

BAB I PENDAHULUAN

1.1 Latar Belakang

Pada saat ini banyak sekali penyakit kronis tidak menular yang berakibat buruk kepada kita dan salah satunya adalah serangan jantung. Menurut (WHO, 2020) bahwa penyakit jantung merupakan salah satu dari 10 penyakit yang menyebabkan kematian terbanyak di dunia sebesar 16%. Menurut (Kemenkes,2019) sekitar 1 juta orang menderita penyakit jantung yang dapat dilihat pada Gambar 1.1. Penyakit jantung mengakibatkan kematian disebabkan oleh tekanan darah, stress, kerja berlebihan, gula darah dan banyak penyebab lainnya. Jantung adalah organ utama yang harus bekerja dengan benar karena berfungsi untuk memompa darah ke seluruh tubuh sehingga oksigen dan zat-zat gizi dapat tersalurkan. Jika jantung tidak bekerja dengan benar, maka otak akan menderita dan kedua organ tersebut akan berhenti bekerja dan menyebabkan kematian pada seseorang (Gerstenblith, 2007).

Provinsi	Penyakit jantung		N tertimbang
	%	95% CI	
Aceh	1,6	1,5 - 1,8	20.244
Sumatera Utara	1,3	1,2 - 1,5	55.351
Sumatera Barat	1,6	1,4 - 1,8	20.663
Riau	1,1	0,9 - 1,2	26.085
Jambi	0,9	0,7 - 1,1	13.692
Sumatera Selatan	1,2	1,0 - 1,4	32.126
Bengkulu	1,3	1,1 - 1,5	7.531
Lampung	1,2	1,1 - 1,4	32.148
Bangka Belitung	1,5	1,2 - 1,8	5.592
Kepulauan Riau	1,5	1,2 - 1,9	8.173
DKI Jakarta	1,9	1,7 - 2,2	40.210
Jawa Barat	1,6	1,5 - 1,7	186.809
Jawa Tengah	1,6	1,5 - 1,7	132.565
DI Yogyakarta	2,0	1,7 - 2,3	14.602
Jawa Timur	1,5	1,4 - 1,6	151.878
Banten	1,4	1,2 - 1,5	48.621
Bali	1,3	1,2 - 1,5	16.481
Nusa Tenggara Barat	0,9	0,7 - 1,1	19.247
Nusa Tenggara Timur	0,7	0,6 - 0,8	20.599
Kalimantan Barat	1,3	1,2 - 1,5	19.190
Kalimantan Tengah	1,3	1,1 - 1,5	10.189
Kalimantan Selatan	1,3	1,1 - 1,5	16.043
Kalimantan Timur	1,9	1,6 - 2,2	13.977
Kalimantan Utara	2,2	1,8 - 2,7	2.733
Sulawesi Utara	1,8	1,5 - 2,0	9.542
Sulawesi Tengah	1,9	1,7 - 2,2	11.548
Sulawesi Selatan	1,5	1,3 - 1,6	33.693
Sulawesi Tenggara	1,4	1,2 - 1,6	10.167
Gorontalo	2,0	1,6 - 2,4	4.547
Sulawesi Barat	1,5	1,2 - 2,0	5.195
Maluku	1,5	1,2 - 1,7	6.801
Maluku Utara	1,1	0,9 - 1,4	4.723
Papua Barat	1,2	0,9 - 1,5	3.588
Papua	0,9	0,7 - 1,0	12.736
INDONESIA	1,5	1,4 - 1,5	1.017.290

Gambar 1.1. Data penderita penyakit jantung di Indonesia tahun 2018

Sumber: Kemenkes, 2019

Sistem informasi klinis berisi dua data utama, yaitu data pasien lengkap dan data pemeriksaan kesehatan. Sistem ini juga dikenal sebagai *Electronic Medical Record* (EMR). EMR adalah sistem di mana data pasien dapat diproses secara elektronik (Al-Rayes, S. A., & Al-Meshkhas, M., 2020; Zhu, H., & Hou, M., 2018). Keuntungan dari EMR adalah dapat meningkatkan kualitas layanan dengan database terintegrasi, mengurangi penggunaan kertas dan pengerjaan menjadi lebih cepat dan mudah (Amalia et al., 2021; Wu et al., 2017; Vo et al, 2018; Halim et al, 2020, Joseph, B., et al, 2020).

