitas di sebagian besar kumpulan data d. LightGBM menunjukkan	a. Penelitian ini menyarankan penggunaan pendekatan ansambel dengan pengklasifikasi seperti LightGBM dan XGBoost untuk menangani kumpulan data yang tidak seimbang secara efektif b. LightGBM direkomendasikan karena kemampuannya menangani data skala besar secara efisien dan spesifitasnya	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradient boosting

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				spesifitas yang lebih baik dalam memprediksi kelas minoritas dibandingkan dengan XGBoost	c. Studi ini menunjukkan bahwa XGBoost mengungguli LightGBM dalam akurasi dan sensitivitas, menunjukkan evaluasi komparatif dari kedua metode	menggunakan optimasi bayesian
10	Pemodelan Prediktif Menggunakan Metode Ensemble Learning XGBoost dalam Peningkatan Akurasi Klasifikasi Penyakit Ginjal	R. Soefitjadjil, Th. Dwiati Wismarini <sup>2</sup> , Sei Eniyati <sup>3</sup> , Sunardi <sup>4</sup>  KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen) Terakreditasi Nomor 204/E/KPT/2022   Vol. 5, No. 4, Oktober (2024), pp. 1866-1875	a. Penelitian ini bertujuan untuk meningkatkan akurasi dalam mengklasifikasikan pasien Penyakit Ginjal Kronis (CKD) menggunakan metode pembelajaran ansambel XGBoost b. Studi ini berfokus pada pengembangan dan evaluasi model prediktif untuk klasifikasi CKD	a. Kinerja model dievaluasi menggunakan akurasi, presisi, ingatan, dan skor F1, memastikan generalisasi yang baik dengan data uji terpisah b. Faktor-faktor yang berkontribusi terhadap kinerja model diidentifikasi, memengaruhi praktik klinis c. Model XGBoost mencapai akurasi tinggi (99,16%) dalam	a. Penelitian ini menyoroti perlunya metode yang lebih baik dalam klasifikasi CKD untuk membantu profesional Kesehatan b. Makalah ini membahas inkonsistensi dalam data prevalensi CKD di berbagai sumber	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			c. Studi ini berusaha untuk meningkatkan manajemen keseluruhan penyakit ginjal kronis	mengklasifikasikan pasien CKD.		forest dan extreme gradien boosting menggunakan optimasi bayesian.
11	A Random Forest Algorithm Combined with Bayesian Optimization for Atmospheric Duct Estimation	C. Yang et al. (2023)	<p>a. Makalah ini membahas inversi saluran atmosfer, penting untuk evaluasi kinerja radar dan sistem komunikasi</p> <p>b. Ini mengembangkan model hutan acak yang terintegrasi dengan optimasi Bayesian untuk prediksi saluran atmosfer</p> <p>c. Studi ini menyoroti pentingnya parameter model dalam meningkatkan kinerja prediksi</p> <p>d. Ini bertujuan untuk membandingkan</p>	<p>a. Makalah ini menyimpulkan bahwa model BO-RF mengungguli metode lain dalam akurasi dan kemampuan anti-noise untuk estimasi parameter saluran</p> <p>b. Ini menekankan pentingnya parameter model dalam pembelajaran mesin untuk meningkatkan kinerja prediksi</p> <p>c. Penggabungan validasi silang K-fold membantu dalam mendapatkan partisipasi model terbaik dan</p>	<p>a. Penelitian masa depan dapat fokus pada peningkatan akurasi estimasi parameter saluran atmosfer menggunakan teknik pembelajaran mesin canggih</p> <p>b. Menyelidiki integrasi model pembelajaran mesin yang berbeda untuk kinerja yang lebih baik dalam prediksi saluran atmosfer disarankan</p> <p>c. Menjelajahi dampak faktor lingkungan pada radar dan sistem komunikasi di</p>	Perbedaan dibandingkan penelitian terdahulu terdapat pada dataset yang digunakan yaitu hasil pemeriksaan Kesehatan Jemaah haji pada embarkasi Balikpapan tahun 2024 dan metode yang digunakan yaitu metode random forest dan extreme gradien boosting menggunakan optimasi bayesian

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			<p>kinerja model BO-RF yang diusulkan dengan metode lain seperti RF klasik dan XGBoost.</p> <p>e. Makalah ini menekankan perlunya analisis akurat guna radar dan parameter saluran atmosfer</p>	<p>mengurangi kelebihan pemasangan</p> <p>d. Hasilnya menunjukkan bahwa model BO-RF efektif untuk prediksi saluran atmosfer.</p>	<p>lingkungan maritim tetap menjadi area yang signifikan untuk studi masa depan</p> <p>d. Pengembangan lebih lanjut dari sistem pemantauan waktu nyata untuk kondisi atmosfer dapat meningkatkan evaluasi kinerja radar</p>	

## **2.3. Landasan Teori**

### **2.3.1 Ensemble Learning**

Teknik ensemble adalah algoritma yang digunakan dalam pembelajaran mesin untuk mencari prediksi yang paling akurat dibandingkan dengan algoritma alternatif. Metode ini memanfaatkan beberapa algoritma pembelajaran untuk meningkatkan kinerja prediktif. Ensemble adalah model pembelajaran mesin yang menggabungkan prediksi dari dua atau lebih model individu. Model yang terlibat dalam ensemble, yang dikenal sebagai anggota ensemble, dapat memiliki jenis yang sama atau berbeda, dan mungkin dilatih pada dataset yang sama atau berbeda (Buslim et al., 2023).

### **2.3.2 Random forest (RF)**

Random forest adalah salah satu algoritma machine learning paling populer untuk tugas klasifikasi dan regresi. Algoritma ini bekerja dengan membangun banyak pohon keputusan secara acak, lalu menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan keputusan akhir yang lebih akurat dan stabil.

### **2.3.3 Extreme Gradient Boosting (XGBoost)**

XGBoost, sebuah teknik pembelajaran mesin yang sangat dihormati, dengan cermat memanfaatkan kekuatan gabungan dari beberapa model lemah, biasanya pohon keputusan, untuk mengembangkan model prediktif yang kuat (Long & Dung, 2024). Algoritma ini menggunakan pendekatan pelatihan iteratif dimana pohon-pohon tersebut dilatih

secara berurutan, memungkinkan mereka untuk memperbaiki kesalahan sebelumnya dan meningkatkan akurasi prediksi. Algoritma ini menawarkan berbagai karakteristik unik, seperti regularisasi, pemangkasan pohon (tree pruning), penanganan nilai yang hilang, kemampuan pemrosesan paralel, dan analisis pentingnya fitur. Berkat kinerja dan efisiensinya yang mengesankan, XGBoost telah menjadi sangat populer di berbagai bidang yang melibatkan tugas klasifikasi, regresi, dan peringkat (ranking).

#### **2.3.4 Bayesian Optimization**

Bayesian optimization adalah metode yang digunakan untuk mengoptimalkan fungsi objektif yang mahal untuk dievaluasi, sering digunakan dalam berbagai bidang seperti ilmu pengetahuan dan rekayasa. Metode ini sangat efisien dalam hal data dan cocok untuk domain kontinu dengan dimensi kurang dari 20, serta dapat menangani kebisingan stokastik dalam evaluasi fungsi (Garnett, 2023).

Bayesian optimization membangun model pengganti untuk fungsi objektif dan mengkuantifikasi ketidakpastian dalam model tersebut menggunakan teknik pembelajaran mesin Bayesian, seperti regresi proses Gaussian. Kemudian, fungsi akuisisi yang didefinisikan dari model pengganti ini digunakan untuk memutuskan di mana sampel berikutnya harus diambil, tiga fungsi akuisisi umum yang digunakan adalah peningkatan yang diharapkan, pencarian entropi, dan gradien pengetahuan (Frazier, 2018).

### 2.3.5 Hyperparameter tuning

Hyperparameter tuning merupakan proses penting dalam pembelajaran mesin yang melibatkan pemilihan set Hyperparameter yang optimal bagi algoritma pembelajaran untuk meningkatkan kinerjanya. Pencarian Grid merupakan metode populer untuk penyctelan hiperparameter, yang menawarkan pendekatan sistematis untuk mengeksplorasi ruang hyperparameter.

### 2.3.6 Accuracy (Akurasi)

Secara teoretis, Akurasi mengukur proporsi total prediksi yang benar (TP dan TN) dari keseluruhan data. Ini adalah metrik paling intuitif untuk menilai ketepatan model secara umum. Meskipun demikian, akurasi memiliki keterbatasan. Metrik ini dapat sangat menyesatkan (misleading) pada dataset yang tidak seimbang (imbalanced). Sebagai contoh, jika 95% jemaah 'Tidak Berisiko' (kelas mayoritas), model yang memprediksi "Tidak Berisiko" untuk semua orang akan memiliki akurasi 95% namun sama sekali tidak berguna untuk mendeteksi jemaah yang berisiko (kelas minoritas).

### 2.3.7 F1-Score (Macro)

Dalam penelitian ini, F1-Score (khususnya `f1_macro`) dipilih sebagai metrik optimasi utama karena beberapa alasan. Pertama, metrik ini mampu mengatasi data tidak seimbang, tidak seperti Akurasi, karena F1-Score fokus pada performa model terhadap kelas positif. Kedua,

dalam konteks medis, terjadi trade-off antara meminimalkan False Negative (FN) yang berbahaya dan False Positive (FP) yang membuang sumber daya. F1-Score menyediakan satu angka tunggal yang menyeimbangkan kedua kepentingan ini. Ketiga, penggunaan 'Macro' Average memastikan bahwa performa model pada kelas minoritas (Berisiko) memiliki bobot yang sama pentingnya dengan performa pada kelas mayoritas (Tidak Berisiko).

### 2.3.8 AUC - ROC (Area Under the Receiver Operating Characteristic Curve)

Secara teoretis, AUC mengukur kemampuan model untuk melakukan *diskriminasi*, yaitu seberapa baik model dapat membedakan antara kelas positif dan kelas negatif. Kurva ROC sendiri memplot *True Positive Rate* (TPR, identik dengan Recall) melawan *False Positive Rate* (FPR, dihitung sebagai  $\frac{FP}{FP+TN}$  pada berbagai *threshold* klasifikasi.

Interpretasi nilai AUC berkisar dari 0 hingga 1. Nilai AUC = 1.0 mengindikasikan pemisah (*classifier*) yang sempurna. Nilai AUC = 0.5 menunjukkan model tidak memiliki kemampuan diskriminatif (sama dengan tebakan acak), sementara nilai AUC < 0.5 berarti model bekerja lebih buruk daripada tebakan acak. AUC sangat berguna karena tidak bergantung pada *threshold* probabilitas tertentu, memberikan gambaran holistik tentang performa diskriminatif model.

## BAB III

### METODE PENELITIAN

#### 3.1 Jenis, Sifat, dan Pendekatan Penelitian

Studi ini mengadopsi jenis penelitian eksperimental dengan sifat penelitian prediktif dan eksploratif, menggunakan pendekatan kuantitatif. Penelitian ini berfokus pada pengembangan dan pengujian model machine learning (Random Forest dan XGBoost) dengan intervensi optimasi (Bayesian Optimization) untuk memprediksi tingkat risiko kesehatan jemaah haji. Pengujian kinerja model melalui perbandingan dengan model tanpa optimasi merupakan karakteristik dari pendekatan eksperimental.

