



TURNAMEN SAINS DATA NASIONAL 2024

PTM-PRe: Aplikasi Prediksi Penyakit Tidak Menular dan Rekomendasi Personal Berbasis AI

STG Finite

Ketua Pelaksana Turnamen Sains Data Nasional 2024

DISPONSORI OLEH



Asosiasi
Data Sains dan AI
Indonesia



DIDUKUNG OLEH



DIDUKUNG OLEH



UNIVERSITAS TRISAKTI



SWISS GERMAN UNIVERSITY



Universitas Bunda Mulia



UNIVERSITAS KATOLIK INDONESIA

ATMA JAYA

Tepercaya Kualitas Lulusannya



Universitas Al Azhar Indonesia



SEKOLAH TINGGI ILMU MANAJEMEN & ILMU KOMPUTER



IAIN
METRO



ITERA



STMIK
PRIMAKARA



TANRI ABENG
UNIVERSITY



UNIVERSITAS
BAKRIE



ISLAM
INDONESIA

UNIVERSITAS
ISLAM
INDONESIA



UNIVERSITAS
MIKROSKIL



Universitas
Pembangunan Jaya



UNIVERSITAS
NASIONAL
PIONIR PERUBAHAN



UNIVERSITAS
TEKNOLOGI
SUMBAWA

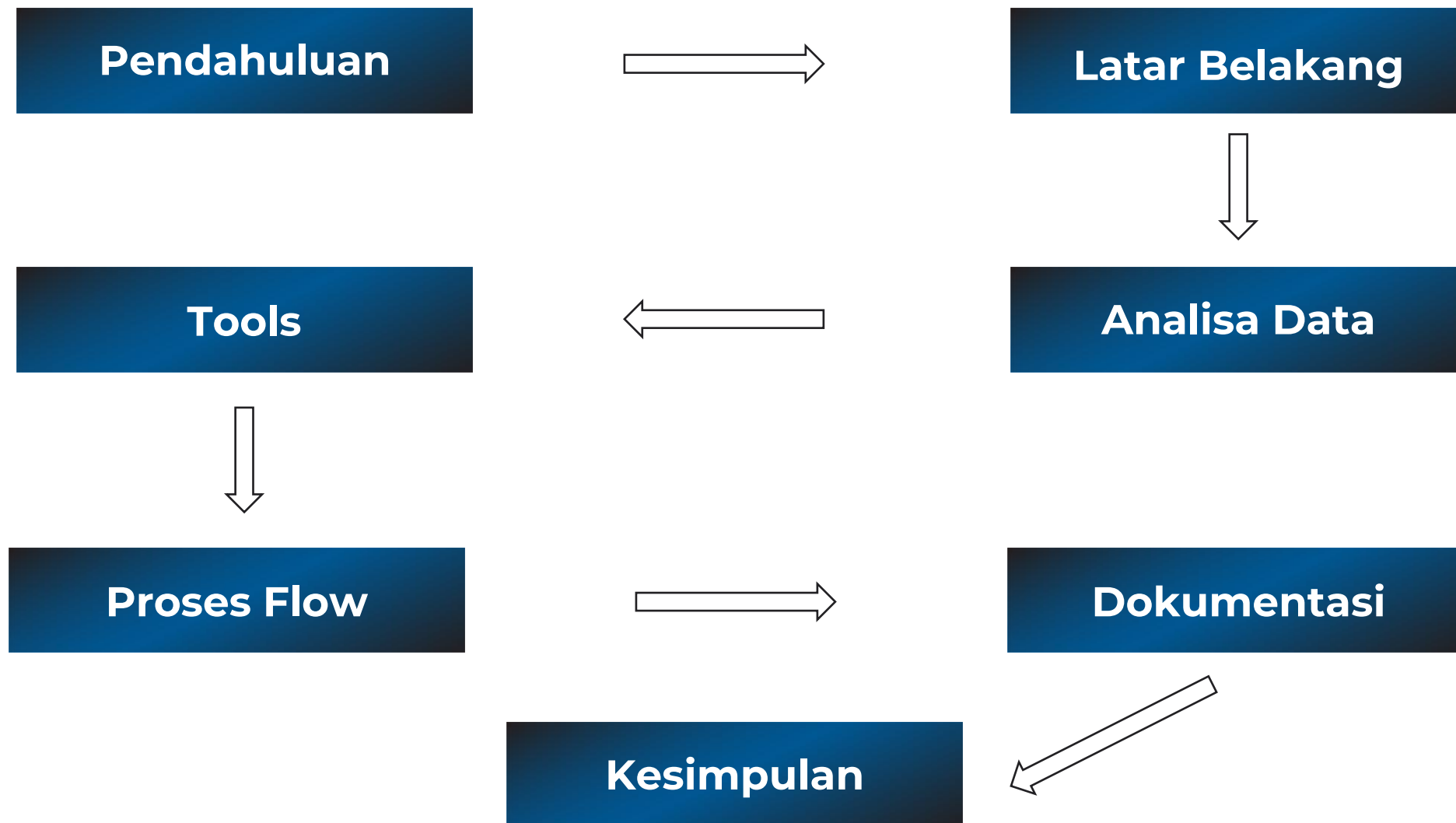


IDB BALI
INSTITUT DESAIN & BISNIS BALI

DIDUKUNG OLEH



Daftar Isi



Informasi Tim



STG Finite

Yogyakarta State University



ML Engineer

Billiarto Sastro Cemerson

Hobi: Futsal, Menulis

Instagram: @billycmrsn



Data Science

Hajran Azbytama Winarya

Hobi: Nonton Film, Membaca

Instagram: @havanjr



Project Manager

Muhammad Goldy Wahyu H.

Hobi: Membaca

Instagram: @mhmmmdgoldy_



Data Science

Risky Anugrah

Hobi: Membaca

Instagram: @risky_zc_

Laporan Hasil Pekerjaan

Pendahuluan



Kesehatan masyarakat adalah salah satu **pilar penting** dalam **pembangunan berkelanjutan (SDGs)** di Indonesia.

(SDGs Poin 3)

Terdapat **enam pilar transformasi** sebagai landasan untuk **memperkuat Sistem Kesehatan**.

Permenkes No. 15 Tahun 2022

Tangguh

Terjangkau

Responsif

Khususnya dalam penanganan **penyakit tidak menular** (PTM) yang meningkat dalam beberapa tahun terakhir.

Penggunaan teknologi canggih seperti **Artificial Intelligence (AI)** dan **Machine Learning (ML)** dalam sistem kesehatan menawarkan berbagai potensi, mulai dari **peningkatan akurasi prediksi** (Gunawan et. al., 2020) hingga **otomatisasi rekomendasi** pengobatan yang **dipersonalisasi** (Hasfera, 2023)

(Gunawan et. al., 2020)
(Hasfera, 2023)

Salah satu solusi inovatif adalah implementasi **Predictive Analytics for Disease Diagnosis** dan **Personalized Medicine**.

Latar Belakang



Pilihan topik use case ini didasarkan pada kondisi kesehatan di Indonesia, di mana **penyakit tidak menular (PTM)** menjadi tantangan besar bagi layanan kesehatan nasional.

Berdasarkan data “Share of Death by Broad Cause in Indonesia 2021 oleh WHO”, angka **kematian akibat PTM** mencapai **9,2 ribu** dibandingkan dengan penyakit menular **6,1 ribu** dan cedera **0,6 ribu**.

