

Project Plan

III Detect - Cardiovascular Diagnosis Detector

Pernyataan Masalah

Penyakit kardiovaskular merupakan salah satu penyebab utama kematian di dunia dan sering kali tidak terdeteksi sejak dini. Banyak individu tidak menyadari risiko yang mereka miliki karena keterbatasan akses ke pemeriksaan kesehatan rutin atau kurangnya pengetahuan tentang faktor risiko. Di sisi lain, teknologi machine learning memiliki potensi untuk mengidentifikasi risiko secara otomatis berdasarkan data kesehatan individu. Namun, hingga saat ini belum banyak tersedia platform yang memungkinkan masyarakat melakukan skrining risiko penyakit kardiovaskular secara mandiri, akurat, dan mudah digunakan. Oleh karena itu, dibutuhkan pengembangan sistem berbasis website yang mampu memprediksi risiko penyakit kardiovaskular dan memberikan saran pencegahan secara tepat.

Pertanyaan Penelitian

1. Bagaimana penerapan machine learning dan front-end back-end dapat membantu dalam mengidentifikasi individu yang berisiko tinggi terkena penyakit kardiovaskular berdasarkan data medis?
2. Bagaimana sistem dapat memprediksi risiko penyakit kardiovaskular serta memberikan saran pencegahan berbasis hasil prediksi tersebut secara akurat dan user-friendly?

ID Tim: CC25-CF225

Anggota Tim:

1. (ML) MC008D5X2348 - Marsha Rasyida Al-Farabi - Universitas Gadjah Mada - Aktif
2. (ML) MC008D5X2087 - Syifa Azzahra Susilo - Universitas Gadjah Mada - Aktif
3. (ML) MC008D5X2388 - Indara Nurwulandari - Universitas Gadjah Mada - Aktif
4. (FEBE) FC492D5Y2169 - Muhammad Novian - Universitas Muhammadiyah Yogyakarta - Aktif
5. (FEBE) FC179D5Y0593- Muh Diaz Nazarudin Rahman - Universitas Ahmad Dahlan - Aktif
6. (FEBE) FC179D5Y1055 - Aditya Navra Erlangga - Universitas Ahmad Dahlan - Aktif

Tema yang Dipilih:

Health Innovation ▾

Judul Proyek:

III Detect – Cardiovascular Diagnosis Detector: "Detect Your Risk. Protect Your Heart"

Ringkasan Eksekutif/Abstrak:

Penyakit kardiovaskular menjadi penyebab utama kematian di dunia, termasuk Indonesia, dengan banyak kasus yang terlambat terdeteksi akibat keterbatasan akses skrining dini dan rendahnya kesadaran masyarakat terhadap faktor risiko. Proyek ini bertujuan mengembangkan sebuah website inovatif yang memungkinkan pengguna melakukan skrining mandiri risiko penyakit kardiovaskular secara praktis dan mudah diakses. Dengan memanfaatkan data medis terpercaya dan algoritma penilaian risiko yang mengintegrasikan gejala, riwayat kesehatan, serta gaya hidup, platform ini dirancang menggunakan pendekatan design thinking untuk memastikan solusi yang benar-benar dibutuhkan masyarakat. Selain memberikan prediksi risiko secara cepat, website ini juga menyajikan edukasi yang komprehensif guna meningkatkan kesadaran dan mendorong tindakan pencegahan sejak dini. Tim kami terdorong untuk menyelesaikan masalah ini karena tingginya angka kematian akibat penyakit jantung dan minimnya solusi skrining yang efektif dan terjangkau. Diharapkan, platform ini dapat menjadi alat penting dalam mengurangi beban penyakit kardiovaskular melalui deteksi dini dan edukasi yang tepat sasaran.

Bagaimana tim Anda menemukan proyek ini?