Tujuan utama penelitian adalah untuk membangun model yang mampu memprediksi tingkat risiko kesehatan jemaah haji di masa depan berdasarkan data pemeriksaan yang ada. Penelitian ini juga bersifat eksploratif dalam mencari kombinasi hyperparameter optimal menggunakan Bayesian Optimization untuk meningkatkan kinerja model, serta dalam mengidentifikasi fitur-fitur yang paling berpengaruh terhadap prediksi risiko.

Pendekatan Kuantitatif diterapkan melalui pengumpulan data numerik (hasil pemeriksaan kesehatan), analisis statistik, serta evaluasi kinerja model menggunakan metrik kuantitatif seperti Accuracy, AUC, Precision dan Recall (F1 Score).

### 3.2 Metode Pengumpulan Data

Dalam penelitian ini, proses pengumpulan data bersumber dari hasil Surveilans Epidemiologi Kesehatan Haji. Data diperoleh melalui pemeriksaan Kesehatan Jemaah haji yang dilakukan secara bertahap dimulai dari Tingkat kecamatan, kab/kota hingga Tingkat provinsi.

Dataset yang dikumpulkan mencakup berbagai hasil pemeriksaan, yaitu Nomor Porsi, Nama Jemaah, Kloter, Rombongan Kloter, Regu Kloter, Umur, No. Paspor, No. Visa, Status Jemaah, Kab/Kota, Jenis Kelamin, WUS (HCG), WUS (PALPASI), ICD 1, ICD 2, ICD 3, ICD 4, ICD 5, STATUS, KR / TK, RISTI / TIDAK, LAIK / TDK LAIK, LANSA / TIDAK. Data ini merupakan hasil pemeriksaan kesehatan jemaah haji pada tahun keberangkatan 2024, yang menjadi fokus utama dalam penelitian ini.

### 3.3 Metode Analisis Data

Metode analisis data dalam penelitian ini dirancang untuk menerapkan dan mengevaluasi pengembangan model prediktif tingkat risiko kesehatan jemaah haji menggunakan Extreme Gradient Boosting (XGBoost) dan Random Forest dengan optimasi Bayesian Optimization (BOXGB dan BO-RF). Proses analisis ini akan melalui beberapa tahapan utama:

#### 1. Pengumpulan Dataset

Dataset yang digunakan adalah hasil pemeriksaan kesehatan jemaah haji Embarkasi Balikpapan tahun 2024, yang telah dijelaskan secara rinci pada bagian Metode Pengumpulan Data. Dataset ini

mencakup berbagai fitur demografi, status kesehatan, dan riwayat penyakit.

## 2. Pra-pemrosesan Data

Tahap awal meliputi pembersihan data untuk memastikan kualitas dan konsistensi. Tahapan ini mencakup identifikasi dan penanganan data duplikat, penanganan nilai yang hilang, serta koreksi inkonsistensi data. Selain itu, proses ini juga akan melibatkan penerjemahan atau pemetaan kode pemeriksaan (misalnya, kode ICD) menjadi representasi yang sesuai jika diperlukan untuk analisis.

Data yang memiliki entri ganda (duplikat) akan diidentifikasi dan dihapus. Keberadaan duplikat dapat menyebabkan bias dalam pelatihan model dan memberikan bobot yang tidak semestinya pada observasi tertentu. Data yang tidak lengkap (memiliki nilai yang hilang atau NaN) akan ditangani. Imputasi Median atau Modus untuk fitur numerik, nilai yang hilang dapat diisi dengan median (nilai tengah) dari kolom tersebut. Untuk fitur kategorikal, nilai yang hilang dapat diisi dengan modus (nilai yang paling sering muncul). Penghapusan Baris/Kolom jika proporsi nilai yang hilang sangat tinggi dalam suatu baris atau kolom, dan imputasi tidak dapat dilakukan secara valid atau berisiko mengurangi kualitas data, maka baris atau kolom tersebut dapat dihapus.

Tahap selanjutnya yaitu Transformasi Data. Setelah pembersihan, data akan ditransformasi agar sesuai dengan format yang

dibutuhkan oleh model machine learning. Kolom-kolom fitur yang bersifat kategorikal (misalnya, 'Jenis Kelamin', 'Status Jemaah') yang direpresentasikan sebagai teks akan diubah menjadi format numerik menggunakan *Label Encoder*. Ini memastikan bahwa model dapat memproses fitur-fitur tersebut.

Selanjutnya dilakukan penskalaan Fitur Numerik. Datasheet yang berjumlah 12, hanya 1 yang numerik, sisanya adalah string. Maka perlu dilakukan perubahan data menjadi numerik. Proses tersebut dilakukan pada tahap label encoding. Selanjutnya, untuk atribut "umur", diperlukan perubahan skala menggunakan *Z Score Standardization*. selain itu, tidak perlu dilakukan feature scaling karena nilai pada atribut tersebut bukan menunjukkan urutan atau jarak antar kategori yang bermakna. Kontribusi yang seimbang dari semua variabel diharapkan akan meningkatkan efektivitas model.

Setelah pra-pemrosesan, dataset akan dibagi menjadi data latih (training data) dan data uji (testing data) dengan rasio 7:3 (70% untuk pelatihan dan 30% untuk pengujian). Pembagian ini dilakukan menggunakan metode *StratifiedKfold* (sebagai bentuk stratified random sampling dalam konteks cross-validation). *StratifiedKfold* memastikan bahwa distribusi kelas target (misalnya, 'Risiko Rendah' dan 'Risiko Tinggi') tetap konsisten antara data latih dan data uji di setiap fold. Ini sangat penting untuk mencegah bias model, terutama jika terdapat ketidakseimbangan kelas dalam dataset.

Pelatihan Model (dengan Optimasi Bayesian): Model Random Forest dan Extreme Gradient Boosting (XGBoost) akan dilatih menggunakan data latih. Untuk mengoptimalkan kinerja kedua model, Bayesian Optimization (BO) akan diintegrasikan untuk hyperparameter tuning. BO adalah metode yang efisien dalam mencari kombinasi hyperparameter optimal (seperti *n\_estimators*, *max\_depth*, *learning\_rate*, dll.) dengan membangun model probabilitas (misalnya, *Gaussian Process*) dari fungsi objektif (misalnya, akurasi validasi) dan secara iteratif memilih hyperparameter berikutnya yang paling menjanjikan untuk dievaluasi. Ini memungkinkan penemuan hyperparameter terbaik secara lebih efisien dibandingkan pencarian grid atau random.

Setelah model dilatih dan dioptimalkan, kinerjanya akan diuji pada data uji yang belum pernah dilihat oleh model. Evaluasi performa akan dilakukan dengan mengidentifikasi nilai parameter yang digunakan sebagai indikator capaian kinerja model. Metrik evaluasi utama meliputi:

Akurasi (Accuracy): Mengukur proporsi total prediksi yang benar.

$$Accuracy = (TP + TN) / (TP + FP + TN + FN) \times 100\% \quad (3.1)$$

Presisi (Precision): Mengukur proporsi instance yang diidentifikasi sebagai positif oleh model yang benar-benar positif.

$$\text{Presisi} = TP / (TP + FP) \times 100\% \quad (3.2)$$

Recall (Recall / Sensitivitas): Mengukur persentase kasus positif sebenarnya yang berhasil diidentifikasi dengan benar oleh model.

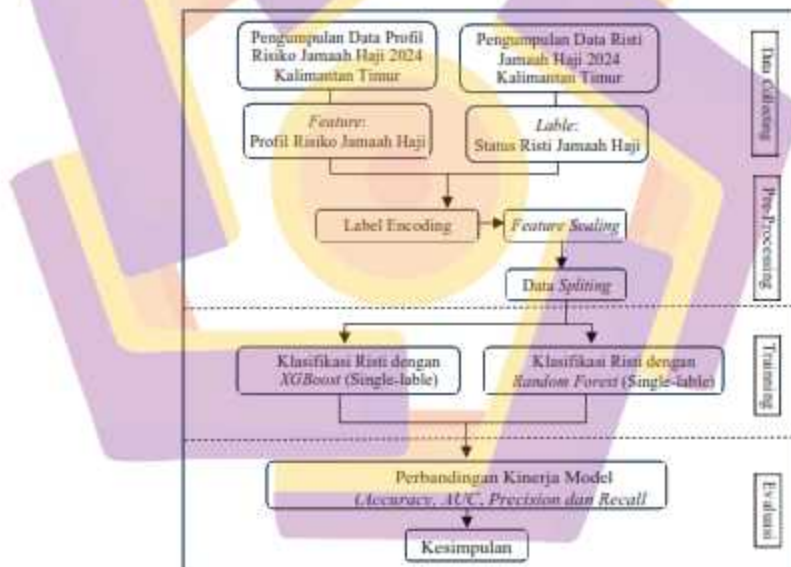
$$\text{Recall} = TP / (TP + FN) \times 100\% \quad (3.3)$$

F1-score (F1): Merupakan rata-rata harmonik dari Presisi dan Recall, memberikan gambaran keseimbangan antara kedua ukuran tersebut.

$$F1 = 2 \times (PRESISI \times RECALL) / (PRESISI + RECALL) \quad (3.4)$$

Ukuran akurasi (Persamaan (3.1)) berfungsi sebagai indikator efektivitas model dalam memprediksi semua kategori target dengan akurat. Selanjutnya, metrik presisi (Persamaan (3.2)) mencerminkan proporsi instance yang diidentifikasi sebagai positif oleh model yang benar-benar positif. Sementara itu, metrik recall (Persamaan (3.3)) mengukur persentase kasus positif sebenarnya yang berhasil diidentifikasi dengan benar oleh model. F1-score (Persamaan (3.4)) bertindak sebagai metrik yang diperoleh dari kombinasi presisi dan recall, memberikan gambaran keseimbangan antara kedua ukuran tersebut dan AUC-ROC (Area Under the Receiver Operating Characteristic Curve) untuk mengukur kemampuan model untuk

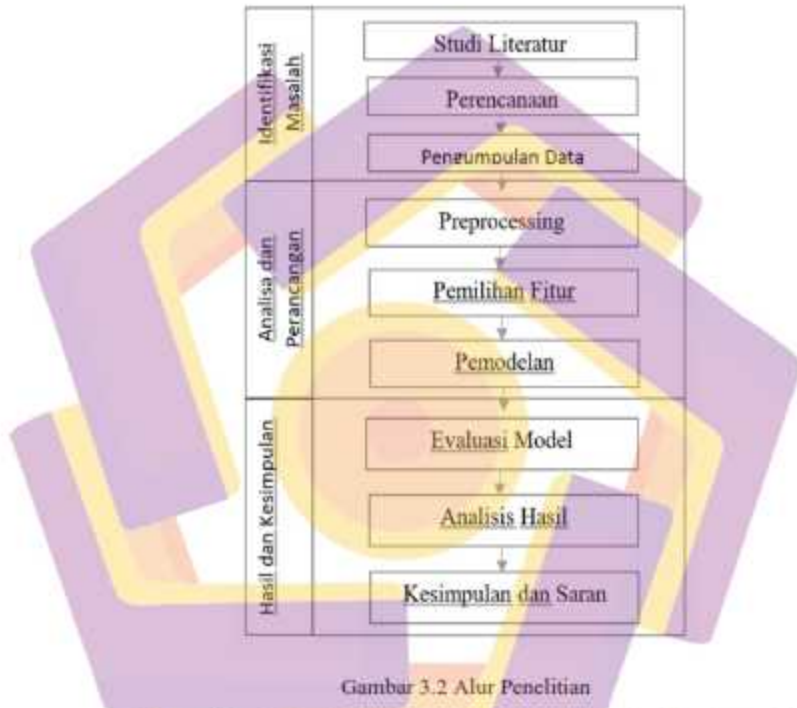
membedakan antara kelas positif dan negatif pada berbagai threshold klasifikasi. Penelitian ini akan membandingkan hasil evaluasi model dengan intervensi optimasi terhadap model standar untuk mengidentifikasi perbedaan dan dampak dari optimasi yang diterapkan. Selain itu, metrik seperti presisi, recall, dan F1-score akan menampilkan nilai untuk setiap kategori atau kelas yang terkait dengan variabel *target* (variabel *dependen*). Ringkasan pendekatan analisis data dapat dilihat pada Gambar 3.1.