Share of deaths by broad cause

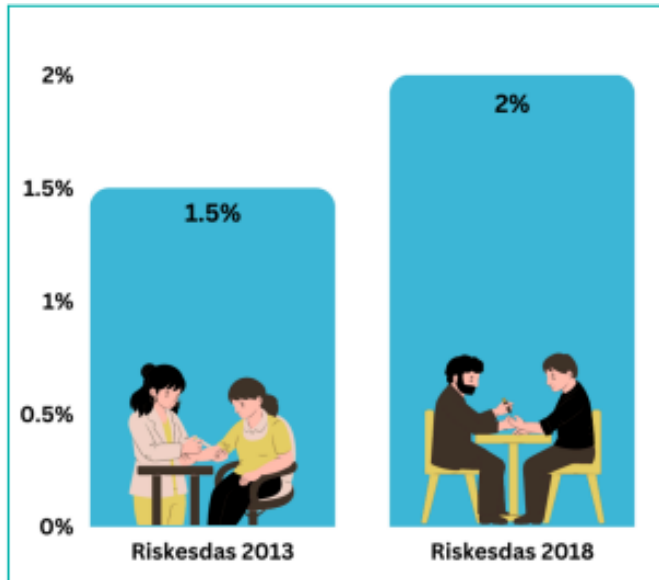
Indonesia, 2021

Cause of Death ▲		% ▼
Noncommunicable diseases	9.2k	<div></div>
Communicable, maternal, perinatal and nutritional conditions	6.1k	<div></div>
Injuries	0.6k	<div></div>

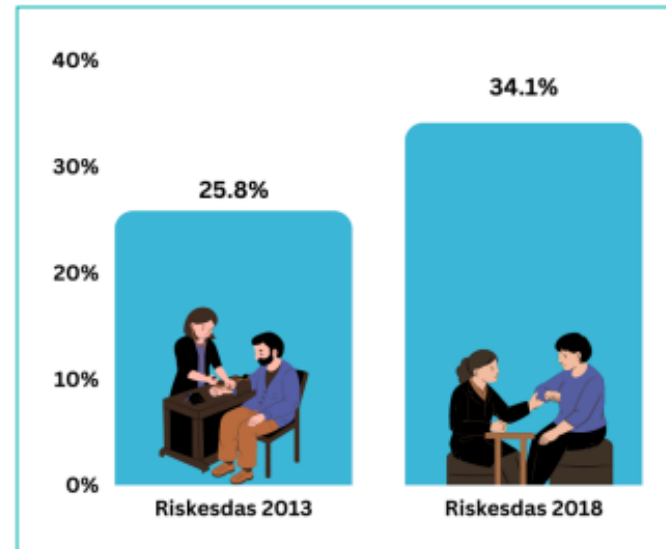
Source: data.who.int

Latar Belakang

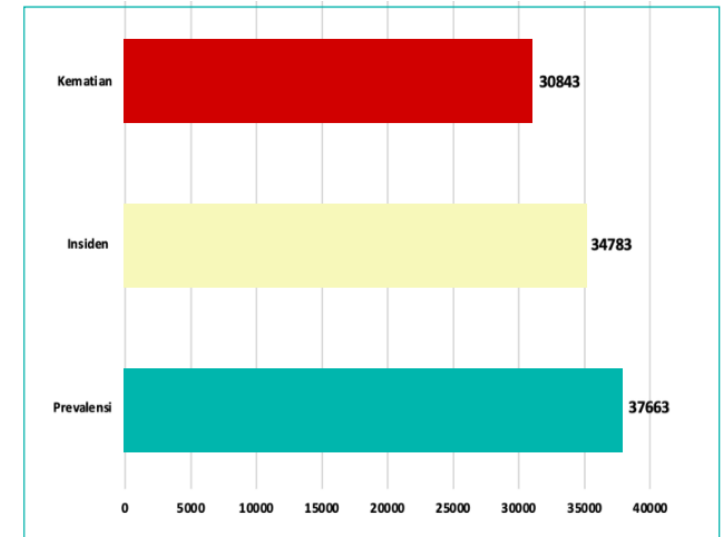
Prevalensi Penyakit Tidak Menular



Hipertensi



Diabetes



Kanker Paru-Paru

Source: Annual Report 2022 Prevention and Control of Non-Communicable Diseases Kemenkes

Prevalensi PTM seperti **hipertensi**, **diabetes**, dan **kanker paru-paru** terus meningkat, terbukti dari survei Riskesdas yang menunjukkan **peningkatan signifikan** dari 2013 hingga 2018.

Latar Belakang



Solusi dalam use case ini adalah sistem prediktif berbasis **AI** yang memanfaatkan **Machine Learning** (ML) dan **Retrieval-Augmented Generation** (RAG) untuk deteksi penyakit dan pemberian rekomendasi yang meliputi pengobatan, pola hidup sehat, serta penanganan lanjutan.

Gender: Perempuan
Umur: 25
Tekanan Darah: 140/60

Predict

Langkah konkret ini dirancang khusus untuk **petugas kesehatan** guna mendukung mereka dalam mengambil keputusan yang **lebih cepat dan akurat**.



Dengan teknologi **Machine learning** dan **Retrieval Augmented Generation** (RAG), sistem ini diharapkan mampu memberikan prediksi diagnosis yang **akurat** serta **rekomendasi** yang lebih disesuaikan dengan **kebutuhan pasien**, sehingga meningkatkan **kualitas** dan **efisiensi** layanan kesehatan di Indonesia.



Analisa Data



Data Prediksi

Data Penyakit Hipertensi

Survey BRFSS 2015

<https://www.kaggle.com/datasets/prosperchuks/health-dataset/data>

Data Penyakit Diabetes

Pima Indians Diabetes Database

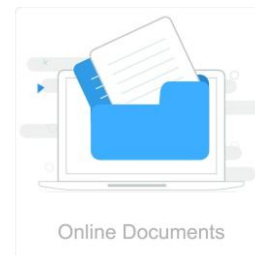
<https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database/data>

Data Penyakit Kanker Paru-Paru

Lung Cancer Dataset

<https://www.kaggle.com/datasets/mysarahmadbhat/lung-cancer>

kaggle



Data RAG

Dokumen Penyakit Hipertensi

Menteri Kesehatan Republik Indonesia
Perhimpunan Dokter Hipertensi Indonesia
Ikatan Dokter Anak Indonesia
Jurnal Artikel Terkait

Dokumen Penyakit Diabetes

Menteri Kesehatan Republik Indonesia
Ikatan Dokter Anak Indonesia
Perkumpulan Endokrinologi Indonesia
Jurnal Artikel Terkait

Dokumen Penyakit Kanker Paru-Paru

Menteri Kesehatan Republik Indonesia
Perhimpunan Dokter Paru Indonesia
Komite Penanggulangan Kanker Nasional
Jurnal Artikel Terkait

Tools

Prediction



Retrieval Augmented Generation (RAG)



LangChain

Gemini

Streamlit

Tools

Why Use?



Interactive Environment

Documented Workflow

Support fo Multiple Libraries



High Performance & Accuracy

Robust to Overfitting

Scalable for Large Dataset



Efficient Model Serialization

Simple to Save & Load

Preserves Model Structure & Performance

Tools



Why Use?