Dalam proses merancang sistem informasi, software yang digunakan adalah sistem operasi Windows 8.1, Visual Studio Code, Python 3, MySQL Database Workbench 8.0, dan Orange Data Mining. Database adalah sistem yang digunakan untuk membuat, mengakses, mengatur, mencari, menghapus, menyalin, memperbarui, dan menyimpan data. Visual Studio Code adalah aplikasi berbasis open source yang mendukung banyak bahasa dan ekstensi. Keuntungan dari Visual Studio Code adalah memiliki banyak fitur seperti mampu melacak kesalahan pengkodean dengan cepat, dan memiliki antarmuka yang mudah digunakan. MySQL digunakan untuk mengelola database dengan jumlah data yang sangat besar. MySQL dapat diakses oleh banyak pengguna dan dapat memprosesnya secara bersamaan (Day N. J. B. et al., 2018; Tongkaw. S., 2016).

Python adalah bahasa pemrograman yang sangat populer dan banyak digunakan oleh programmer di dunia. Keuntungan python sederhana dan mudah dipelajari, komunitas yang mendukung, digunakan dalam *Big Data* dan *Machine Learning*, efisien, open-source, dapat digunakan pada semua OS (Windows, macOS, Linux), dan memiliki banyak kepastakaan di dalamnya (Sayeth et al., 2019; Stancin, I., & Jovic, A., 2019; Mehata, S. et al., 2019). Bahasa Python dalam penelitian ini digunakan untuk membuat GUI desktop dengan bantuan Python, tkinter, dan kepastakaan lainnya.

Machine Learning adalah salah satu cabang dari Artificial Intelligence yang digunakan untuk prediksi dan klasifikasi. ML merupakan studi yang memberi kemampuan kepada komputer untuk mempelajari tanpa harus diprogram berulang. Hasil yang diperoleh yaitu berasal dari data yang dipelajari sebelumnya oleh

komputer secara berulang dan digunakan oleh pengguna untuk mengambil sebuah keputusan dalam memecahkan sebuah masalah (Guruvayur, S. R., & Suchithra, R., 2017; Singh, P., et al, 2019).

Ber macam macam algoritma *Machine Learning* digunakan untuk memprediksi penyakit jantung baik itu supervised *Machine Learning* atau unsupervised *Machine Learning* (Guruvayur, S. R., & Suchithra, R., 2017). Namun akurasi untuk tiap algoritma memiliki nilai yang berbeda pada berbagai penelitian yang telah dilakukan oleh (Obasi, T., & Omair Shafiq, M., 2019; Keya et al, 2021; Kavitha, M., et al, 2021; Alex P., et al, 2019; Tougui, I., 2020; Kannan, R., & Vasanthi, V., 2018; Ed-Daoudy, A., & Maalmi, K., 2019) seperti terlihat pada Tabel 1.1.

Tabel 1.1.

Perbandingan Algoritma Machine Learning

Judul	Algoritma ML	Akurasi
Towards comparing and using <i>Machine Learning</i> techniques for detecting and predicting Heart Attack and Diseases	<i>Logistic Regression</i>	59.7 %
	Naive Bayes	61.9 %
	<i>Random Forest</i>	92.4 %
Measuring the Heart Attack Possibility using Different Types of <i>Machine Learning</i> Algorithms	<i>Logistic Regression</i>	80 %
	<i>Random Forest</i>	75 %
	Bagging	74 %
	MLP	78 %
	Decision Tree	70 %
Heart Disease Prediction using Hybrid <i>Machine Learning</i> Model	Decision Tree	79 %
	<i>Random Forest</i>	81 %
	DT + RF	88 %
Performance evaluation of machine learning based big data processing framework for prediction of heart disease	SVM	85.82 %
	Decision Tree	82.80 %
	<i>Random Forest</i>	87.50 %
	<i>Logistic Regression</i>	85.70 %

Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique	SVM	85.88 %
	<i>Random Forest</i>	85.88 %
	KNN	83.21 %
	ANN	92.21 %
Heart disease classification using data mining tools and <i>Machine Learning</i> techniques	KNN	67 %
	ANN	79.46 %
	Naive Bayes	83.16 %
	<i>Random Forest</i>	81.82 %
<i>Machine Learning</i> Algorithms with ROC Curve for Predicting and Diagnosing the Heart Disease	<i>Logistic Regression</i>	86.5 %
	<i>Random Forest</i>	80.8 %
	Gradient Boosting	84.2 %
	SVM	79.7 %

Algoritma *Machine Learning* yang ada tersebut digunakan untuk meningkatkan kinerja dan akurasi dalam memprediksi penyakit serangan jantung. Namun, peningkatan masih perlu dilakukan untuk mendapatkan akurasi terbaik. Sehingga penelitian ini menyajikan sistem yang menggabungkan sistem informasi rekam medis dan *Machine Learning* dengan menggunakan algoritma *Logistic Regression*, *Random Forest*, dan algoritma hybrid.