Gambar 3.1 Alur Model Yang Diusulkan

### 3.4 Alur Penelitian

Penelitian ini menyajikan detail lengkap tentang proses penelitian, memberikan ringkasan skenario pelaksanaan penelitian yang diilustrasikan dalam Gambar 3.2.



Gambar 3.2 Alur Penelitian

Alur penelitian ini terdiri dari tiga tahap utama, yaitu Identifikasi Masalah, Analisis dan Perancangan, serta Hasil dan Kesimpulan, yang dijelaskan sebagai berikut:

#### 1. Identifikasi Masalah

- Studi Literatur: merupakan tahap awal dalam pengumpulan dan peninjauan komprehensif terhadap referensi dari penelitian terdahulu.

Fokusnya adalah pada studi terkait prediksi risiko kesehatan menggunakan algoritma machine learning, serta teknik optimasi hyperparameter dalam bidang kesehatan. Tujuannya adalah untuk membangun pemahaman mendalam tentang kondisi terkini penelitian, mengidentifikasi gap pengetahuan, dan menemukan justifikasi ilmiah untuk pendekatan yang diusulkan.

- b. **Perencanaan:** Berdasarkan hasil studi literatur, tahap ini mencakup penyusunan metodologi penelitian yang terperinci. Ini meliputi pemilihan dataset yang relevan (data jemaah haji Embarkasi Balikpapan tahun 2024), penentuan algoritma machine learning yang akan digunakan (Random Forest dan XGBoost), serta spesifikasi teknik optimasi yang diterapkan (Bayesian Optimization).
- c. **Pengumpulan Data:** Mengumpulkan data kesehatan jemaah haji dari sumber yang relevan dan berwenang, yaitu hasil Surveilans Epidemiologi Kesehatan Haji tahun 2024.

## **2. Analisis dan Perancangan**

- a. **Pra-pemrosesan Data (Preprocessing):** Melakukan serangkaian langkah untuk menyiapkan data mentah. Ini mencakup pembersihan data (penganganan duplikat, nilai hilang, inkonsistensi), serta transformasi data (encoding fitur kategorikal dan penskalaan fitur numerik) untuk memastikan kualitas dan format data yang optimal untuk pemodelan.

- b. **Pemilihan Fitur (Feature Selection):** Mengidentifikasi dan memilih fitur-fitur yang paling relevan dan berpengaruh dalam memprediksi tingkat risiko kesehatan jemaah haji. Tahap ini dapat melibatkan analisis statistik, metode berbasis model (seperti feature importance dari model tree-based), atau teknik wrapper/filter. Pemilihan fitur yang tepat dapat meningkatkan akurasi model, mengurangi waktu komputasi, dan meningkatkan interpretasi model.
- c. **Pemodelan:** Mengembangkan dan melatih model prediktif. Tahap ini melibatkan penggunaan model Extreme Gradient Boosting (XGBoost) dan Random Forest dengan integrasi Bayesian Optimization (BOXGB dan BO-RF) untuk menemukan hyperparameter optimal. Model akan dilatih pada data latih yang telah diproses.

### **3. Hasil dan Kesimpulan**

- a. **Evaluasi Model:** Mengukur performa model yang telah dilatih dan dioptimasi menggunakan metrik kuantitatif seperti Accuracy, AUC, Precision dan Recall (F1 Score). Evaluasi ini akan dilakukan pada data uji yang terpisah untuk menilai kemampuan generalisasi model.
- b. **Analisis Hasil:** Menganalisis secara mendalam hasil evaluasi model. Ini mencakup interpretasi kinerja model, perbandingan efektivitas BOXGB dan BO-RF, serta identifikasi fitur-fitur yang paling berpengaruh terhadap prediksi risiko kesehatan jemaah haji. Analisis ini bertujuan untuk memberikan wawasan komprehensif mengenai kekuatan dan keterbatasan pendekatan yang diusulkan.

- c. Kesimpulan dan Saran: Menyimpulkan temuan utama penelitian, menjawab rumusan masalah, dan memberikan saran untuk pengembangan model yang lebih baik di masa depan. Saran dapat mencakup eksplorasi fitur tambahan, perluasan dataset, atau penerapan metode machine learning lain.

Alur penelitian ini bertujuan untuk mengoptimalkan prediksi risiko kesehatan jemaah haji guna meningkatkan kualitas layanan kesehatan selama ibadah haji.



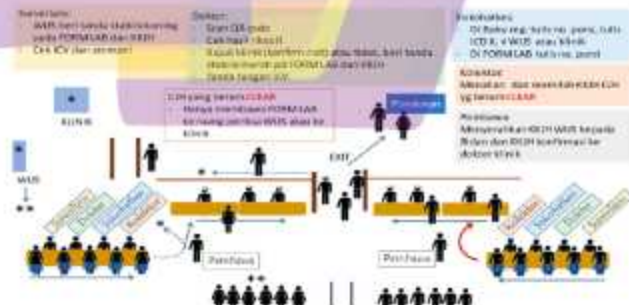
## BAB IV

### HASIL PENELITIAN DAN PEMBAHASAN

Bab ini menyajikan hasil dari penelitian yang telah dilakukan serta pembahasan mendalam terkait hasil tersebut. Pembahasan mencakup proses pengumpulan data, skenario penelitian yang digunakan, langkah-langkah pre-processing data, hingga hasil dan evaluasi model machine learning untuk memprediksi risiko kesehatan pada jemaah haji menggunakan metode random forest dan extreme gradient boosting menggunakan optimasi bayesian berdasarkan data jemaah haji embarkasi Balikpapan tahun 2024.

#### 4.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari hasil surveilans kesehatan jemaah haji yang dilaksanakan oleh tim Kantor Kesehatan Pelabuhan (KKP) Embarkasi Balikpapan. Proses surveilans ini dilakukan sebagai bagian dari prosedur pemeriksaan kesehatan standar sebelum keberangkatan, yang alurnya diilustrasikan pada skema Digambar 4.1



Gambar 4.1 Alur Skema Pengumpulan Data

Proses pengumpulan data ini mengikuti alur surveilans yang sistematis pada saat jemaah haji memasuki asrama embarkasi. Alur ini dirancang untuk memverifikasi status kesehatan dan kelengkapan dokumen setiap jemaah. Secara prosedural, jemaah haji yang tiba akan melalui beberapa pos pemeriksaan utama. Pos-pos ini mencakup:

- a. Pos Surveilans: Petugas melakukan verifikasi kelengkapan vaksinasi (ICV) dan stempel. Di pos ini, jemaah Wanita Usia Subur (WUS) diberi penanda khusus (stabilo kuning) pada FORM LAB dan Kartu Kesehatan Jemaah Haji (KKJH) untuk pemeriksaan lebih lanjut.
- b. Pos Dokter: Petugas dokter melakukan validasi data (misalnya, scan QR code), memeriksa hasil Rikes II (pemeriksaan kesehatan tahap kedua), dan mengonfirmasi status risiko (risti) jemaah. Jemaah yang tidak memerlukan rujukan klinik akan diberi tanda stabilo pada dokumen kesehatan dan stempel ICV. Jemaah yang terindikasi risti akan dirujuk ke KLINIK.
- c. Pos Siskohatkes: Pos ini berperan krusial dalam digitalisasi data. Petugas Siskohatkes melakukan entri data ke dalam sistem (Buku reg) dan FORM LAB. Data vital seperti nomor porsi, status WUS (v WUS), dan rujukan klinik dicatat di sini. Yang terpenting, kode diagnostik komorbiditas jemaah dicatat sebagai ICD X, yang kemudian menjadi basis data mentah untuk fitur ICD 1 hingga ICD 5 dalam penelitian ini.

Kumpulan data (dataset) final yang berhasil dikumpulkan untuk penelitian ini mencakup **6.037 record** data jemaah haji. Dataset ini terdiri dari dua komponen utama yang digunakan untuk pemodelan:

- Data Fitur (Variabel Independen)** Merupakan data prediktor yang digunakan untuk mengidentifikasi risiko (lihat tabel 4.1). Fitur-fitur ini mencakup variabel demografis dan klinis, yaitu Umur, Status Jemaah, Jenis Kelamin, Pemeriksaan WUS (WUS HCG, WUS PALPASI), Riwayat komorbiditas (direpresentasikan sebagai ICD 1 hingga ICD 5), STATUS KR/TK (Indikator penggunaan Tongkat atau Kursi Roda).