LangChain

Gemini



Streamlit



Flexible Framework

Modular & Scalable

Optimized for LLM Pipeline



High Quality Embedding

Powerful LLM & Free

Efficient Performance



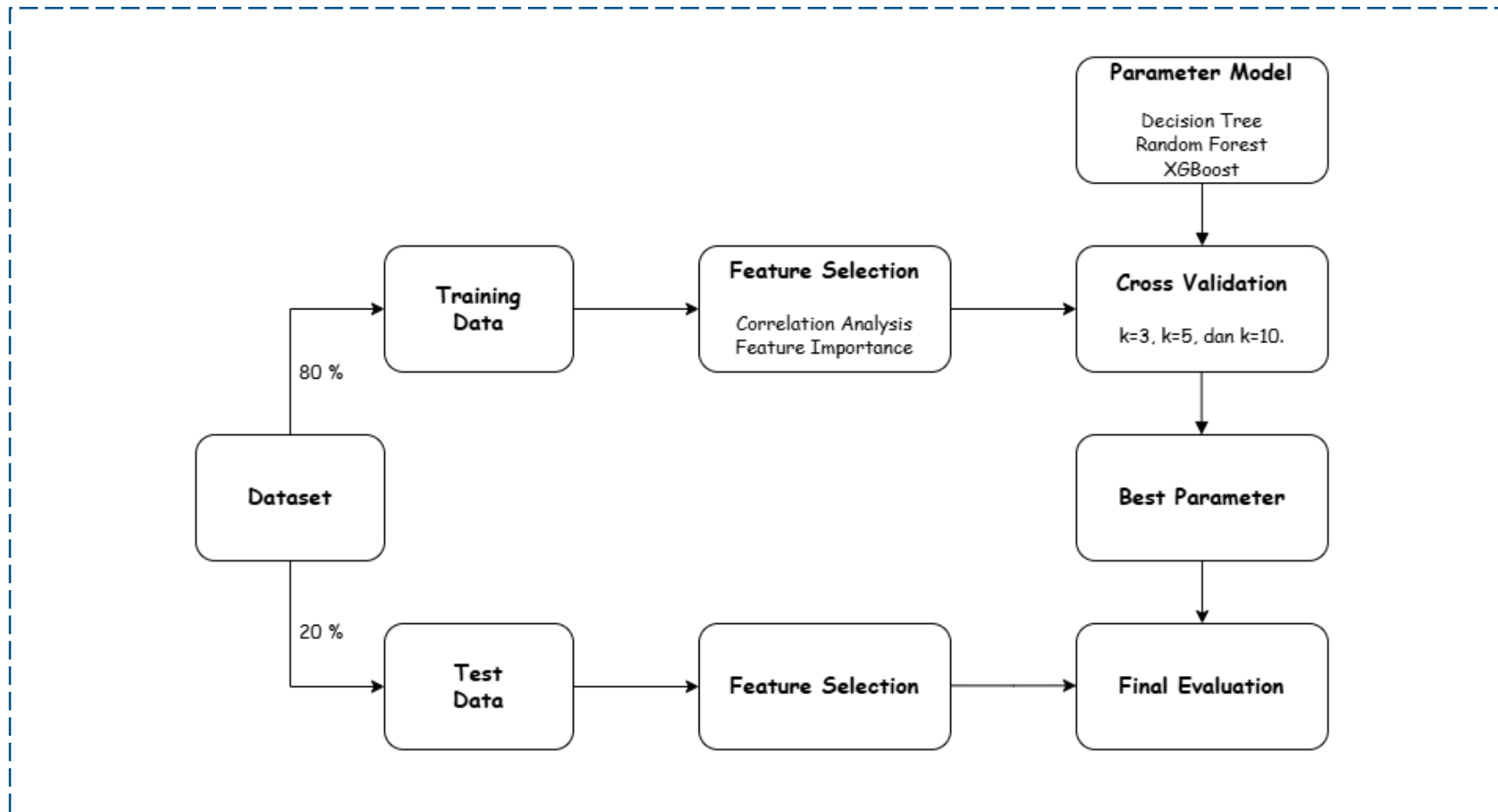
User-Friendly Interface

Fast Prototyping

*Interactive Widgets
for Seamless User Experience*

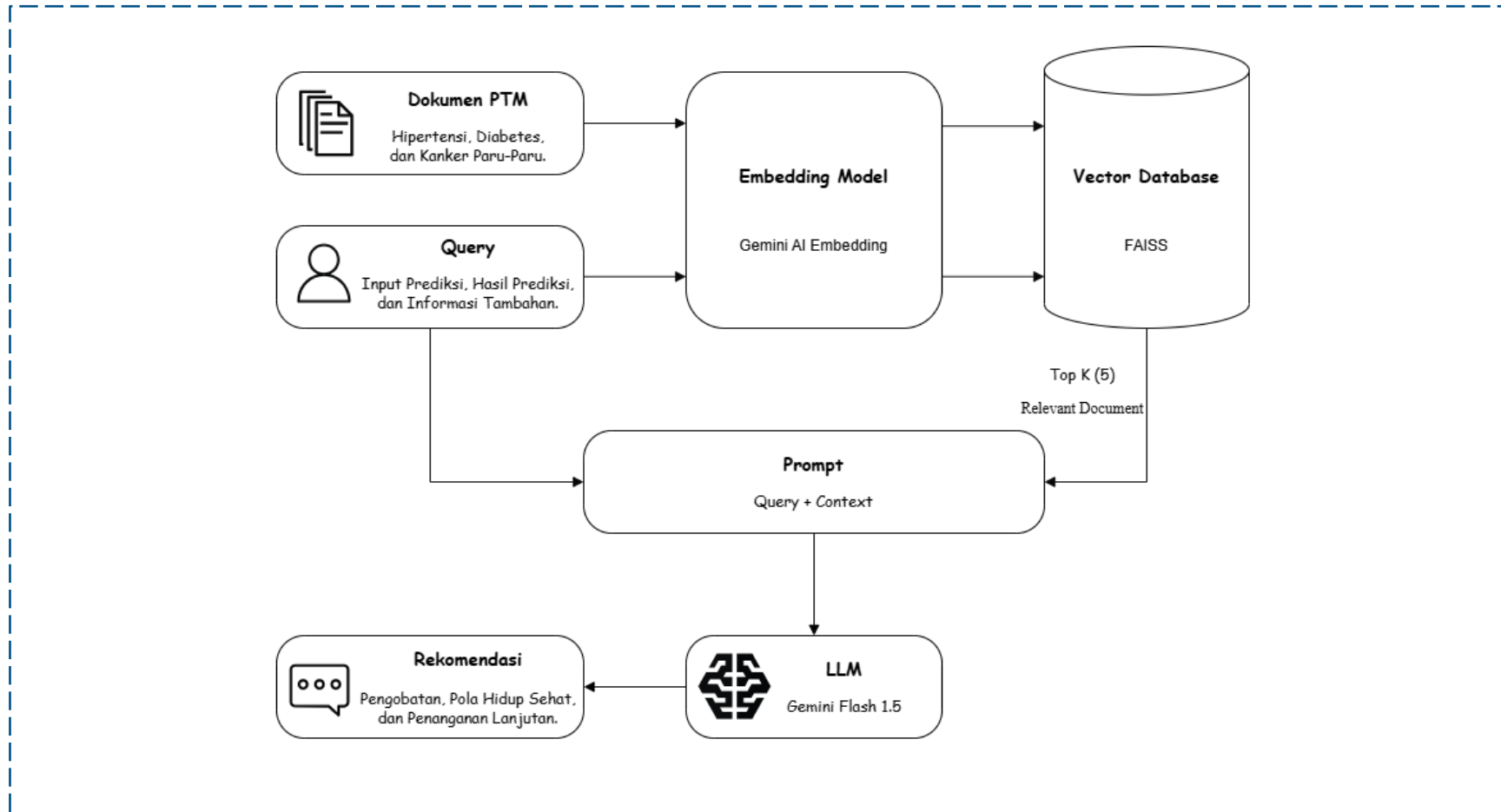
Proses Flow

Prediction



Proses Flow

Retrieval Augmented Generation (RAG)



