

BAB I PENDAHULUAN

1.1 LATAR BELAKANG

Penyakit kardiovaskular terus menjadi penyebab kematian nomor satu di dunia (19 juta kematian pada tahun 2022, yang mencakup hampir sepertiga dari seluruh kematian di dunia). Sebagian besar kematian ini disebabkan oleh serangan jantung (infark miokard), yang merupakan bentuk penyakit kardiovaskular akut yang sangat serius. Mungkin yang paling mengkhawatirkan adalah, tiga perempat dari jumlah kematian ini terjadi di negara-negara berpenghasilan rendah dan menengah, dengan sebagian besar adalah anak-anak atau individu di bawah usia 70 tahun. Pola ini mengungkapkan titik buta yang penting: model prediksi risiko saat ini tidak memadai dalam mengidentifikasi orang-orang yang berisiko tinggi terkena serangan jantung, terutama di antara populasi yang lebih beragam.[1].

Secara klinis, hipertensi, kolesterol LDL tinggi, diabetes, obesitas, dan merokok merupakan faktor risiko serangan jantung yang sudah dikenal. Pedoman klinis terbaru, termasuk yang dikeluarkan oleh ACC/AHA, menekankan bahwa modifikasi gaya hidup adalah dasar dari pencegahan primer.[2] Selain itu, sifat hubungan ini semakin didukung oleh bukti yang menjelaskan peran kausal kolesterol LDL dalam aterosklerosis dan menemukan bahwa pengurangan LDL yang berkepanjangan menurunkan kejadian kardiovaskular seperti serangan jantung.[3].

Terlepas dari penanda klinis ini, perilaku gaya hidup juga secara langsung memengaruhi risiko serangan jantung. Apa yang sudah diketahui tentang hal ini: Studi berbasis akselerometer skala besar menunjukkan bahwa bahkan aktivitas fisik tingkat rendah secara substansial dan bergantung pada dosis menurunkan risiko kematian akibat semua penyebab, sementara durasi perilaku sedentari yang lebih lama pada orang dewasa meningkatkan risiko penyakit kardiovaskular.[4] Secara spesifik, waktu duduk yang berkepanjangan ($\geq 10,6$ jam per hari) dikaitkan dengan risiko gagal jantung dan kematian kardiovaskular yang lebih tinggi.[5] Demikian pula, kebiasaan tidur yang tidak teratur, atau variabilitas dalam durasi dan waktu tidur, telah dikaitkan dengan peningkatan risiko kardiovaskular secara independen dari total durasi tidur.[6] Komponen Kuantitatif dan Aplikasi yang Efektif Konsumsi alkohol menunjukkan manfaat pada risiko kardiovaskular, dalam pola respons dosis dari efek setiap hari dan hampir mencapai batas kristal melalui hasil kekuatan koktail, pengurangan konsumsi merupakan tanda-tanda terukur untuk peningkatan kebugaran kardiovaskular.[7] Berdasarkan temuan ini, penyertaan variabel yang berkaitan dengan gaya hidup sangat penting untuk meningkatkan kemampuan prediksi risiko serangan jantung.

Selain itu, distribusi dan tingkat keparahan risiko serangan jantung sangat dipengaruhi oleh determinan sosial ekonomi kesehatan (SDoH). Perilaku kesehatan dan hasil klinis dipengaruhi oleh faktor-faktor termasuk tingkat pendidikan, akses ke perawatan kesehatan, kondisi pekerjaan, dan paparan terhadap lingkungan. Menurut American Heart Association, SDoH merupakan faktor mediasi baik untuk faktor risiko klasik maupun ketidaksetaraan dalam hasil kardiovaskular.[8]Tingkat pendidikan yang lebih rendah juga dikaitkan dengan insiden dan mortalitas yang lebih tinggi akibat kejadian kardiovaskular, seperti serangan jantung di negara-negara dengan sumber daya terbatas, sebagaimana dibuktikan oleh studi epidemiologi prospektif perkotaan-pedesaan (PURE).[9]Oleh karena itu, konteks sosial ekonomi sangat penting untuk meningkatkan akurasi dan keadilan model prediktif.

Namun, model prediksi risiko standar seringkali tidak memperhitungkan konteks yang lebih luas ini.[10], [11]Lebih dari sekadar pembacaan fisiologis, risiko terkena serangan jantung bergantung pada bagaimana individu menjalani kehidupan sehari-hari mereka: seberapa aktif mereka, berapa lama mereka tidak aktif, apakah mereka tidur secara teratur setiap malam, dan apakah mereka terlibat dalam perilaku berbahaya seperti merokok atau minum alkohol berlebihan. Metode statistik tradisional kesulitan memodelkan hubungan multidimensi semacam itu.