Tabel 4.1 Data Fitur

Umur	Status Jemaah	Jenis Kelamin	WUS HCG	WUS PALPASI	ICD 1	ICD 2	ICD 3	ICD 4	ICD 5	STATUS	KR/TK
42	TPHIL	L	TIDAK PERIKSA	TIDAK PERIKSA	General examination and investigation of persons with out complaint and reported diagnosis	NEGATIF	NEGATIF	NEGATIF	NEGATIF	BUKANLANJIA	NORMAL

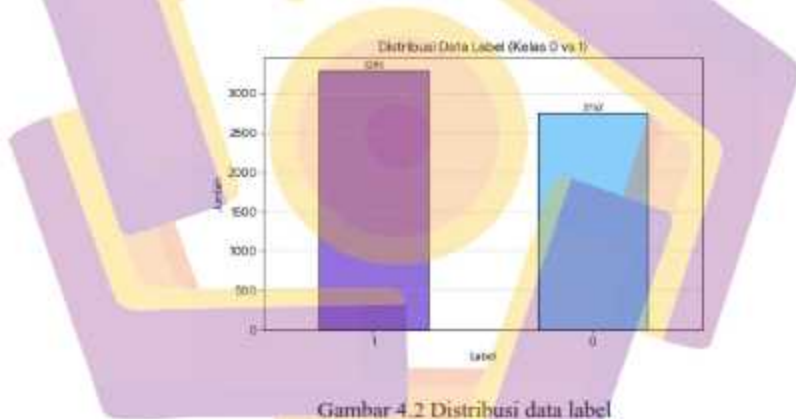
Tabel 4.2 Data Fitur (lanjutan)

Umur	Status Jemaah	Jenis Kelamin	WUS HCG	WUS PALPASI	ICD 1	ICD 2	ICD 3	ICD 4	ICD 5	STATUS	KR TK
65	KAROM	L	TIDAK PERIKSA	TIDAK PERIKSA	Senility	Asthma	NEGATIF	NEGATIF	NEGATIF	LAN SIA	NOR MAL
31	KARU	L	TIDAK PERIKSA	TIDAK PERIKSA	Asthma	NEGATIF	NEGATIF	NEGATIF	NEGATIF	BUKAN LAN SIA	NOR MAL
58	Jemaah	P	TIDAK PERIKSA	TIDAK PERIKSA	Hypertensive heart disease	NEGATIF	NEGATIF	NEGATIF	NEGATIF	BUKAN LAN SIA	NOR MAL
35	Jemaah	P	NEGATIF	NEGATIF	Asthma	NEGATIF	NEGATIF	NEGATIF	NEGATIF	BUKAN LAN SIA	NOR MAL

2. **Data Label (Variabel Dependen):** Merupakan data target atau *ground truth* dari file Data Label.xlsx. Data yang ditunjukkan pada table 4.3 ini menunjukkan klasifikasi risiko jemaah (misalnya, 'Berisiko Tinggi' dengan nilai 1 atau 'Tidak Berisiko' dengan nilai 0) yang akan diprediksi oleh model.

Tabel 4.3 Data Label

Prediksi Risti	Keterangan
0	Tidak Beresiko
0	Tidak Beresiko
0	Tidak Beresiko
0	Tidak Beresiko
...	...
1	Beresiko Tinggi
1	Beresiko Tinggi
1	Beresiko Tinggi
1	Beresiko Tinggi
1	Beresiko Tinggi
0	Tidak Beresiko
1	Beresiko Tinggi



Gambar 4.2 Distribusi data label

Distribusi data label yang ditunjukkan pada gambar 4.2 memperlihatkan Kelas 1 sebanyak 3.285 data (54,4%) dan Kelas 0 sebanyak 2.752 data (45,6%). Pengumpulan data yang komprehensif ini sejalan dengan penelitian yang dilakukan oleh (Armis Sibuea et al., 2025) yang menemukan asosiasi kuat antara usia, riwayat medis, dan morbiditas pada jemaah haji, yang menegaskan bahwa faktor demografis (usia),

komorbiditas (riwayat penyakit/ICD), dan status fungsional adalah prediktor utama dalam asesmen kesehatan jemaah haji.

#### 4.2 Skenario Penelitian

Penelitian ini mengusulkan skenario pemodelan prediktif untuk mengklasifikasikan risiko kesehatan jemaah haji. Skenario ini dirancang untuk menemukan model dengan performa terbaik melalui beberapa tahapan utama:

##### 1. Algoritma Klasifikasi

Penelitian ini membandingkan dua algoritma ensemble learning yang *state-of-the-art* karena kemampuannya menangani data tabular kompleks yaitu XGBoost (Extreme Gradient Boosting) dan Random Forest.

##### 2. Optimasi Hiperparameter

Alih-alih menggunakan Grid Search atau Random Search yang komputasionalnya mahal, penelitian ini mengadopsi Bayesian Optimization (BO) menggunakan library *scikit-optimize*. BO dipilih karena lebih efisien dalam menemukan hiperparameter optimal dengan jumlah iterasi yang lebih sedikit.

##### 3. Validasi Model

Untuk mendapatkan estimasi performa model yang *robust* dan tidak *overfitting*, digunakan skema *Stratified 10-Fold Cross-Validation*.

#### 4. Stratified K-Fold Cross-Validation

Validasi model adalah tahapan krusial untuk mengukur seberapa baik model dapat digeneralisasi pada data baru yang belum pernah dilihat sebelumnya. Metode *K-Fold Cross-Validation* standar membagi dataset  $D$  secara acak menjadi  $k$  lipatan (folds) yang berukuran sama. Model kemudian dilatih pada  $k - 1$  lipatan dan diuji pada 1 lipatan yang tersisa (disebut *validation set*).

Proses ini diulang  $k$  kali, sehingga setiap lipatan mendapat kesempatan satu kali menjadi *validation set*. Namun, dalam kasus data medis seperti risiko kesehatan jemaah haji, sering terjadi ketidakseimbangan kelas (*class imbalance*), di mana jumlah jemaah 'Berisiko Tinggi' (kelas minoritas) jauh lebih sedikit daripada jemaah 'Tidak Berisiko' (kelas mayoritas). Jika menggunakan *K-Fold* standar, pembagian acak berpotensi menghasilkan beberapa lipatan yang memiliki proporsi kelas yang sangat berbeda dari dataset asli, atau bahkan tidak memiliki sampel kelas minoritas sama sekali. Hal ini dapat menyebabkan estimasi performa model yang bias dan tidak stabil. Untuk mengatasi masalah ini, penelitian ini mengadopsi *Stratified K-Fold Cross-Validation*.

Metode ini bekerja dengan memodifikasi proses *sampling* untuk memastikan bahwa setiap lipatan (*fold*) merupakan representasi yang proporsional dari distribusi kelas pada keseluruhan dataset. Secara prosedural, jika dataset  $D$  memiliki  $N$

sampel, dengan  $N_c$  adalah jumlah sampel untuk kelas  $c \in C$  (dimana  $C$  adalah himpunan semua kelas), maka *Stratified K-Fold* akan mempartisi data menjadi  $k$  lipatan ( $F_1, F_2, \dots, F_k$ ). Untuk setiap lipatan  $F_i$ , jumlah sampel dari kelas  $c$  di dalamnya adalah  $\approx \frac{N_c}{k}$ . Dengan  $k = 10$  (seperti dalam penelitian ini), dataset dibagi menjadi 10 lipatan, di mana masing-masing lipatan memiliki proporsi jemaah 'Berisiko Tinggi' dan 'Tidak Berisiko' yang sama (atau hampir sama) dengan proporsi pada 6.037 data asli.

Model dilatih 10 kali, dan performa akhirnya dihitung sebagai rata-rata dari 10 iterasi tersebut, menghasilkan evaluasi yang jauh lebih reliabel. Inisialisasi *StratifiedKFold* dilakukan menggunakan *library* *scikit-learn*. Parameter `n_splits=10` mendefinisikan jumlah lipatan, `shuffle=True` mengacak data sebelum dibagi, dan `random_state=50` memastikan hasil pembagian data dapat direproduksi.

```
# 1. Inisialisasi StratifiedKFold
from sklearn.model_selection import StratifiedKFold
skf = StratifiedKFold(n_splits=10, shuffle=True,
                      random_state=50)
```

Proses *looping* kemudian digunakan untuk membagi data fitur  $X$  dan label  $Y$  menjadi set pelatihan ( $X_{train}, y_{train}$ ) dan set pengujian ( $x_{test}, y_{test}$ ) untuk setiap iterasi *fold*.

```
# 2. Iterasi melalui setiap fold
# (X dan Y adalah dataset fitur dan label lengkap)
for train_idx, test_idx in skf.split(X, Y):
    # Membagi data berdasarkan indeks yang dihasilkan
```

```

X_train, X_test = X.iloc[train_idx],
X.iloc[test_idx]
y_train, y_test = Y[train_idx], Y[test_idx]

# Model (misal: opt.fit) dilatih pada X_train, y_train
# Model dievaluasi pada X_test, y_test

```

### 4.3 Pre-processing Data

Sebelum data dimasukkan ke dalam model, dilakukan tahap *pre-processing* untuk memastikan data bersih dan dapat diolah oleh algoritma *machine learning*. Ada dua teknik utama yang digunakan yaitu *Label Encoding* dan *Standard Scaling*.

#### 1. Label Encoding

Algoritma *machine learning* tidak dapat memproses data kategorikal (teks) secara langsung (Rodriguez et al., 2018). *Label Encoding* adalah proses mengubah data kategorikal (misalnya: "Pria", "Wanita") menjadi nilai numerik (misalnya: 0, 1).

```

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
categorical_columns =
df.select_dtypes(include=['object',
'category']).columns
for col in categorical_columns:
df_encoded[col] = le.fit_transform(df[col])

```

Langkah ini diterapkan pada semua kolom yang terdeteksi sebagai tipe *'object'* atau *'category'* dalam dataset fitur yang ditunjukkan pada table 4.5.

Tabel 4.4 Hasil setelah Label Encoding

	UMUR	STATUS JEMAH	JENIS KELAMIN	WUS (WUS PALPASI)	IC1	IC2	IC3	IC4	IC5	STATUS US	KRITIK	
0	42	6	0	1	2	71	106	74	47	13	0	3
1	50	7	0	1	2	71	106	74	47	13	0	3
2	36	4	0	0	1	71	106	74	47	13	0	3
3	43	4	1	0	1	71	106	74	47	13	0	3
4	47	4	1	1	2	71	106	74	47	13	0	3

## 2. Standard Scaling

Fitur numerik seperti *Umur* memiliki rentang nilai yang berbeda dengan fitur lainnya, jika tidak diskalakan maka fitur dengan rentang nilai besar dapat mendominasi proses *learning* model. *Standard Scaling* (Standardisasi) mentransformasi data sehingga memiliki rata-rata ( $\mu$ ) 0 dan standar deviasi ( $\sigma$ ) 1 seperti pada persamaan 4.1:

$$z = \frac{(x - \mu)}{\sigma} \quad (4.1)$$

Dimana:

$z$  adalah nilai fitur setelah *scaling* (nilai standar).

$x$  adalah nilai asli dari *data point*.

$\mu$  adalah rata-rata (mean) dari seluruh nilai pada fitur tersebut.

$\sigma$  adalah standar deviasi (standard deviation) dari fitur tersebut.