Evaluasi Model Prediksi

	Precision	Recall	Accuracy	ROC-AUC
Decision Tree	0,78	0,83	0,89	0,84
Random Forest	0,78	0,78	0,89	0,93
XGBoost	0,75	0,77	0,88	0,91

Kanker Paru Paru

	Precision	Recall	Accuracy	ROC-AUC
Decision Tree	0,65	0,66	0,68	0,65
Random Forest	0,70	0,71	0,73	0,82
XGBoost	0,71	0,72	0,73	0,79

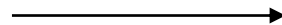
Diabetes

	Precision	Recall	Accuracy	ROC-AUC
Decision Tree	1,00	1,00	1,00	1,00
Random Forest	1,00	1,00	1,00	1,00
XGBoost	1,00	1,00	1,00	1,00

Hipertensi

Evaluasi Model Prediksi (Cross Validation)

Hipertensi



	ROC-AUC
k=3	1,0000
k=5	1,0000
k=10	1,0000

Model k=10 dipilih

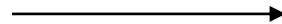
Diabetes



	ROC-AUC
k=3	0,8910
k=5	0,8939
k=10	0,8905

Model k=5 dipilih

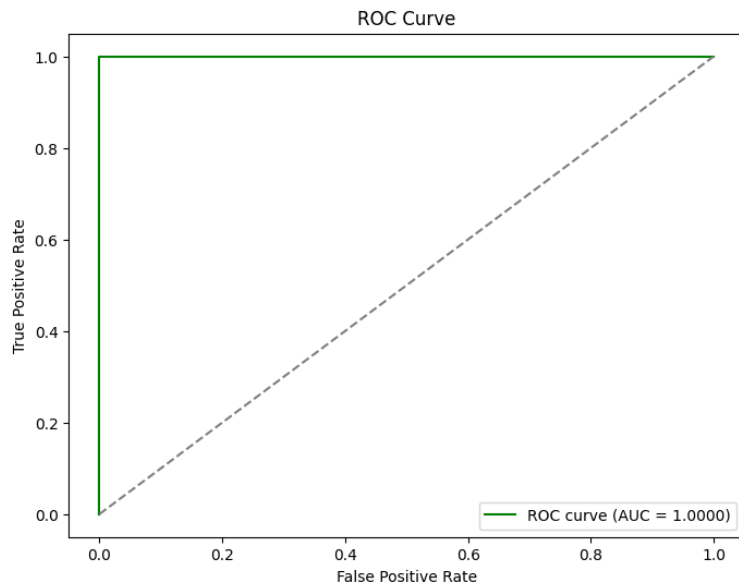
Kanker Paru Paru



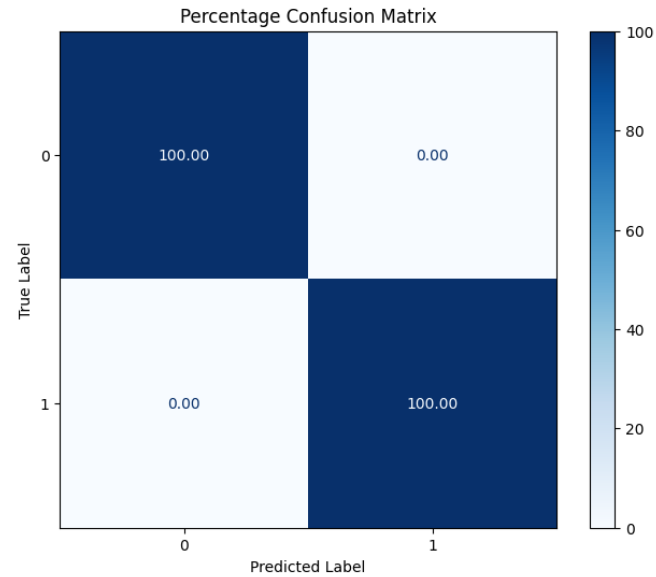
	ROC-AUC
k=3	0,9773
k=5	0,9786
k=10	0,9809

Model k=10 dipilih

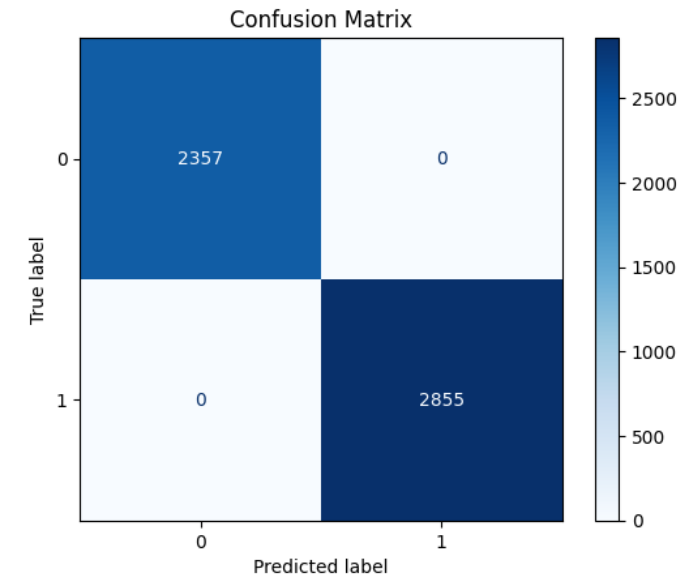
Hasil Klasifikasi Hipertensi



ROC Curve

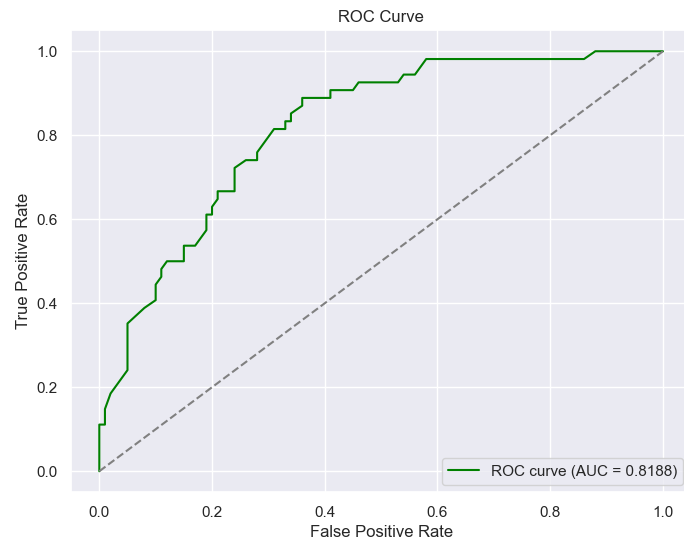


Confusion Matrix
(dalam persenan kasus)

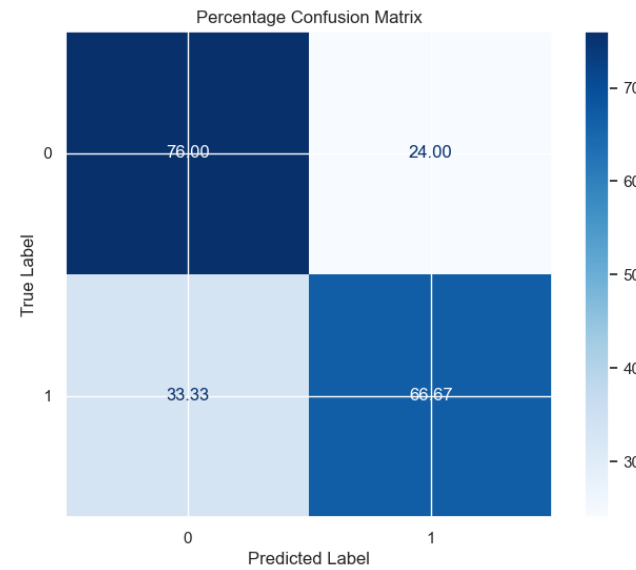


Confusion Matrix
(dalam jumlah kasus)