Proyek ini bertujuan untuk mengidentifikasi dan mengembangkan solusi inovatif berdasarkan diskusi kelompok yang dilakukan melalui platform Zoom. Setiap anggota tim menyampaikan riwayat proyek yang pernah mereka kerjakan, yang kemudian dikompilasi menjadi daftar ide potensial. Setelah itu, dilakukan voting untuk memilih proyek yang paling relevan dan dibutuhkan dalam konteks saat ini. Proses ini bertujuan untuk memastikan bahwa solusi yang dihasilkan tidak hanya berfokus pada fitur tambahan, tetapi juga pada masalah nyata yang dihadapi oleh pengguna atau masyarakat. Pertanyaan penelitian yang diangkat adalah, "Bagaimana cara mengidentifikasi dan mengembangkan solusi yang paling relevan dan berdampak dari berbagai riwayat proyek yang ada?" Latar belakang dari proyek ini berfokus pada pentingnya kolaborasi dalam pengembangan solusi yang berdasarkan pada pengalaman praktis dan kebutuhan nyata, serta menggunakan

teknik seperti design thinking untuk meminimalkan risiko kesalahan dalam pemilihan ide. Tim kami ingin menyelesaikan masalah ini untuk menciptakan solusi yang dapat memberikan nilai tambah yang nyata dan efektif, bukan hanya menambah fitur tanpa mempertimbangkan kebutuhan utama pengguna.

Ruang Lingkup & Hasil Proyek:

Proyek ini bertujuan mengembangkan aplikasi web fungsional yang memungkinkan pengguna memasukkan data terkait gejala, riwayat kesehatan, dan gaya hidup untuk mendapatkan prediksi risiko penyakit kardiovaskular, disertai rekomendasi umum berbasis edukasi. Berikut adalah rincian tugas dan pembagian tanggung jawab tim:

Pembagian Tugas:

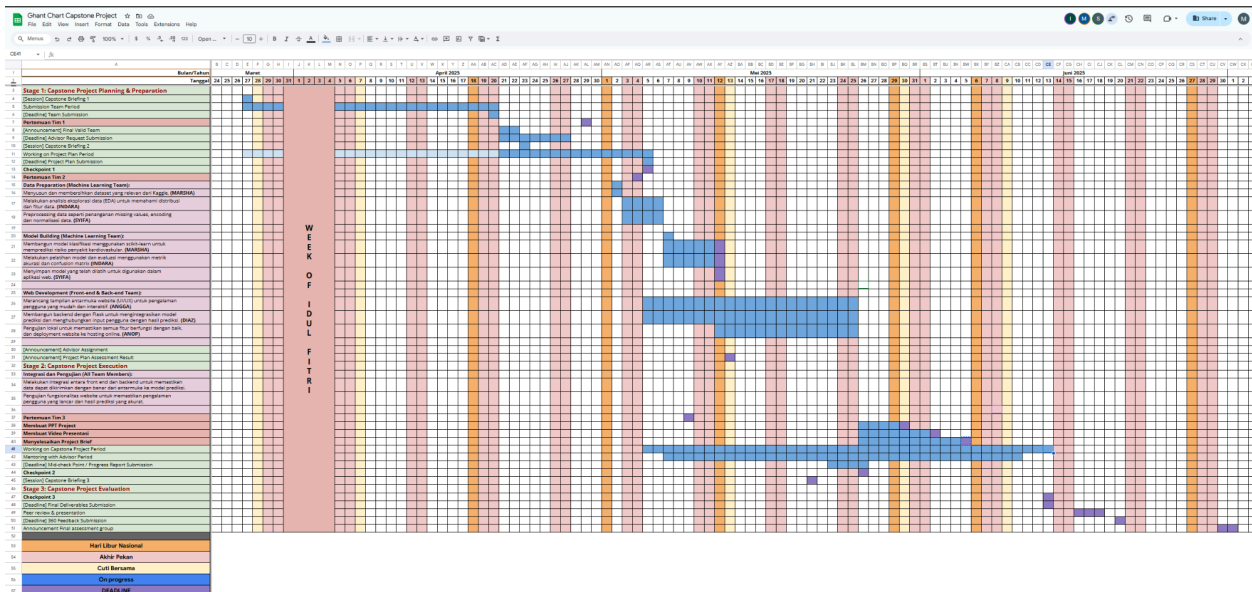
1. Data Preparation (Machine Learning Team):
 - Menyusun dan membersihkan dataset yang relevan dari Kaggle.
 - Melakukan analisis eksplorasi data (EDA) untuk memahami distribusi dan fitur data.
 - Preprocessing data seperti penanganan missing values, encoding, dan normalisasi data.
2. Model Building (Machine Learning Team):
 - Membangun model klasifikasi menggunakan scikit-learn untuk memprediksi risiko penyakit kardiovaskular.
 - Melakukan pelatihan model dan evaluasi menggunakan metrik akurasi dan *confusion matrix*
 - Menyimpan model yang telah dilatih untuk digunakan dalam aplikasi web (menggunakan joblib atau pickle).
3. Web Development (Front-end & Back-end Team):
 - Merancang tampilan antarmuka website (UI/UX) untuk pengalaman pengguna yang mudah dan interaktif.
 - Membangun backend dengan Flask untuk mengintegrasikan model prediksi dan menghubungkan input pengguna dengan hasil prediksi.
 - Pengujian lokal untuk memastikan semua fitur berfungsi dengan baik, dan deployment website ke hosting online.
4. Integrasi dan Pengujian (All Team Members):
 - Melakukan integrasi antara front end dan backend untuk memastikan data dapat dikirimkan dengan benar dari antarmuka ke model prediksi.