Pembelajaran mesin (ML) merupakan alternatif yang menjanjikan yang memungkinkan pemodelan interaksi kompleks dan non-linier antara faktor klinis, gaya hidup, dan sosioekonomi. Di antara metode-metode tersebut, Extreme Gradient Boosting (XGBoost) telah menunjukkan kinerja yang baik dalam menangani data medis terstruktur.[12]Namun demikian, prediksi dengan akurasi tinggi saja tidak cukup. Model juga harus dapat diinterpretasikan dalam praktik klinis untuk membangun kepercayaan dan mendorong penerapannya. Prediksi yang dibuat oleh model kotak hitam berdasarkan algoritma yang tidak transparan — berapa pun akurasinya — kemungkinan besar tidak akan diterima di dunia perawatan kesehatan nyata.

Dengan memanfaatkan metode kecerdasan buatan yang dapat dijelaskan (XAI) seperti SHapley Additive exPlanations (SHAP)[13]Kita tidak hanya dapat mengurangi keterbatasan ini tetapi juga menawarkan kerangka kerja yang solid untuk memahami prediksi model. SHAP memungkinkan penjelasan global dan lokal, menentukan faktor mana yang paling berkontribusi terhadap peningkatan risiko serangan jantung. Yang terpenting, SHAP menekankan parameter yang dapat dimodifikasi seperti merokok dan olahraga, mendorong kesadaran risiko dan manajemen risiko bagi dokter maupun pasien.

Penelitian ini mengusulkan model prediksi risiko serangan jantung yang dapat dijelaskan berbasis XGBoost yang dikombinasikan dengan SHAP. Tujuan: Kami mengusulkan kerangka

kerja pembelajaran mesin yang dapat diinterpretasikan untuk memprediksi hasil AD berdasarkan pengukuran klinis, perilaku gaya hidup, dan faktor sosioekonomi guna memberikan prediksi yang akurat dan dapat diinterpretasikan. Kami mengevaluasi model dengan berbagai metrik kinerja yang dirancang untuk memfasilitasi deteksi dini, meningkatkan komunikasi risiko, dan memungkinkan penargetan intervensi. Dalam semua pekerjaan ini, pada akhirnya kami mencoba untuk memberikan sesuatu yang bermanfaat dan transparan bagi dokter, pasien, dan pembuat kebijakan dalam hal bagaimana kita dapat membuat serangan jantung lebih bebas masalah.

1.2 Rumusan Masalah

1. Bagaimana keterbatasan model prediksi risiko serangan jantung konvensional dalam mengakomodasi faktor klinis, gaya hidup, dan determinan sosial ekonomi kesehatan?
2. Bagaimana pengaruh integrasi variabel gaya hidup dan faktor sosioekonomi terhadap peningkatan akurasi prediksi risiko serangan jantung?
3. Bagaimana kinerja metode *Extreme Gradient Boosting* (XGBoost) dalam memprediksi risiko serangan jantung dibandingkan pendekatan statistik tradisional?
4. Bagaimana penerapan metode *Explainable Artificial Intelligence* (XAI) menggunakan SHAP dalam meningkatkan interpretabilitas model prediksi risiko serangan jantung?
5. Faktor-faktor apa saja yang paling berkontribusi terhadap risiko serangan jantung berdasarkan hasil interpretasi model?

1.3 Tujuan Penelitian

1. Menganalisis keterbatasan model prediksi risiko serangan jantung konvensional.
2. Mengembangkan model prediksi risiko serangan jantung berbasis XGBoost dengan mengintegrasikan faktor klinis, gaya hidup, dan determinan sosial ekonomi kesehatan.
3. Mengevaluasi kinerja model yang dikembangkan menggunakan metrik evaluasi yang relevan untuk deteksi dini.
4. Mengimplementasikan metode SHAP untuk meningkatkan interpretabilitas model prediksi.
5. Mengidentifikasi dan menganalisis faktor-faktor dominan yang memengaruhi risiko serangan jantung secara global maupun individual.

1.4 Batasan Masalah

1. Penelitian difokuskan pada prediksi risiko serangan jantung (infark miokard) sebagai salah satu bentuk penyakit kardiovaskular.
2. Evaluasi model dilakukan berdasarkan metrik performa statistik tanpa uji implementasi klinis secara langsung.
3. Variabel yang digunakan meliputi faktor klinis (misalnya tekanan darah, kadar kolesterol LDL, diabetes), gaya hidup (aktivitas fisik, perilaku sedentari, pola tidur, merokok, konsumsi alkohol), dan determinan sosial ekonomi (pendidikan, akses kesehatan, kondisi pekerjaan)

1.5 Manfaat Penelitian

1. Manfaat Akademis : Memberikan kontribusi dalam pengembangan ilmu di bidang data sains kesehatan, khususnya dalam penerapan pembelajaran mesin untuk prediksi penyakit kardiovaskular
2. Manfaat Tenaga Medis : Membantu dalam pengambilan keputusan klinis melalui model prediksi yang akurat dan dapat diinterpretasikan.