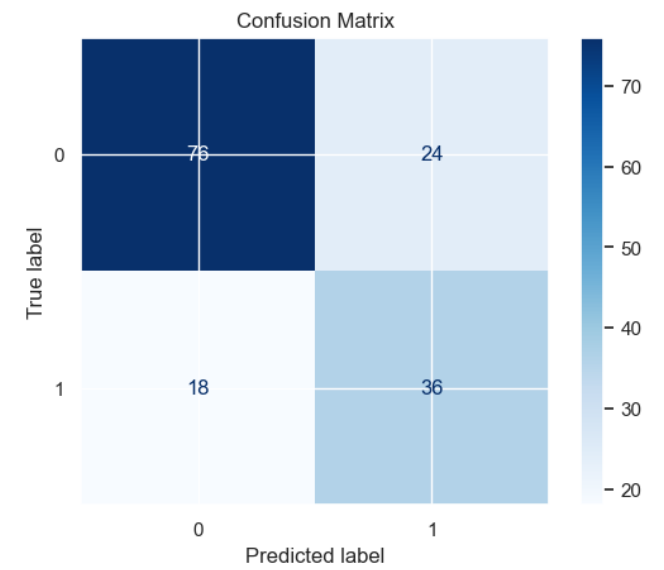
Hasil Klasifikasi Diabetes



ROC Curve

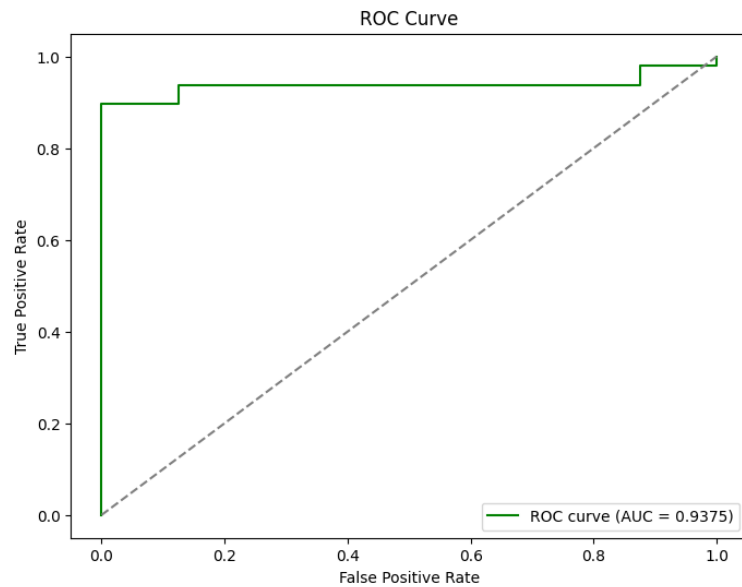


Confusion Matrix
(dalam persenan kasus)

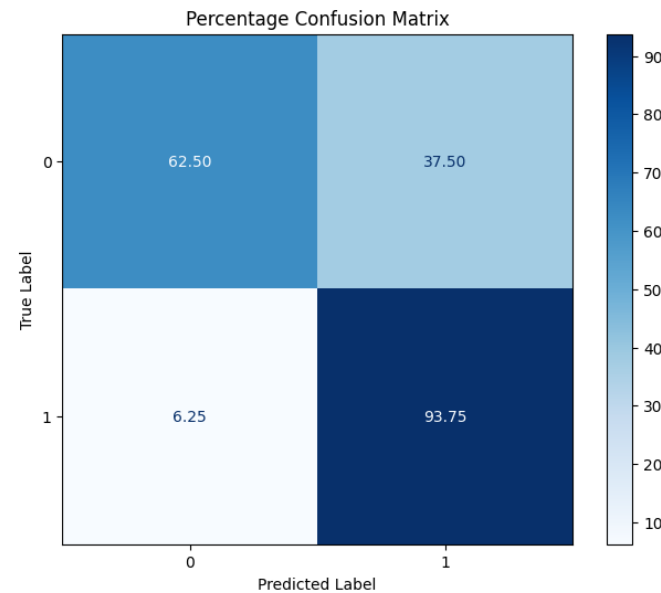


Confusion Matrix
(dalam jumlah kasus)

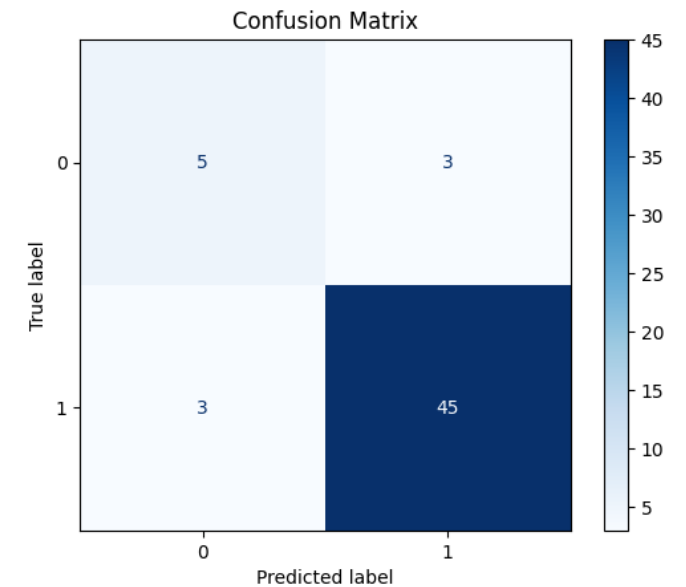
Hasil Klasifikasi Kanker Paru Paru



ROC Curve



Confusion Matrix
(dalam persenan kasus)



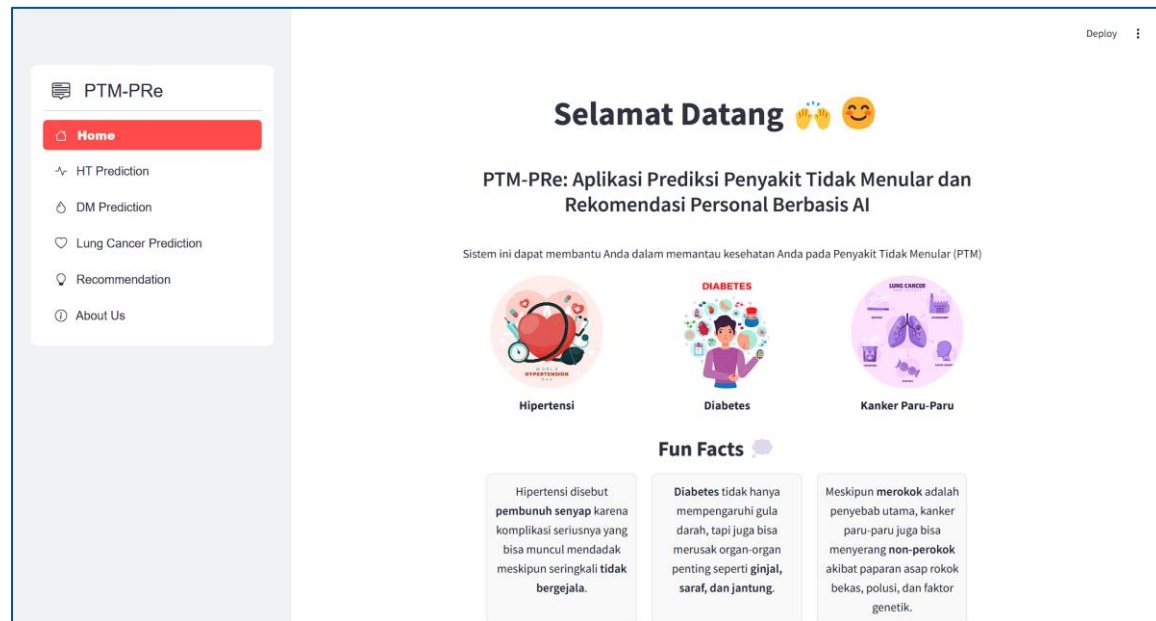
Confusion Matrix
(dalam jumlah kasus)