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
df_encoded['Umur'] =
scaler.fit_transform(df_encoded[['Umur']])
```

Tabel 4.5 hasil setelah dilakukan scaler pada data umur

#	Umur	Status Jenazah	Jenis Kelamin	WUS S HC G	WUS PALPA SI	IC D 1	IC D 2	IC D 3	IC D 4	IC D 5	STA TUS	K R T K
0	-0.908322	6	0	1	2	71	106	74	47	13	0	3
1	-0.246388	7	0	1	2	71	106	74	47	13	0	3
2	-1.404772	4	1	0	1	71	106	74	47	13	0	3
3	-0.825580	4	1	0	1	71	106	74	47	13	0	3
4	-0.494613	4	0	1	2	71	106	74	47	13	0	3
...	...	...	...	...	...	...	...	...	...	...	...	...
6032	-1.404772	0	1	0	1	68	59	74	47	13	0	3
6033	0.332603	0	1	1	2	85	84	74	47	13	0	3
6034	-0.080905	0	1	1	2	68	59	74	47	13	0	3
6035	0.994737	0	0	1	2	153	46	48	47	13	0	3
6036	0.498287	0	1	1	2	29	106	74	47	13	0	3

6037 rows × 12 columns

Dalam penelitian ini, *scaling* hanya diterapkan pada fitur *Umur* menggunakan *Z Score Standardization* terlihat pada table 4.6. Selain itu, tidak perlu dilakukan *feature scaling* karena nilai pada atribut

tersebut bukan menunjukkan urutan atau jarak antar kategori yang bermakna.

#### 4.4 Hasil dan Evaluasi Model

Evaluasi performa model klasifikasi secara fundamental didasarkan pada Confusion Matrix (Matriks Kebingungan). Matriks ini memetakan hasil prediksi model terhadap nilai aktual (*ground truth*). Untuk studi kasus ini (Risiko Positif (1) vs Negatif (0)), komponennya terdiri dari: True Positive (TP), yaitu jemaah yang *benar-benar berisiko* dan diprediksi dengan benar sebagai *berisiko*; True Negative (TN), yaitu jemaah yang *tidak berisiko* dan diprediksi dengan benar sebagai *tidak berisiko*; False Positive (FP) / Kesalahan Tipe I, yang merupakan jemaah *tidak berisiko* namun salah diprediksi sebagai *berisiko*; dan False Negative (FN) / Kesalahan Tipe II, yaitu jemaah yang *benar-benar berisiko* namun salah diprediksi sebagai *tidak berisiko*. Berdasarkan empat komponen ini, metrik evaluasi standar diturunkan:

##### 4.4.1 Model 1 XGBoost dengan Bayesian Optimization

XGBoost adalah implementasi *Gradient Boosting* yang canggih. Model ini membangun pohon keputusan secara sekuensial, di mana setiap pohon baru dikhususkan untuk memperbaiki kesalahan (residu) dari pohon sebelumnya. *Bayesian Optimization* (BO) digunakan untuk menyetel hiperparameter XGBoost (Nguyen, 2019) seperti `max_depth`, `n_estimators`, `learning_rate`, dll yang ditunjukkan pada table 4.8. BO membangun model probabilistik (Fungsi Surrogate) dari fungsi

objektif (dalam hal ini, F1-Score) dan menggunakan Fungsi Akuisisi untuk secara cerdas memilih kombinasi hiperparameter berikutnya yang paling menjanjikan (Rong et al., 2020).

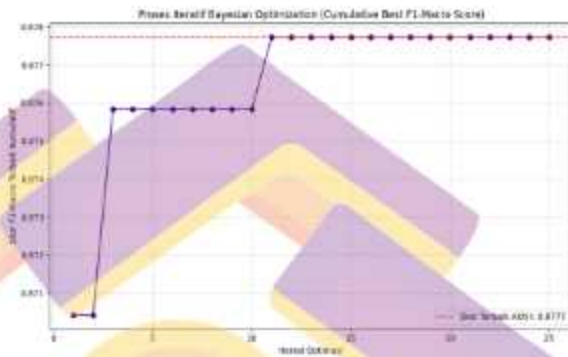
Tabel 4.6 Search Space XGBoost

#	Hyperparameter	Default value	Search space
XGBOOST	n_estimators	100	(100, 600)
	max_depth	none	(3, 10)
	learning_rate	1	(0.01, 0.3)
	subsample	1	(0.6, 1.0)
	colsample_bytree	-	(0.6, 1.0)
	Gamma	-	(0, 5)
	min_child_weight	-	(1, 10)
	reg_alpha	-	(1e-5, 1.0)
reg_lambda	-	(1e-5, 1.0)	

### 1. Analisis Proses Optimasi XGBoost dan BO

Grafik "Proses Iteratif Bayesian Optimization" untuk XGBoost memberikan wawasan tentang efisiensi proses pencarian hiperparameter. Proses dimulai dengan skor *F1-Macro* yang relatif rendah (Iterasi Awal 1-3), di mana algoritma mengeksplorasi kombinasi parameter awal. Kemudian, terjadi dua lompatan (step) performa yang signifikan (Lompatan Performa Iterasi 4 & 11), yang menunjukkan bahwa BO berhasil menemukan area yang menjanjikan dalam ruang pencarian. Setelah iterasi ke-11, skor *F1-Macro* kumulatif terbaik tidak lagi meningkat (Konvergensi Iterasi 12-25) dan stabil di nilai akhir **0.8777**. Ini mengindikasikan bahwa sisa 14 iterasi digunakan

untuk *eksploitasi* di sekitar solusi optimal, dan 25 iterasi sudah cukup untuk mencapai konvergensi seperti yang ditunjukkan pada gambar 4.2.



Gambar 4.3 Proses Iteratif Bayesian Optimization XGBoost (Cumulative Best F1-Macro Score)

## 2. Hasil Evaluasi 10-Fold CV XGBoost dan BO

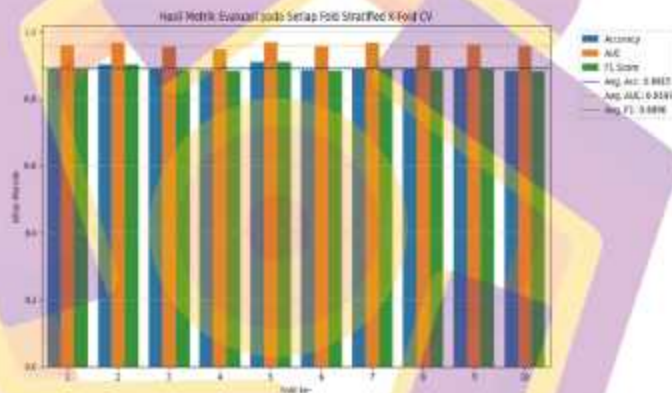
Hasil rata-rata dari 10-fold cross-validation pada table 4.9 dan 4.10 menggunakan model XGBoost yang telah dioptimasi adalah sebagai berikut:

Tabel 4.7 Rata-rata 10-Fold CV XGBoost + BO

FOLD	Subset Accuracy	AUC Score	F1 Score (macro)
FOLD 1	0.8907	0.9611	0.8892
FOLD 2	0.9023	0.9672	0.9014
FOLD 3	0.8874	0.9542	0.8853
FOLD 4	0.8858	0.9477	0.8836
FOLD 5	0.9106	0.9676	0.9095
FOLD 6	0.8858	0.9555	0.8833
FOLD 7	0.8907	0.9658	0.8891

Tabel 4.8 Rata-rata 10-Fold CV XGBoost + BO (Lanjutan)

FOLD	Subset Accuracy	AUC Score	F1 Score (macro)
FOLD 8	0.8872	0.9595	0.8848
FOLD 9	0.8905	0.9620	0.8886
FOLD 10	0.8839	0.9561	0.8813
NILAI RATA-RATA	0.8915	0.9597	0.8896



Gambar 4.4 Hasil Metrik Evaluasi pada Setiap Fold Stratified K-Fold CV XGBoost

grafik yang ditunjukkan pada gambar 4.4 menunjukkan bahwa hasil metrik evaluasi pada setiap Fold Stratified K-Fold CV yang dievaluasi memiliki kinerja yang baik dan sangat stabil. Model ini memiliki kemampuan yang sangat kuat untuk memisahkan kelas (AUC tinggi) dan kemampuan yang baik untuk memprediksi kelas secara akurat (Akurasi & F1-Score baik).

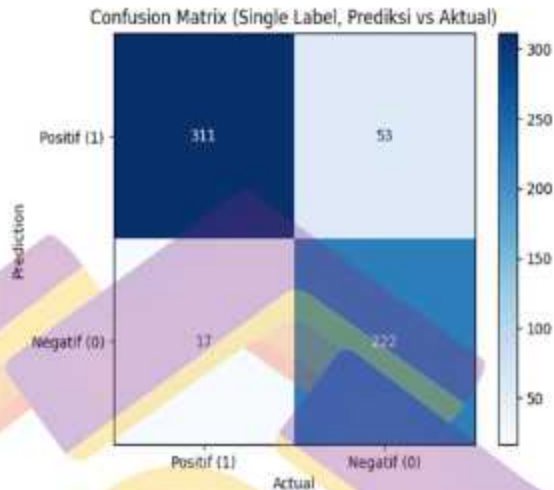
### 3. Analisis Confusion Matrix XGBoost dan BO

Evaluasi mendalam dilakukan terhadap hasil prediksi pada fold terakhir (603 data uji) untuk memahami karakteristik keputusan model. Berdasarkan Gambar 4.5, detail performa model adalah sebagai berikut:

*True Positive (TP)*: 311 Jemaah Model berhasil mengidentifikasi 311 jemaah yang benar-benar berisiko tinggi. Ini mencerminkan tingkat *Recall* (Sensitivitas) sebesar 94,8% (311 dari total 328 kasus positif aktual). Tingginya angka ini menunjukkan kemampuan model yang sangat baik dalam "menangkap" jemaah yang membutuhkan perhatian medis.

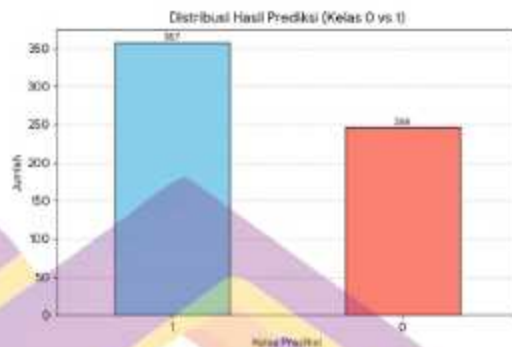
*True Negative (TN)*: 222 Jemaah Model berhasil mengklasifikasikan 222 jemaah sehat dengan tepat. *False Positive (FP)*: 53 Jemaah (*Over-estimation*) Terdapat 53 jemaah sehat yang diprediksi berisiko. Meskipun ini adalah kesalahan, dalam konteks skrining kesehatan massal, *False Positive* seringkali merupakan "biaya" yang dapat diterima untuk memastikan keamanan (*Safety Bias*). Lebih baik sistem menduga sakit (untuk diperiksa ulang) daripada melewatkan yang sakit.

*False Negative (FN)*: 17 Jemaah (*Under-estimation*) Hanya terdapat 17 kasus di mana model gagal mendeteksi risiko (jemaah berisiko dianggap sehat). Angka ini ditekan seminimal mungkin karena merupakan risiko paling fatal.



Gambar 4.5 Confusion Matrix XGBoost (Single Label, Prediksi vs Aktual)

Analisis lebih lanjut terhadap distribusi data prediksi menunjukkan bahwa model XGBoost yang dioptimasi memiliki kecenderungan untuk bersikap Agresif/Sensitif. Total Aktual Positif sebanyak 328 jemaah (54,4% dari data uji). Total Prediksi Positif (TP + FP) sebanyak 364 jemaah (60,4% dari data uji). Selisih ini (Prediksi > Aktual) mengindikasikan bahwa model cenderung memperluas cakupan deteksinya yaitu menghasilkan lebih banyak prediksi Kelas 1 sejumlah 357 data (59,2%) dan kelas 0 sejumlah 246 data (40,8%) seperti pada gambar 4.6 untuk meminimalisir peluang lolosnya jemaah berisiko.



Gambar 4.6 Distribusi Hasil Prediksi Model 1

Perilaku ini sangat selaras dengan tujuan operasional kesehatan haji yang memprioritaskan keselamatan jemaah (*Zero Accident*). Model ini menunjukkan jumlah *False Negative* (FN) yang relatif rendah (17), yang sangat penting dalam konteks medis. Gagal mendeteksi jemaah berisiko (FN) jauh lebih berbahaya daripada salah mengidentifikasi jemaah sehat sebagai berisiko (FP).

#### 4.4.2 Model 2 Random Forest dengan Bayesian Optimization

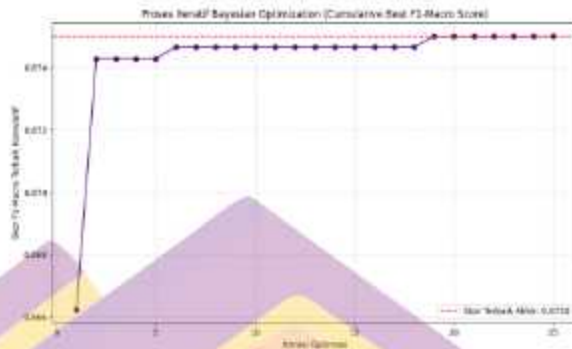
Random Forest (RF) adalah algoritma *ensemble* yang bekerja dengan membangun banyak pohon keputusan (hutan) pada *subset* data yang di-*bootstrap* (bagging) dan *subset* fitur acak (Xu et al., 2017). Keputusan akhir diambil melalui *voting* (klasifikasi) atau rata-rata (regresi) dari semua pohon. Mirip dengan XGBoost, RF dioptimasi menggunakan *BayesSearchCV* dengan ruang pencarian yang disesuaikan untuk hiperparameter RF (lihat ditabel 4.11).

