

**Prediksi Risiko Penyakit Jantung Menggunakan Algoritma
Klasifikasi Pada Dataset Medis**



Disusun Oleh :

Faathir El Tasleem

A11.2022.14624

A11.4504

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO
SEMARANG 2024

DESKRIPSI

Proyek ini berfokus pada penggunaan algoritma klasifikasi untuk menentukan risiko penyakit jantung pada pasien berdasarkan data medis seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, dan indikator kesehatan lainnya. Kumpulan data yang digunakan adalah kumpulan data kesehatan yang tersedia di Kaggle dengan tujuan membantu dalam menentukan risiko penyakit jantung dan memberikan panduan kepada tenaga medis untuk mendiagnosis pasien.

MASALAH

Tujuan utama dari proyek ini adalah untuk membuat model prediksi yang akurat untuk mengidentifikasi risiko penyakit jantung menggunakan peralatan medis yang tersedia. Penyebab utama kematian di dunia adalah penyakit jantung, oleh karena itu deteksi dan prediksi yang akurat dapat menjadi sangat penting dalam memeranginya.

TUJUAN

1. Mengembangkan algoritma prediktif yang dapat mengidentifikasi kemungkinan penyakit jantung pada pasien.
2. Menganalisis dan memahami variabel yang paling signifikan dalam menentukan risiko penyakit jantung.
3. Memberikan rekomendasi untuk beberapa prosedur pencegahan atau penanganan berdasarkan hasil klasifikasi.

ALUR/TAHAPAN/KERANGKA EKSPERIMEN

- **Pengumpulan Data:**
 - Mengunduh dan memahami dataset dari Kaggle.
- **Preprocessing Data:**
 - Memanfaatkan data dari data yang tidak normal atau berkualitas rendah.
 - Menstandarisasi atau melakukan normalisasi fitur sesuai kebutuhan model.
- **Pemilihan Algoritma Klasifikasi:**
 - Menentukan algoritma klasifikasi seperti K-Nearest Neighbors.
- **Pelatihan Model:**
 - Melatih model dengan data yang sudah diproses dan memilih parameter terbaik.
- **Evaluasi Model:**
 - Mengevaluasi model menggunakan metrik seperti akurasi, presisi, recall, dan F1-score untuk menilai performa.
- **Interpretasi dan Analisis Hasil:**
 - Menginterpretasikan hasil prediksi dan analisis fitur yang paling signifikan.
- **Visualisasi dan Kesimpulan:**
 - Menyajikan hasil dalam bentuk visual dan menarik kesimpulan terkait prediksi dan variabel penting.

PENJELASAN DATASET

Dataset **Heart Disease Classification Dataset** dari Kaggle mengandung beberapa fitur penting yang digunakan untuk prediksi, antara lain:

- **Age:** Usia pasien.
- **Sex:** Jenis kelamin pasien.
- **Chest Pain Type:** Jenis nyeri dada, yang merupakan salah satu indikator risiko.
- **Resting Blood Pressure:** Tekanan darah pasien saat beristirahat.
- **Cholesterol:** Kadar kolesterol dalam tubuh.
- **Fasting Blood Sugar:** Kadar gula darah puasa.
- **Electrocardiographic Results:** Hasil elektrokardiografi.
- **Maximum Heart Rate Achieved:** Denyut jantung maksimum yang dicapai.
- **Exercise Induced Angina:** Indikator apakah pasien mengalami angina saat berolahraga.
- **Oldpeak:** Depresi ST yang diinduksi oleh olahraga.
- **Slope of the Peak Exercise ST Segment:** Kemiringan segmen ST pada puncak latihan.
- **Number of Major Vessels:** Jumlah pembuluh darah utama.
- **Thal:** Indikator thalassemia.

Timeline Eksperimen

No	Kegiatan	Minggu ke-1	Minggu ke-2	Minggu ke-3	Minggu ke-4	Minggu ke-5	Minggu ke-6	Minggu ke-7
1.	Perencanaan dan Pemahaman Data	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				
2.	Persiapan Data		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				
3.	Permodelan Data			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
4.	Evaluasi Model				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
5.	Visualisasi dan Pelaporan					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>