Evaluasi Model RAG (Rekomendasi Pengobatan)

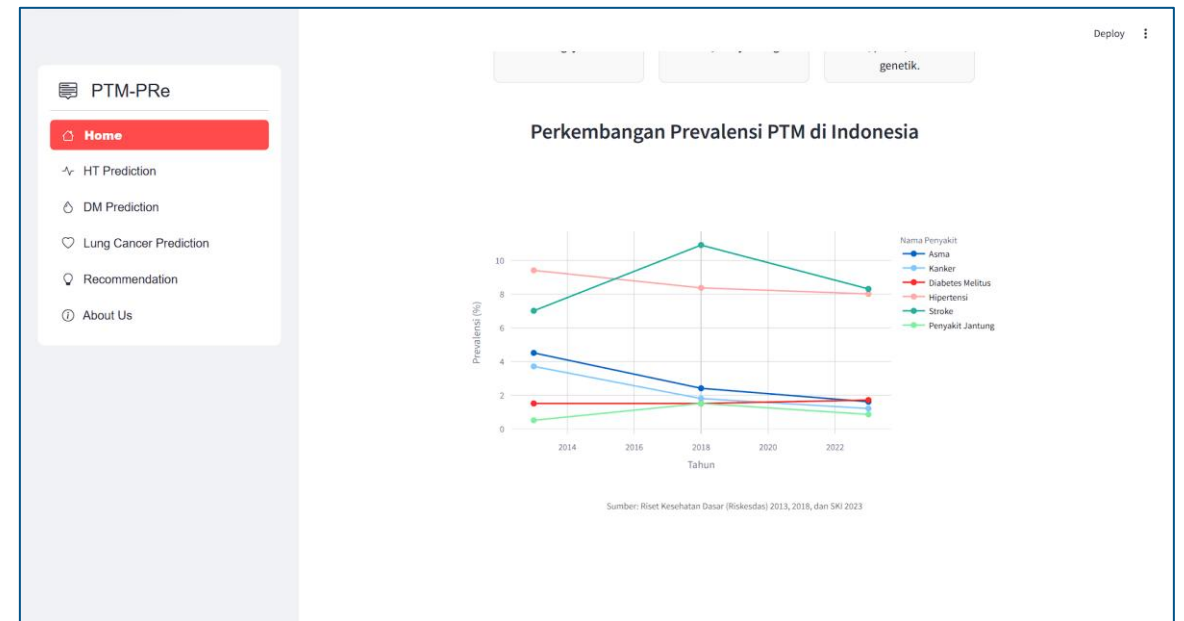
Penyakit	Expected Answer	RAG Answer	Evaluation	Respon Time
Hipertensi	Diuretik, Beta-blocker, ACE inhibitor.	ACE inhibitor, Beta-blocker	Sebagian Tepat (66,7%).	4,03 s.
Diabetes	Agonis reseptor GLP-1, Inhibitor SGLT-2, Metformin, Insulin, Sulfonilurea.	Insulin, Metformin, Sulfonilurea	SebagianTepat (60%).	4,26 s.
Kanker Paru-Paru	Operasi, Kemoterapi, Terapi target, Imunoterapi, Radioterapi, Pereda nyeri.	Kemoterapi, Terapi target, Imunoterapi, Radioterapi, Pereda nyeri	Hampir Lengkap (83,3%).	4,12 s.

Evaluasi dilakukan langsung **tanpa RAGAS** dan sejenisnya, karena metode tersebut hanya mendukung embedding dan LLM dari OpenAI.

Hasil Web App



Tampilan Home (bagian 1)



Tampilan Home (bagian 2)