- Pengujian fungsionalitas website untuk memastikan pengalaman pengguna yang lancar dan hasil prediksi yang akurat.

Hasil yang Diharapkan:

- Sebuah website yang dapat melakukan skrining risiko penyakit kardiovaskular secara mandiri bagi pengguna.
- Sebuah website yang memberikan prediksi berdasarkan data yang dimasukkan, serta menyarankan langkah-langkah pencegahan untuk menjaga kesehatan jantung.
- Model prediksi yang dievaluasi menggunakan metrik akurasi yang memadai, serta website dengan antarmuka yang mudah digunakan oleh pengguna.

Jadwal Proyek:



Gantt Chart Capstone Project

Berdasarkan pengetahuan tim Anda, alat/IDE/Library, dan sumber daya apa yang akan digunakan untuk menyelesaikan masalah?

Untuk menyelesaikan masalah prediksi penyakit kardiovaskular, tim kami akan memanfaatkan berbagai alat, IDE, dan *library* yang mendukung dari tahap awal hingga akhir. Pada tahap eksplorasi data, *preprocessing*, dan pelatihan model, kami menggunakan *Jupyter Project Notebook* pada aplikasi *Visual Studio Code* (VSCode). *Library* utama yang akan digunakan adalah *scikit-learn* untuk membangun model *machine learning* dan

preprocessing, pandas untuk manipulasi data, serta numpy untuk operasi numerik. Visualisasi akan menggunakan *library* matplotlib dan seaborn, sedangkan penyimpanan model dilakukan dengan Tensorflow.js. Pengolahan data berbasis gambar atau *deep learning*, kami juga mempertimbangkan menggunakan *library* Keras, cv2, dan TensorFlow.

Untuk pengembangan front end, tim kami akan membangun antarmuka pengguna yang responsif dan interaktif dengan menggunakan framework modern serta module bundler seperti webpack atau Vite. Front end akan melakukan panggilan jaringan (API calls) untuk berinteraksi dengan backend secara efisien. Pada sisi back-end, kami akan membangun API RESTful yang mendukung aplikasi front end, dengan kemampuan menyimpan data menggunakan database atau tanpa database sesuai kebutuhan. Integrasi fitur AI/ML sebagai bagian utama aplikasi juga akan diimplementasikan di back-end. Seluruh pengembangan akan dilakukan tanpa menggunakan generator web otomatis agar kontrol penuh atas kode tetap terjaga dan aplikasi berjalan stabil tanpa crash.

Dari sisi platform, tim kami akan menggunakan GitHub untuk menyimpan dan mengelola kode sumber secara kolaboratif, serta Google Collab sebagai alternatif IDE berbasis cloud yang memungkinkan penggunaan GPU gratis untuk mempercepat proses pelatihan model. Untuk pengembangan front end, tim akan menggunakan framework modern yang mendukung pembuatan antarmuka responsif dan interaktif, serta memastikan integrasi yang mulus dengan back-end melalui API. Pada sisi back-end, kami akan membangun API RESTful yang aman dan efisien, serta mengelola penyimpanan data dengan sistem database yang sesuai kebutuhan. Seluruh proses pengembangan front end dan back-end dilakukan dengan kontrol penuh tanpa menggunakan generator otomatis agar aplikasi tetap stabil dan mudah dikembangkan. Sumber daya yang kami gunakan mencakup dataset rekam medis yang relevan dengan penyakit kardiovaskular dari Kaggle, serta referensi dari dokumentasi resmi seperti artikel penelitian dan jurnal medis untuk memastikan pendekatan yang kami gunakan sesuai dengan praktik medis yang dapat dipertanggungjawabkan secara ilmiah.