Tabel 4.9 Search Space Random Forest

#	Hyperparameter	Default value	Search space
RF	n_estimators	100	(100, 500)
	max_depth	none	(3, 10)
	min_samples_split	1	(2, 10)
	min_samples_leaf	1	(1, 5)
	max_features	Max Value (factor number)	(0.3, 1.0)

### 1. Analisis Proses Optimasi Random Forest dan BO

Proses optimasi untuk Random Forest (Gambar 4.7) menunjukkan pola yang sedikit berbeda namun sama-sama efisien. Model ini mengalami lompatan performa yang sangat tajam pada iterasi awal (Lompatan Awal Iterasi 2-3), menunjukkan bahwa BO dengan cepat menemukan kombinasi hiperparameter yang jauh lebih baik daripada *default*. Setelah itu, terdapat peningkatan-peningkatan kecil (Peningkatan Minor Iterasi 4-19), sebelum akhirnya mencapai nilai puncak. Model mencapai skor F1-Macro kumulatif terbaiknya, **0.8750**, di sekitar iterasi ke-20 dan kemudian stabil (Konvergensi Iterasi 20-25).



Gambar 4.7 Proses Iteratif Bayesian Optimization (Cumulative Best F1-Macro Score Random Forest)

## 2. Hasil Evaluasi 10-Fold CV Random Forest dan BO

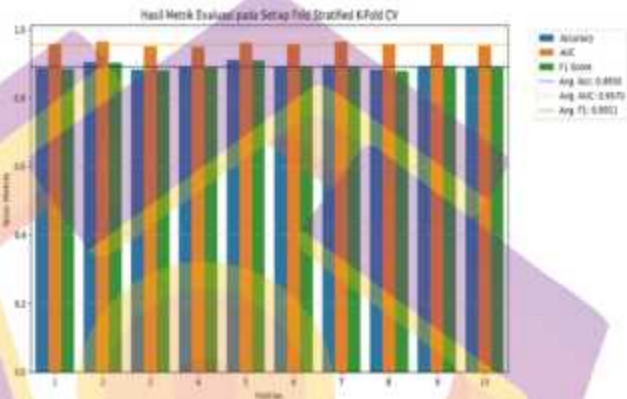
Hasil rata-rata dari 10-fold cross-validation pada tabel 4.12 menggunakan model Random Forest yang telah dioptimasi: a.

**Rata-rata Subset Accuracy: 0.8930 (89.30%)** b. **Rata-rata AUC Score: 0.9570 (95.70%)** c. **Rata-rata F1-Score (Macro): 0.8911 (89.11%)**

Tabel 4.10 Rata-rata 10-Fold CV RF + BO

FOLD	Subset Accuracy	AUC Score	F1 Score (macro)
FOLD 1	0.8874	0.9581	0.8855
FOLD 2	0.9040	0.9643	0.9027
FOLD 3	0.8825	0.9507	0.8802
FOLD 4	0.8924	0.9488	0.8903
FOLD 5	0.9106	0.9607	0.9094
FOLD 6	0.8924	0.9575	0.8903
FOLD 7	0.8940	0.9644	0.8921
FOLD 8	0.8823	0.9569	0.8797
FOLD 9	0.8939	0.9564	0.8921

FOLD	Subset Accuracy	AUC Score	F1 Score (macro)
FOLD 10	0.8905	0.9522	0.8886
NILAI RATA-RATA	0.8930	0.9570	0.8911



Gambar 4.8 Hasil Metrik Evaluasi pada Setiap Fold Stratified K-Fold CV Random Forest

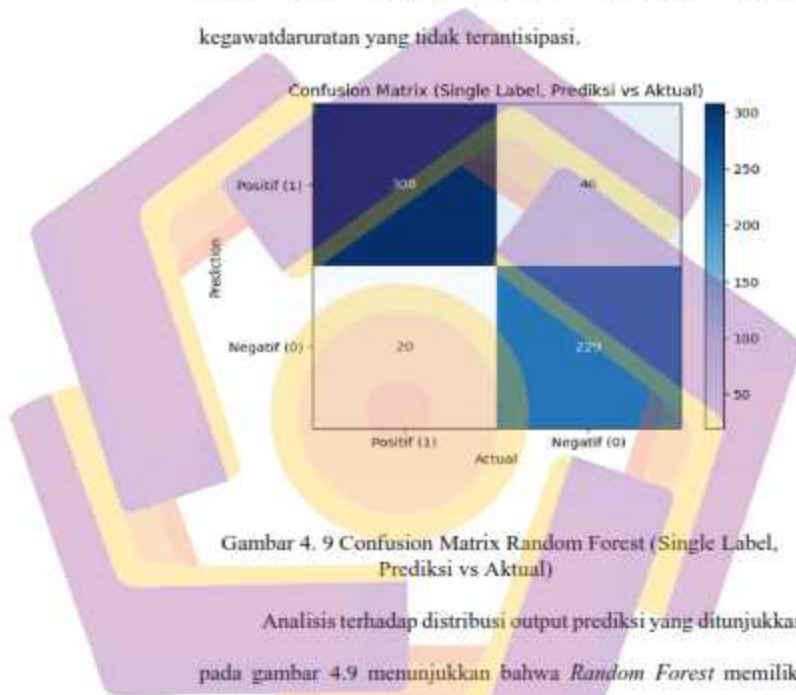
Pada grafik yang ditunjukkan pada gambar 4.8, garis putus-putus horizontal merepresentasikan nilai rata-rata dari setiap metrik: Akurasi pada 89.30% (biru), AUC pada 95.70% (orange), dan F1-Score pada 89.11% (hijau). Secara visual, grafik menunjukkan bahwa model memiliki performa yang sangat stabil, di mana tinggi batang untuk ketiga metrik relatif seragam antar fold. AUC (warna orange) secara konsisten menjadi metrik tertinggi di setiap lipatan, mendekati angka 0.96, yang menegaskan kemampuan diskriminatif model yang sangat kuat. Variabilitas performa antar fold sangat minim meskipun terdapat sedikit fluktuasi seperti Fold 5 yang

menunjukkan performa sedikit di atas rata-rata dan Fold 8 yang sedikit di bawahnya perbedaannya tidak signifikan. Konsistensi ini mengindikasikan bahwa model Random Forest yang dikembangkan bersifat robust dan tidak bias terhadap subset data tertentu.

### 3. Analisis Confusion Matrix Random Forest dan BO

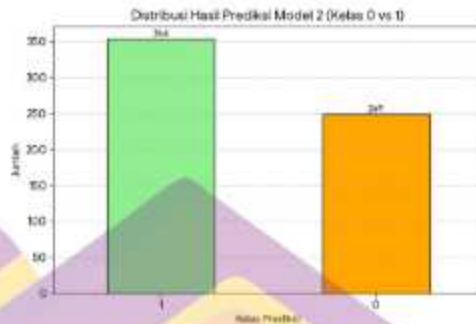
Hasil evaluasi pada *fold* terakhir (merujuk Gambar 4.7), kinerja *Random Forest* menunjukkan profil keputusan yang lebih konservatif dibandingkan XGBoost, dengan rincian True Positive (TP) sebanyak 308 Jemaah, model berhasil mendeteksi 308 jemaah berisiko tinggi dengan benar. Tingkat *Recall* tercatat sebesar 93,9% (308 dari 328). Angka ini sedikit di bawah XGBoost (311 TP). Menunjukkan bahwa *Random Forest* "melewatkan" 3 jemaah berisiko lebih banyak dibandingkan model XGBoost. True Negative (TN) sebanyak 229 Jemaah, keunggulan utama RF terlihat pada kemampuannya mengenali jemaah sehat. Model ini berhasil mengklasifikasikan 229 jemaah sehat dengan tepat, lebih tinggi dibandingkan XGBoost (222 TN), yang berarti RF lebih jarang melakukan "salah tuduh" pada jemaah sehat. False Positive (FP) sebanyak 46 Jemaah, jumlah kesalahan prediksi positif (sehat dianggap sakit) lebih rendah, yaitu 46 kasus (dibandingkan 53 pada XGBoost). RF terbukti lebih efisien karena mengurangi jumlah jemaah sehat yang harus menjalani pemeriksaan ulang yang tidak perlu. Model ini lebih "selektif" dalam memberikan label berisiko.

False Negative (FN) sebanyak 20, terdapat 20 jemaah berisiko tinggi yang gagal terdeteksi oleh sistem. Dalam konteks *critical safety*, selisih 3 orang (dibandingkan 17 FN pada XGBoost) adalah angka yang signifikan karena menyangkut potensi kegawatdaruratan yang tidak terantisipasi.



Gambar 4. 9 Confusion Matrix Random Forest (Single Label, Prediksi vs Aktual)

Analisis terhadap distribusi output prediksi yang ditunjukkan pada gambar 4.9 menunjukkan bahwa *Random Forest* memiliki karakter Konservatif/Seimbang dengan hasil Total Aktual Positif sebanyak 328 jemaah dan Total Prediksi Positif (TP + FP): 354 jemaah. Selisih antara prediksi positif (354) dan kondisi aktual (328) pada RF lebih kecil dibandingkan XGBoost. Hal ini menunjukkan bahwa Random Forest tidak "seagresif" XGBoost dalam menandai risiko.



Gambar 4.10 Distribusi Hasil Prediksi Model 2

Pada gambar 4.10, Random Forest berusaha menyeimbangkan antara presisi (ketepatan deteksi) dan *recall* (cakupan deteksi), sehingga menghasilkan lebih sedikit alarm palsu tetapi dengan risiko melewatkan sedikit lebih banyak kasus nyata. Jika pada distribusi hasil prediksi model 1 (gambar 4.6) terlihat kecenderungan agresif dalam memberikan label '1', data pada distribusi hasil prediksi model 2 (gambar 4.10) menunjukkan distribusi yang lebih ketat. Random Forest memprediksi total 354 jemaah sebagai "Berisiko" (TP+FP), jumlah ini lebih sedikit dibandingkan XGBoost yang memprediksi 364 jemaah. Selisih 10 prediksi positif ini membuktikan bahwa Random Forest tidak semudah XGBoost dalam memvonis status "Berisiko". Model ini membutuhkan bukti fitur yang lebih kuat (mayoritas *vote* dari pohon keputusan) sebelum menetapkan label positif. Hal ini menjelaskan mengapa angka *False Positive* pada RF lebih rendah

(data output menunjukkan lebih banyak label '0' yang tepat sasaran pada jemaah sehat), namun konsekuensinya adalah *Recall* yang sedikit turun.

#### 4.5 Pembahasan

Bagian ini membandingkan temuan dari kedua model dan mendiskusikan implikasinya dalam konteks studi kasus.

##### 4.5.1 Perbandingan Performa Model

Tabel 4.13 merangkum perbandingan metrik rata-rata dari kedua skenario model setelah optimasi:

Tabel 4.11 Tabel Perbandingan metrik rata-rata kedua skenario

Model	Accuracy	AUC	F1-Score (Macro)
XGBoost + BO	0.8915	<b>0.9597</b>	0.8896
Random Forest + BO	<b>0.8930</b>	0.9570	<b>0.8911</b>

Performa Kedua model menunjukkan performa yang sangat tinggi dan sebanding. Rata-rata *Accuracy* dan *F1-Score* berada di sekitar 89%, dan AUC di atas 95%. Ini menunjukkan bahwa data fitur yang digunakan (demografi dan ICD) sangat prediktif terhadap risiko kesehatan jemaah.

Random Forest (RF) sedikit unggul dalam metrik *Accuracy* (+0.15%) dan *F1-Score* (+0.15%). *F1-Score*, yang menyeimbangkan *precision* dan *recall*, sering dianggap sebagai metrik yang lebih baik untuk evaluasi keseluruhan pada data yang mungkin tidak seimbang

sedangkan XGBoost unggul tipis dalam metrik *AUC* (+0.27%). Ini menunjukkan XGBoost sedikit lebih baik dalam membedakan secara absolut antara jemaah berisiko dan tidak berisiko.

Kedua model mencapai titik konvergensi (performa puncak) dalam 25 iterasi yang dialokasikan. Model XGBoost mencapai konvergensi di sekitar iterasi ke-11, sementara Random Forest di sekitar iterasi ke-20. Ini menghemat waktu komputasi yang signifikan dibandingkan *Grid Search* tradisional yang mungkin memerlukan ratusan kombinasi.

#### 4.5.2 Analisis Efisiensi Komputasi (*Computation Cost*)

Selain membandingkan efektivitas akurasi, penelitian ini juga menyoroti aspek efisiensi komputasi (*computation cost*) yang krusial untuk implementasi sistem real-time.