Dokumentasi



Hasil Web App

PTM-Pre

Home

HT Prediction

DM Prediction

Lung Cancer Prediction

Recommendation

About Us

Prediksi Hipertensi

Tipe Nyeri Data: Tanpa Gejala

Detak Jantung Maksimum: 50.00

Angina Dipicu Oleh Olahraga: Tidak

Depresi ST Akibat Olahraga: 0.00

Kemiringan segemen ST: Menanjak

Jumlah Pembuluh Utama: 0

Thalassemia: Normal

Prediksi

Tampilan HT Prediction

PTM-Pre

Home

HT Prediction

DM Prediction

Lung Cancer Prediction

Recommendation

About Us

Prediksi Hipertensi

Tipe Nyeri Data: Typical Angina

Detak Jantung Maksimum: 140.00

Angina Dipicu Oleh Olahraga: Tidak

Depresi ST Akibat Olahraga: 2.00

Kemiringan segemen ST: Menanjak

Jumlah Pembuluh Utama: 2

Thalassemia: Cacat

Prediksi

Probabilitas Negatif: 0.6

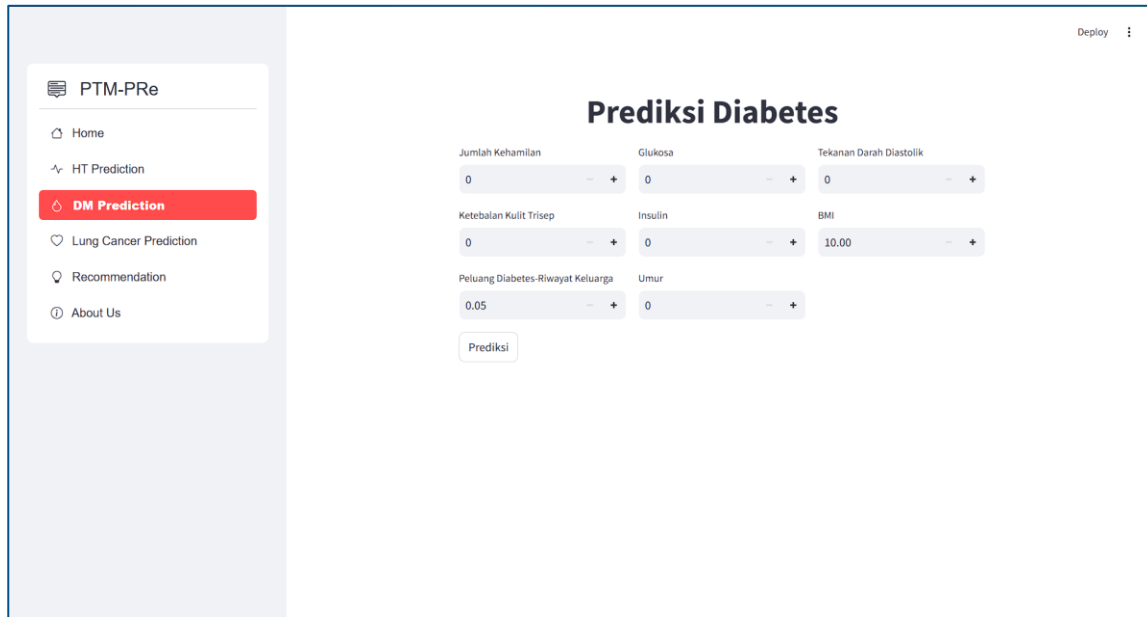
Probabilitas Positif: 0.4

Hasil Prediksi: Negatif

Dapatkan Rekomendasi

Tampilan HT Prediction
setelah prediksi

Hasil Web App



PTM-PRe

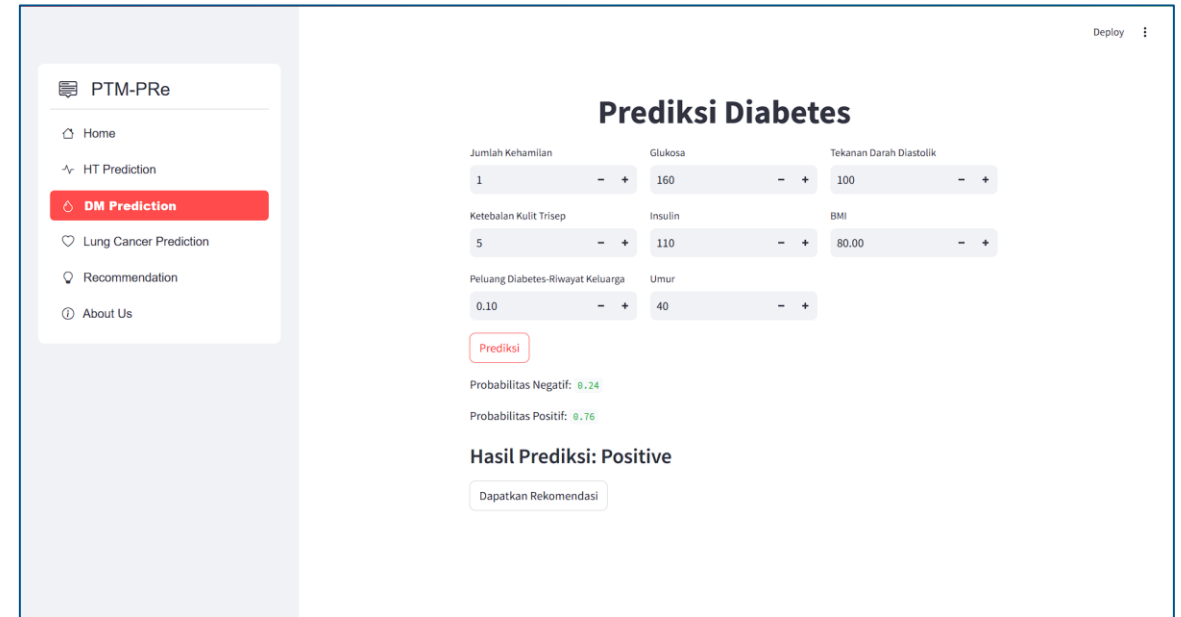
- Home
- HT Prediction
- DM Prediction**
- Lung Cancer Prediction
- Recommendation
- About Us

Prediksi Diabetes

Jumlah Kehamilan	Glukosa	Tekanan Darah Diastolik
0	0	0
Ketebalan Kulit Trisep	Insulin	BMI
0	0	10.00
Peluang Diabetes-Riwayat Keluarga	Umur	
0.05	0	

Prediksi

Tampilan DM Prediction



PTM-PRe

- Home
- HT Prediction
- DM Prediction**
- Lung Cancer Prediction
- Recommendation
- About Us

Prediksi Diabetes

Jumlah Kehamilan	Glukosa	Tekanan Darah Diastolik
1	160	100
Ketebalan Kulit Trisep	Insulin	BMI
5	110	80.00
Peluang Diabetes-Riwayat Keluarga	Umur	
0.10	40	

Prediksi

Probabilitas Negatif: 0.24

Probabilitas Positif: 0.76

Hasil Prediksi: Positive

Dapatkan Rekomendasi

Tampilan DM Prediction
setelah prediksi

Dokumentasi



Hasil Web App

PTM-PRe

- Home
- HT Prediction
- DM Prediction
- Lung Cancer Prediction**
- Recommendation
- About Us

Prediksi Kanker Paru-Paru

Jenis Kelamin	Umur	Merokok
Wanita	0	Tidak
Jari Kuning	Mengalami Kecemasan	Tekanan Teman
Tidak	Tidak	Tidak
Memiliki Penyakit Kronis?	Mengalami Kelelahan?	Memiliki Alergi?
Tidak	Tidak	Tidak

Prediksi

Tampilan Lung Cancer Prediction

PTM-PRe

- Home
- HT Prediction
- DM Prediction
- Lung Cancer Prediction**
- Recommendation
- About Us

Prediksi Kanker Paru-Paru

Jenis Kelamin	Umur	Merokok
Pria	50	Ya
Jari Kuning	Mengalami Kecemasan	Tekanan Teman
Tidak	Ya	Ya
Memiliki Penyakit Kronis?	Mengalami Kelelahan?	Memiliki Alergi?
Ya	Ya	Ya

Prediksi

Probabilitas Negatif: 0.03

Probabilitas Positif: 0.97

Hasil Prediksi: Positif

Dapatkan Rekomendasi

Tampilan Lung Cancer Prediction setelah prediksi

Hasil Web App

PTM-PRe

Home

HT Prediction

DM Prediction

Lung Cancer Prediction

Recommendation

About Us

Deploy

Rekomendasi Berbasis RAG

Pilih Penyakit untuk Rekomendasi:

Kanker Paru-Paru

Rekomendasi untuk Kanker Paru-Paru

	GENDER	AGE	SMOKING	YELLOW_FINGERS	ANXIETY	PEER_PRESSURE	CHRONIC_DISEASE	FAT
	0	1	50	2	1	2	2	2

Informasi Tambahan

Memiliki alergi kacang, telur, daging, dan nasi

Pilih Jenis Rekomendasi:

☒ Rekomendasi Pengobatan

☐ Rekomendasi Pola Hidup Sehat

☐ Rekomendasi Tindak Lanjut

Dapatkan Rekomendasi

Tampilan Rekomendasi

PTM-PRe

Home

HT Prediction

DM Prediction

Lung Cancer Prediction

Recommendation

About Us

Deploy

Pilih Jenis Rekomendasi:

☒ Rekomendasi Pengobatan

☐ Rekomendasi Pola Hidup Sehat

☐ Rekomendasi Tindak Lanjut

Dapatkan Rekomendasi

5.422456582914429

Rekomendasi Rekomendasi Pengobatan: ## Rekomendasi Pengobatan untuk Pasien Kanker Paru-Paru

Informasi Pasien:

- Jenis Kelamin: Laki-laki (GENDER: 1)
- Usia: 50 tahun (AGE: 50)
- Merokok: Ya (SMOKING: 2)
- Alergi: Kacang, telur, daging, nasi
- Hasil Prediksi: Positif