1.2 Rumusan Masalah

Setelah mengetahui latar belakang dari tugas akhir ini, adapun rumusan masalah dalam penyusunan tugas akhir sebagai berikut :

1. Bagaimana cara membangun sistem informasi rekam medis berbasis aplikasi desktop ?
2. Bagaimana cara merancang program *Machine Learning* untuk memprediksi penyakit serangan jantung ?
3. Bagaimana perbandingan antara prediksi penyakit serangan jantung dengan menggunakan algoritma *Logistic Regression*, *Random Forest*, dan algoritma hybrid ?.

1.3 Tujuan Penelitian

Tujuan dalam penelitian tugas akhir ini sebagai berikut :

1. membangun sistem informasi rekam medis berbasis aplikasi *desktop*;
2. merancang program *Machine Learning* untuk memprediksi penyakit serangan jantung;
3. melakukan perbandingan prediksi penyakit serangan jantung dengan algoritma *Machine Learning* menggunakan *Logistic Regression*, *Random Forest* dan algoritma hybrid.

1.4 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Bagi Peneliti

Peneliti diharapkan mendapatkan pengetahuan baru mengenai pembuatan sistem informasi dan prediksi penyakit serangan jantung. Selain itu, peneliti mengetahui proses penerapan metode *Logistic Regression*, *Random Forest*, dan algoritma hybrid untuk pengolahan data dalam memprediksi penyakit serangan jantung.

2. Bagi Pihak Lain

Hasil penelitian ini membantu praktisi dan peneliti sebagai referensi mengenai prediksi penyakit serangan jantung berbasis algoritma *Machine Learning* dan sebagai solusi untuk prediksi penyakit serangan jantung yang lebih akurat sehingga dapat mengobati penderita penyakit jantung secepat mungkin di masa mendatang.

1.5 Batasan Masalah

Batasan masalah dari penelitian ini di antaranya sebagai berikut.

1. dataset yang diolah terbagi menjadi 13 kategori yaitu *sex*, *cp*, *trestbps*, *chol*, *FBS*, *resting*, *thali*, *exang*, *oldpeak*, *slope*, *ca*, *thal* dan *output*;
2. algoritma *Machine Learning* yang dipakai hanya *Logistic Regression* dan *Random Forest*;
3. bahasa pemrograman untuk sistem informasi yang dibangun menggunakan bahasa pemrograman *Python* dan *MySQL* sebagai database;

4. tidak membahas mengenai biaya yang harus dikeluarkan untuk mengimplementasikan sistem informasi dan pengolahan data;
5. sistem informasi rekam medis dibangun tanpa melakukan pengetesan lapangan di instansi terkait seperti klinik, puskesmas dan rumah sakit.

1.6 Sistematika Penulisan

Sistematika penulisan tugas akhir ini terdiri atas lima bab, dan disusun secara per bab. Pembagian bab tersebut adalah sebagai berikut.

1. BAB I : Pendahuluan

Pada bab pendahuluan ini, dikemukakan latar belakang penelitian, rumusan masalah penelitian, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan penelitian.

2. BAB II : Kajian Pustaka

Pada bab ini, akan dijelaskan landasan-landasan teori yang digunakan dalam dilakukannya penelitian ini. Di antaranya adalah mengenai teori sistem informasi, teori penyakit serangan jantung, teori mengenai *Machine Learning*, dan teori bahasa pemrograman python.

3. BAB III : Metode Penelitian

Pada bab ini, akan dikemukakan metode penelitian, tahapan penelitian, dan desain sistem yang digunakan dalam merancang dan membuat sistem informasi rekam medis dan *machine learning* untuk memprediksi penyakit serangan jantung.

4. BAB IV : Desain dan Analisis Penelitian

Pada bab ini, akan dibahas hasil dari penelitian yang dilakukan. Di antaranya adalah mengenai hasil pembuatan sistem informasi rekam medis berbasis desktop, hasil perancangan program *Machine Learning*, perbandingan hasil algoritma yang digunakan, serta kinerja dari algoritma prediksi yang digunakan.

5. BAB V : Kesimpulan dan Saran

Pada bab ini, akan dikemukakan kesimpulan dari hasil proses penelitian yang didapat serta saran untuk penelitian sejenis untuk kedepannya.