##### 1. Efisiensi Konvergensi Pelatihan

Berdasarkan grafik proses iteratif Bayesian Optimization, XGBoost menunjukkan efisiensi pencarian hiperparameter yang lebih unggul. Model ini mencapai titik konvergensi (skor F1 optimal) pada iterasi ke-11 (Gambar 4.2), sedangkan Random Forest baru mencapai stabilitas performa puncak pada iterasi ke-20 (Gambar 4.5). Hal ini mengindikasikan bahwa XGBoost membutuhkan sumber daya komputasi dan waktu pelatihan yang lebih sedikit untuk menemukan konfigurasi terbaiknya.

## 2. Efisiensi Arsitektur Model

Secara arsitektur, XGBoost menggunakan pendekatan boosting (membangun pohon secara sekuensial untuk mengoreksi error) yang dilengkapi dengan mekanisme regularization. Hal ini cenderung menghasilkan model yang lebih ringkas (lean) dengan kedalaman pohon yang lebih terkontrol dibandingkan Random Forest yang menggunakan pendekatan bagging (membangun ratusan pohon yang dalam dan kompleks secara independen). Implikasinya, meskipun Random Forest mudah diparalelisasi saat pelatihan, XGBoost seringkali lebih efisien dari sisi penggunaan memori dan kecepatan prediksi (inference speed) saat digunakan pada aplikasi akhir.

### 4.5.3 Implikasi Penelitian

Hasil penelitian ini memiliki implikasi praktis yang signifikan untuk Embarkasi Haji Balikpapan:

1. **Kelayakan Sistem Prediksi:** Penelitian ini membuktikan secara empiris bahwa pengembangan sistem pendukung keputusan (SPK) berbasis *machine learning* untuk skrining risiko kesehatan jemaah haji sangat layak dilakukan. Model RF, dengan akurasi 89.3%, dapat mengidentifikasi sebagian besar jemaah berisiko tinggi secara otomatis.
2. **Prioritas Intervensi:** Dengan mengidentifikasi jemaah berisiko (Prediksi "Positif (1)"), tim medis dapat memprioritaskan sumber

daya, melakukan pemeriksaan lebih lanjut, dan merancang intervensi kesehatan preventif sebelum keberangkatan.

#### 4.5.4 Keterbatasan dan Arah Penelitian Selanjutnya

Meskipun hasilnya menjanjikan, terdapat beberapa keterbatasan yang perlu dibahas:

1. **Fokus pada *False Negativer*** Dalam konteks medis, *False Negative* (FN) adalah kesalahan paling fatal (jemaah berisiko dibidang sehat). Model XGBoost menunjukkan FN yang sedikit lebih rendah (17) daripada RF (20) dalam *fold* terakhir. Penelitian selanjutnya dapat berfokus pada penyetelan model (misalnya, dengan mengubah *class weights* atau menyesuaikan *threshold* probabilitas) untuk meminimalkan FN, bahkan jika itu harus mengorbankan sedikit *Accuracy* atau *False Positive*.
2. **Generalisasi Model:** Model ini dilatih pada data dari Embarkasi Balikpapan. Diperlukan pengujian lebih lanjut untuk memvalidasi apakah model ini dapat digeneralisasi ke embarkasi lain dengan profil demografi dan kesehatan yang mungkin berbeda.

#### 4.6 Perbandingan dengan Penelitian Terdahulu

Untuk memvalidasi posisi penelitian ini dalam ranah akademik, hasil yang diperoleh dibandingkan dengan beberapa studi terkait yang menggunakan metode atau objek penelitian serupa. Perbandingan ini berfokus pada kinerja algoritma (*performance*), metode optimasi, dan

konteks medis yang digunakan. Meskipun akurasi 89% ini masih di bawah penelitian (Soelistijadi et al., 2025)(99% pada kasus Ginjal), hal ini dapat dijelaskan oleh perbedaan jenis data. Penelitian Soelistijadi menggunakan data hasil laboratorium presisi (Ureum, Kreatinin), sedangkan penelitian ini menggunakan data surveilans administratif yang memiliki tingkat *noise* lebih tinggi. Namun, sebagai alat skrining awal (*pre-screening*) di embarkasi, performa ini sudah sangat memadai untuk memisahkan populasi jemaah berisiko tinggi dari populasi umum secara cepat dan otomatis.

#### 4.6.1 Efektivitas Optimasi Hyperparameter

Penelitian ini menunjukkan bahwa penerapan *Bayesian Optimization* secara signifikan meningkatkan performa model. Hal ini sejalan dengan temuan (Abdurrahman et al., 2022) dalam klasifikasi penyakit diabetes, di mana penerapan *hyperparameter tuning* (Grid Search dan Random Search) mampu meningkatkan akurasi dari 75% menjadi 95%. Lebih lanjut, penelitian oleh (Khurshid et al., 2025) membandingkan XGBoost dengan Grid Search dan Bayesian Optimization untuk prediksi diabetes. Mereka menemukan bahwa XGBoost dengan Bayesian Optimization menghasilkan performa yang sedikit lebih unggul (Akurasi 97.26%, F1-score 95.72%) dibandingkan Grid Search (Akurasi 97.24%). Temuan ini mendukung hasil penelitian ini yang membuktikan bahwa *Bayesian Optimization* adalah metode

yang efisien dan efektif untuk mencapai konvergensi optimal dalam iterasi yang terbatas (kurang dari 25 iterasi), menawarkan alternatif yang lebih efisien dibandingkan Grid Search yang komputasionalnya mahal seperti yang digunakan oleh (Soelistijadi et al., n.d.) pada kasus Penyakit Ginjal Kronis (Akurasi 99,16%).

#### 4.6.2 Perbandingan Algoritma XGBoost vs Random Forest

Dalam penelitian ini, ditemukan bahwa Random Forest (dengan Bayesian Optimization) memiliki performa yang sedikit lebih tinggi atau sebanding dengan XGBoost (RF Akurasi 89,30% vs XGBoost 89,15%). Hal ini berbeda dengan beberapa literatur lain yang menempatkan XGBoost sebagai model yang superior. Sebagai contoh, (Aditya Ananta Wisnu Wardana & Mizwar Rahim, 2024) dalam klasifikasi kesehatan mental menemukan bahwa XGBoost (Akurasi 99,82%) mengungguli Random Forest (99,04%). Demikian pula, (Tomita et al., 2023) dalam diagnosis asma dewasa melaporkan bahwa XGBoost (Akurasi 81%, AUC 0,85) bekerja lebih baik daripada Random Forest (Akurasi 76%, AUC 0,79).

Perbedaan temuan ini menegaskan prinsip *No Free Lunch Theorem* dalam *machine learning*, di mana tidak ada satu algoritma yang selalu unggul untuk setiap jenis dataset. Pada dataset jemaah haji Embarkasi Balikpapan yang digunakan dalam penelitian ini, karakteristik data mungkin lebih cocok ditangani oleh struktur *bagging*

dari Random Forest dibandingkan *boosting* dari XGBoost, meskipun perbedaannya sangat tipis. Namun, XGBoost dalam penelitian ini tetap menunjukkan keunggulan dalam metrik AUC (0.9597), yang konsisten dengan temuan Tomita, dkk. (2023) mengenai kemampuan diskriminatif XGBoost yang kuat pada data medis.



## BAB V

### PENUTUP

Bab ini menyajikan kesimpulan yang ditarik dari hasil penelitian dan pembahasan pada bab sebelumnya, serta merumuskan saran yang relevan untuk implementasi praktis dan pengembangan penelitian selanjutnya.

#### 5.1 Kesimpulan

Merujuk pada rumusan masalah dan hasil pengujian empiris terhadap data jemaah haji Embarkasi Balikpapan tahun keberangkatan 2024, penelitian ini menyimpulkan hal-hal sebagai berikut:

1. Pengembangan model klasifikasi menggunakan algoritma *Random Forest* dan *Extreme Gradient Boosting (XGBoost)* yang diintegrasikan dengan skema *Bayesian Optimization* terbukti secara signifikan mampu meningkatkan performa prediksi tingkat risiko kesehatan jemaah haji. Pendekatan ini berhasil mengatasi keterbatasan metode penalaran *hyperparameter* konvensional (seperti *grid search*) dengan cara mengeksplorasi ruang parameter yang kompleks secara lebih efisien melalui model probabilistik. Hasil visualisasi proses iteratif menunjukkan bahwa kedua model mampu mencapai konvergensi performa optimal dalam jumlah iterasi yang efisien (kurang dari 25 iterasi), menghasilkan model prediksi yang tidak hanya akurat tetapi juga stabil (*robust*) saat divalidasi.

2. Model yang diusulkan, yakni *Random Forest* berbasis *Bayesian Optimization* (BO-RF) dan *Extreme Gradient Boosting* berbasis *Bayesian Optimization* (BOXGB), terbukti sangat efektif dalam mengklasifikasikan tingkat risiko kesehatan jemaah haji, dibuktikan dengan metrik evaluasi yang tinggi:

- a. Model BO-RF menunjukkan efektivitas tertinggi secara keseluruhan dengan performa yang seimbang, mencapai rata-rata Akurasi 89,30% dan F1-Score 89,11%. Tingginya nilai F1-Score menegaskan kemampuan model ini dalam menangani ketidakseimbangan kelas (*imbalanced data*) secara efektif, meminimalisir kesalahan deteksi pada kelompok risiko tinggi.
- b. Model BOXGB menunjukkan efektivitas yang kompetitif dengan rata-rata Akurasi 89,15% dan keunggulan spesifik pada kemampuan diskriminatif pemisahan kelas, yang ditunjukkan oleh nilai AUC-ROC sebesar 95,97%.