Berdasarkan pengetahuan dan eksplorasi, apa yang tim Anda butuhkan untuk dukungan?

Pertama, dari sisi data, kami membutuhkan dataset yang bersih, relevan, dan cukup besar yang mewakili berbagai faktor risiko kardiovaskular dengan jumlah missing value yang minimal. Selain itu, panduan atau best practice dalam preprocessing data medis sangat

dibutuhkan, mengingat data medis seringkali memiliki format yang kompleks seperti data kategorikal, missing values, hingga outlier. Dukungan dari mentor machine learning yang berpengalaman di bidang kesehatan juga sangat penting agar validasi fitur dan model yang dikembangkan tetap sesuai dengan kaidah medis dan diagnosa klinis yang logis.

Dari sisi pengembangan front end, kami memerlukan panduan dan praktik terbaik dalam membangun antarmuka pengguna yang responsif, mudah digunakan, dan dapat berinteraksi lancar dengan backend. Sedangkan untuk back-end, kami membutuhkan dukungan dalam merancang API yang aman, efisien, serta integrasi yang tepat antara model AI/ML dengan sistem backend agar aplikasi berjalan stabil dan dapat diandalkan. Mentor yang berpengalaman dalam pengembangan full-stack dan pengelolaan data medis juga akan sangat membantu untuk memastikan aplikasi memenuhi standar keamanan dan performa yang dibutuhkan.

Berdasarkan pengetahuan dan eksplorasi, informasikan kepada kami bagian Machine Learning Anda!

Membangun model klasifikasi untuk memprediksi apakah seseorang mengidap penyakit kardiovaskular. Proses *machine learning* mencakup tahap pre-processing, analisis data eksploratif (EDA), pembagian data menjadi data latih, validasi, dan uji, pelatihan model, evaluasi menggunakan metrik akurasi dan *confusion matrix*, serta penyimpanan model untuk inferensi.

Berdasarkan pengetahuan dan eksplorasi, informasikan kepada kami bagian Web/Front end/Backend Anda!

Membangun website interaktif dimana website dapat memprediksi penyakit kardiovaskular, proses pengerjaan dimulai dari perancangan tampilan antarmuka, dilanjutkan dengan pembuatan backend menggunakan Node.js untuk membuat model prediksi, integrasi form input dengan backend, lalu dilakukan pengujian lokal dan opsional deployment ke hosting online.

Berdasarkan perencanaan tim, apakah ada potensi risiko atau masalah yang dapat menghambat proyek?

Dari sisi Front end Back-End, mulai dari proses integrasi data dari Front End ke Back-End. model prediksi yang tidak akurat karena data pelatihan terbatas, kendala dalam deployment jika tidak familiar dengan platform hosting. Selain itu, validasi input yang

kurang ketat dapat menyebabkan error saat prediksi, dan kurangnya dokumentasi atau pembagian tugas yang jelas bisa memperlambat pengembangan. Risiko lain keterbatasan waktu pengerjaan dan mungkin error saat memuat model .pkl jika struktur input tidak konsisten.

Dari sisi machine learning, potensi risiko meliputi kualitas data yang rendah, overfitting model, serta ketidaksesuaian proses preprocessing antara pelatihan dan inferensi. Untuk mengidentifikasi dan mengatasi hal ini, tim kami akan melakukan EDA, penerapan regularisasi, dan standarisasi pipeline. Jika terkendala resource, akan digunakan teknik model pruning untuk efisiensi.

Catatan/Tambahan lain yang perlu dipertimbangkan dalam aplikasi tim Anda

- Penting untuk menekankan kepada pengguna bahwa aplikasi ini adalah alat skrining atau edukasi dan bukan pengganti diagnosis medis profesional. Hasil prediksi harus selalu dikonsultasikan dengan dokter.
- Pertimbangkan aspek etika dan privasi data pengguna secara serius sejak awal perancangan.