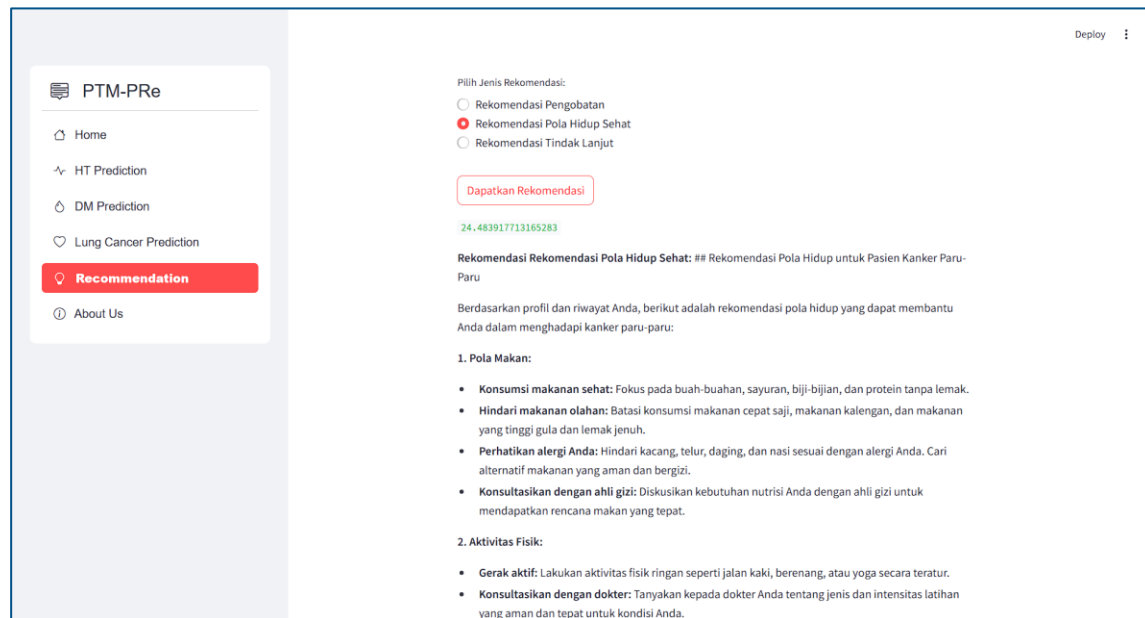
Rekomendasi Pengobatan:

1. Obat yang Disarankan:

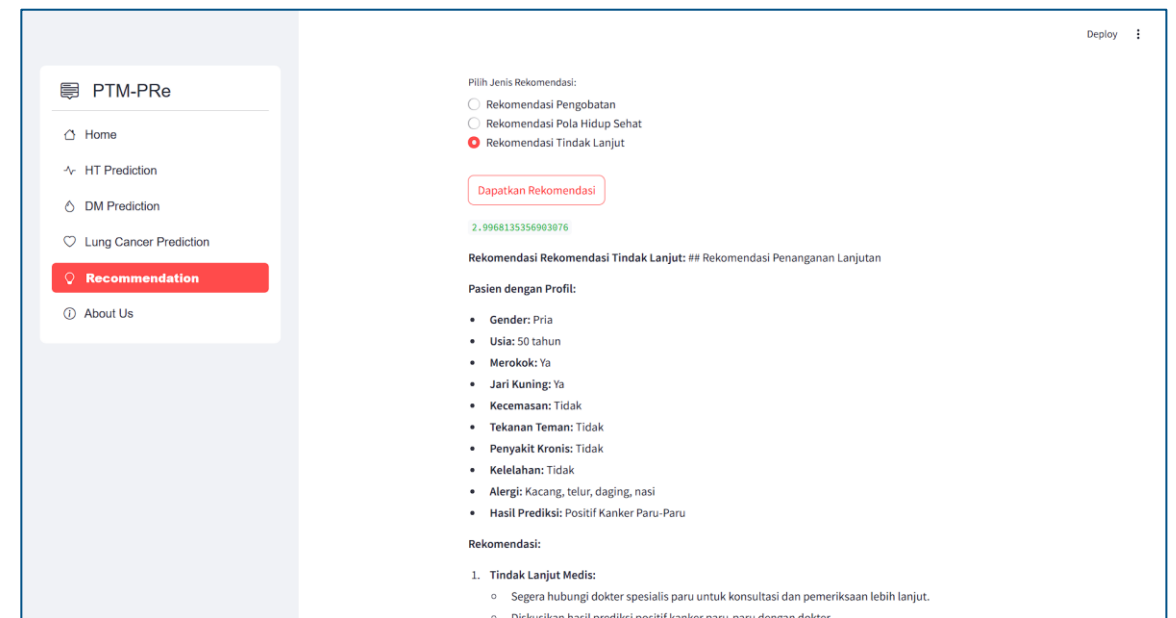
- Kemoterapi:
 - Dosis dan Jenis: Dosis dan jenis kemoterapi akan ditentukan berdasarkan stadium kanker, kondisi kesehatan pasien, dan respons terhadap pengobatan.
 - Contoh:
 - Pemetrexed: Dosis dan frekuensi disesuaikan dengan respons pasien.
 - Cisplatin: Dosis dan frekuensi disesuaikan dengan respons pasien.
 - Docetaxel: Dosis dan frekuensi disesuaikan dengan respons pasien.

Tampilan Rekomendasi bagian jenis Rekomendasi Pengobatan

Hasil Web App



Tampilan Rekomendasi bagian jenis Rekomendasi Pola Hidup Sehat



Tampilan Rekomendasi bagian jenis Rekomendasi Tindak Lanjut

Dokumentasi



Demo Web App

Streamlit Web App

<https://ptm-pre.streamlit.app/>



Youtube Video

https://youtu.be/AJOI_F9nKXs



Kesimpulan



Modell Prediction

Berdasarkan hasil evaluasi model yang tercantum pada tabel evaluasi model prediksi, model **Random Forest** menunjukkan kinerja terbaik dibandingkan dengan model lainnya, yaitu **Decision Tree** dan **XGBoost**.

Pada setiap metrik evaluasi yang digunakan, model **Random Forest** memperoleh nilai tertinggi, yang mengindikasikan bahwa model ini memiliki kemampuan klasifikasi yang lebih baik dalam mengidentifikasi pasien yang menderita penyakit dibandingkan kedua model lainnya.

Modell RAG

Rekomendasi yang dihasilkan telah **dibandingkan** dengan panduan resmi, seperti dokumen dari **Kementerian Kesehatan** dan pustaka kesehatan lainnya, menunjukkan kesesuaian yang signifikan.

Waktu respon rata-rata model cepat, yaitu sekitar **4 detik**, sehingga mendukung efisiensi penggunaannya dalam sistem rekomendasi berbasis AI.

Daftar Pustaka

- Bhat, M. A. (2021). *Lung Cancer Dataset*. Kaggle. <https://www.kaggle.com/datasets/mysarahmadbhat/lung-cancer>
- Chuks, P. (2022). *Diabetes, Hypertension and Stroke Prediction*. Kaggle. <https://www.kaggle.com/datasets/prosperchuks/health-dataset/data>
- Gunawan, M. I., Sugiarto, D., & Mardianto, I. (2020). Peningkatan Kinerja Akurasi Prediksi Penyakit Diabetes Mellitus Menggunakan Metode Grid Search pada Algoritma Logistic Regression. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(3), 280. <https://doi.org/10.26418/jp.v6i3.40718>
- Hasfera, I. P. (2023). Artificial Intelligence untuk Layanan Kesehatan Di Perpustakaan Unbrah. *Al-Ma'arif: Jurnal Ilmu Perpustakaan Dan Informasi Islam ISSN 0740-8188 Imca*, 92-108.
- Kementerian Kesehatan RI. (2013). Riset Kesehatan Dasar 2013. In *Laporan Nasional Riskesdas 2013*.
- Kementerian Kesehatan RI. (2018). Riset Kesehatan Dasar 2018. In *Laporan Nasional Riskesdas 2018*.
- Kementerian Kesehatan RI. (2022). *Permenkes No 15 Tahun 2022 tentang Pedoman Penggunaan Dana Dekonsentrasi Kementrian Kesehatan Tahun Anggaran 2022* (Issue 146914).
- Kementerian Kesehatan RI. (2023). Survei Kesehatan Indonesia 2023. In *Laporan Nasional SKI 2023*.
- UCI Machine Learning. (2016). *Pima Indians Diabetes Database*. Kaggle. <https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database/data>



Terima Kasih