Kedua model ini dinilai layak dan andal untuk digunakan sebagai instrumen prediksi dalam konteks kesehatan haji.

## 5.2 Saran

Berdasarkan kesimpulan yang ditarik dan batasan masalah yang membatasi ruang lingkup penelitian ini, diajukan saran-saran sebagai berikut:

### 5.2.1 Saran Akademis (Pengembangan Penelitian Lanjutan)

1. Penelitian ini dibatasi hanya pada Jemaah haji Embarkasi Balikpapan dan data tahun 2024, penelitian selanjutnya sangat disarankan untuk:

- a. Menggunakan dataset *time-series* (data multi-tahun) untuk menangkap tren dinamis kesehatan jemaah yang mungkin berubah antar musim haji.
  - b. Memperluas cakupan populasi dengan menyertakan data dari embarkasi lain di Indonesia. Langkah ini krusial untuk menguji generalisasi model terhadap variabilitas demografis dan epidemiologis yang lebih luas.
2. Penelitian ini membatasi metode pada karakteristik intrinsik algoritma tunggal dengan optimasi Bayesian, disarankan bagi peneliti selanjutnya untuk:
- a. Mengeksplorasi pendekatan *Ensemble Learning* tingkat lanjut seperti teknik *Stacking* atau *Voting*, yang menggabungkan prediksi dari BO-RF dan BOXGB untuk potensi akurasi yang lebih tinggi.
  - b. Melakukan studi komparasi dengan metode optimasi meta-heuristik lain (misalnya *Genetic Algorithm* atau *Particle Swarm Optimization*) untuk memvalidasi efisiensi relatif dari *Bayesian Optimization*.
  - c. Menerapkan teknik penanganan *imbalanced data* di level data (seperti SMOTE atau ADASYN) sebelum proses optimasi model untuk lebih menekan angka *False Negative*.

### 5.2.2 Saran Praktis (Implementasi Operasional)

1. Hasil pengembangan model prediksi (khususnya BO-RF) disarankan untuk diintegrasikan ke dalam sistem informasi kesehatan di Kantor Kesehatan Pelabuhan (KKP) Embarkasi Balikpapan. Model ini dapat difungsikan sebagai modul Sistem Pendukung Keputusan (SPK) cerdas yang memberikan "opini kedua" secara otomatis saat input data pemeriksaan tahap akhir, membantu tim medis dalam melakukan triase prioritas pemantauan jemaah.
2. Menyadari bahwa dataset 2024 mungkin tidak merepresentasikan kondisi masa depan sepenuhnya, disarankan agar instansi pengguna menetapkan protokol pembaruan model secara berkala. Proses *retraining* model dengan data jemaah terbaru setiap tahunnya mutlak diperlukan untuk mempertahankan relevansi dan akurasi prediksi terhadap perubahan pola penyakit jemaah haji di masa mendatang.

## DAFTAR PUSTAKA

- Abdurrahman, G., Oktavianto, H., & Sintawati, M. (2022). Optimasi Algoritma XGBoost Classifier Menggunakan Hyperparameter Gridsearch dan Random Search Pada Klasifikasi Penyakit Diabetes. In *Informatics Journal* (Vol. 7, Issue 3).
- Aditya Ananta Wisnu Wardana, K., & Mizwar Rahim, A. A. (2024). *Analisis Perbandingan Algoritma XGBoost Dan Algoritma Random Forest Untuk Klasifikasi Data Kesehatan Mental*. <https://www.kaggle.com/datasets/bhavikjikadara/mental-health-dataset>
- Afiatuddin, N., Pratiwi, F., Septia, R., & Hendrawan, H. (2024). Evaluation of Data Mining in Heart Failure Disease Classification. *COGITO Smart Journal*, 10(2).
- Alhazmi, H. N. (2024). A Prediction Triage System for Emergency Department During Hajj Period using Machine Learning Models. *IJCSNS International Journal of Computer Science and Network Security*, 24(7), 11. <https://doi.org/10.22937/IJCSNS.2024.24.7.2>
- Armis Sibuea, N., Otniel Ketaren, S., Manurung, J., Nababan, D., Arwina Bangun, H., Studi Magister Kesehatan Masyarakat, P., & Pascasarjana, D. (2025). *Faktor-Faktor Yang Berhubungan Dengan Morbiditas Pada Jemaah Haji Kota Medan Tahun 2024 Factors Associated with Morbidity Among Hajj Pilgrims in Medan City in 2024*. <https://ejournal.medistra.ac.id/index.php/JKG>

- Budholiya, K., Shrivastava, S. K., & Sharma, V. (2022). An optimized XGBoost based diagnostic system for effective prediction of heart disease. *Journal of King Saud University - Computer and Information Sciences*, 34(7), 4514–4523. <https://doi.org/10.1016/j.jksuci.2020.10.013>
- Buslim, N., Zulfiandri, & Lee, K. (2023). Ensemble learning techniques to improve the accuracy of predictive model performance in the scholarship selection process. *Journal of Applied Data Sciences*, 4(3), 264–275. <https://doi.org/10.47738/jads.v4i3.112>
- Frazier, P. I. (2018). *A Tutorial on Bayesian Optimization*. <http://arxiv.org/abs/1807.02811>
- Gao, L., & Ding, Y. (2020). Disease prediction via Bayesian hyperparameter optimization and ensemble learning. *BMC Research Notes*, 13(1). <https://doi.org/10.1186/s13104-020-05050-0>
- Garnett, R. (2023). Introduction. In *Bayesian Optimization* (pp. 1–14). Cambridge University Press. <https://doi.org/10.1017/9781108348973.002>
- Jaya Kusuma, E., Nurmandhani, R., Aryani, L., Pantuwati, I., Fajar Shidik, G., Dian Nuswantoro, U., & Korespondensi, P. (2025). *OPTIMASI MODEL EXTREME GRADIENT BOOSTING DALAM UPAYA PENENTUAN TINGKAT RISIKO PADA IBU HAMIL BERBASIS BAYESIAN OPTIMIZATION (BOXGB) MACHINE LEARNING OPTIMIZATION IN DETERMINING THE MATERNAL RISK LEVEL BASED ON BAYESIAN OPTIMIZATION*. 12(1). <https://doi.org/10.25126/jtiik.2025129001>

- Khurshid, M. R., Manzoor, S., Sadiq, T., Hussain, L., Khan, M. S., & Dutta, A. K. (2025). Unveiling diabetes onset: Optimized XGBoost with Bayesian optimization for enhanced prediction. *PLoS ONE*, 20(1 January). <https://doi.org/10.1371/journal.pone.0310218>
- Long, P. D., & Dung, N. T. (2024). *Proceedings of the 5th International Conference on Geotechnics for Sustainable Infrastructure Development: GEOTEC2023; 14–15 Dec; Hanoi, Vietnam*. Springer Nature Singapore. <https://books.google.co.id/books?id=vpkTEQAAQBAJ>
- Murdiansyah, D. T. (2024). Prediksi Stroke Menggunakan Extreme Gradient Boosting. *JIKO (Jurnal Informatika Dan Komputer)*, 8(2), 419. <https://doi.org/10.26798/jiko.v8i2.1295>
- Nguyen, V. (2019). Bayesian optimization for accelerating hyper-parameter tuning. *Proceedings - IEEE 2nd International Conference on Artificial Intelligence and Knowledge Engineering, AIKE 2019*, 302–305. <https://doi.org/10.1109/AIKE.2019.00060>
- Rodriguez, P., Bautista, M. A., González, J., & Escalera, S. (2018). *Beyond One-hot Encoding: lower dimensional target embedding*. <https://doi.org/10.1016/j.imavis.2018.04.004>
- Rong, G., Alu, S., Li, K., Su, Y., Zhang, J., Zhang, Y., & Li, T. (2020). Rainfall induced landslide susceptibility mapping based on bayesian optimized random forest and gradient boosting decision tree models—a case study of shuicheng county, china. *Water (Switzerland)*, 12(11), 1–22. <https://doi.org/10.3390/w12113066>

- Soelistjadi, R., Dwiati Wismarini, T., Eniyati, S., & 14, S. (n.d.). *Pemodelan Prediktif Menggunakan Metode Ensemble Learning XGBoost dalam Peningkatan Akurasi Klasifikasi Penyakit Ginjal* (Vol. 5, Issue 4).
- Tomita, K., Yamasaki, A., Katou, R., Ikeuchi, T., Touge, H., Sano, H., & Tohda, Y. (2023). Construction of a Diagnostic Algorithm for Diagnosis of Adult Asthma Using Machine Learning with Random Forest and XGBoost. *Diagnostics*, *13*(19). <https://doi.org/10.3390/diagnostics13193069>
- Wang, S., Zhuang, J., Zheng, J., Fan, H., Kong, J., & Zhan, J. (2021). Application of Bayesian Hyperparameter Optimized Random Forest and XGBoost Model for Landslide Susceptibility Mapping. *Frontiers in Earth Science*, *9*. <https://doi.org/10.3389/feart.2021.712240>
- Xu, W., Zhang, J., Zhang, Q., & Wei, X. (2017). Risk prediction of type II diabetes based on random forest model. *2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, 382–386. <https://doi.org/10.1109/AEEICB.2017.7972337>
- Yang, C., Wang, Y., Zhang, A., Fan, H., & Guo, L. (2023). A Random Forest Algorithm Combined with Bayesian Optimization for Atmospheric Duct Estimation. *Remote Sensing*, *15*(17). <https://doi.org/10.3390/rs15174296>
- Yang, J., & Guan, J. (2022). A Heart Disease Prediction Model Based on Feature Optimization and Smote-Xgboost Algorithm. *Information (Switzerland)*, *13*(10). <https://doi.org/10.3390/info13100475>