

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Dalam beberapa tahun terakhir, penelitian di bidang deteksi aritmia berbasis pembelajaran mendalam mengalami perkembangan pesat seiring dengan meningkatnya ketersediaan dataset medis berskala besar seperti PTB-XL [16]. Model konvensional berbasis *CNN* dan *LSTM* telah banyak digunakan untuk mengenali pola pada sinyal elektrokardiogram (ECG), namun kinerjanya sering kali terhambat oleh kompleksitas temporal dan variasi morfologi antarindividu serta hanya memanfaatkan kedua label yaitu *Abnormal* dan *Normal*. Oleh karena itu, penelitian terbaru mulai berfokus pada arsitektur hibrid dan berlapis yang tidak hanya memproses fitur spasial tetapi juga memahami hubungan jangka panjang antar sinyal. Dalam penelitian ini, pendekatan *ResNet Bidirectional LSTM Attention* digunakan untuk fokus representasi fitur, meningkatkan sensitivitas terhadap pola irama abnormal, serta menekankan bagian sinyal yang paling informatif untuk klasifikasi aritmia [20]. serta mengeksplorasi pengujian berbagai parameter arsitektur dan fitur, seperti jumlah *filter CNN*, ukuran *kernel*, jumlah unit *Bi-LSTM*, serta parameter pada mekanisme atensi untuk memperoleh konfigurasi model yang paling optimal. Evaluasi dilakukan secara sistematis menilai dampak dari setiap parameter terhadap performa model, berupa peningkatan *F1-score* dan sensitivitas pada kelas aritmia minoritas serta analisis perbandingan terhadap hasil eksperimen untuk mengidentifikasi kombinasi parameter yang memberikan keseimbangan terbaik Selain itu, integrasi dengan model *RoBERTa* pada laporan medis berbasis teks memperkuat konteks diagnostik, menggabungkan pemahaman fisiologis dari sinyal dengan deskripsi klinis dari laporan [21]. Pendekatan multimodal ECG dan teks medis ini diharapkan menghasilkan sistem deteksi aritmia yang lebih komprehensif, adaptif, dan akurat, sebagaimana disajikan pada Tabel 2.1 yang menjadi acuan dalam pemilihan arsitektur terbaik untuk penelitian ini.

Tabel 2.1 Penelitian terdahulu berdasarkan metode yang digunakan

No	Tahun	Masalah	Dataset	Model	Result (Akurasi/F1)	Main Findings
<i>Studi Literatur Metode ResNet–Bidirectional LSTM–Attention (Sinyal ECG)</i>						
1	2022	<i>Hybrid ResNet + BiLSTM untuk klasifikasi aritmia</i>	<i>Benchmark ECG (multi-dataset)</i>	<i>Res-BiANet (ResNet + BiLSTM + Attention)</i>	Acc 92.22%, F1 86.75% (lap. penulis)	<i>Res-BiANet menunjukkan peningkatan akurasi vs ablation (ResNet) [22].</i>
2	2023	<i>2D-ECG + ResNet dengan SE modules untuk multi-class arrhythmia</i>	<i>PTB-XL / 12-lead 2D transform</i>	<i>DSE-ResNet (ResNet + detached SE)</i>	(lap. paper: per-class mtr; overall akurasi tinggi di eksperimen)	representasi 2D antar-lead meningkatkan ekstraksi fitur [23].
3	2024	<i>ResNet + SE + BiLSTM untuk MIT-BIH arrhythmia</i>	<i>MIT-BIH</i>	<i>ResNet + SE + BiLSTM + Attention</i>	<i>Acc up to 99% pada beberapa kelas (lap. penulis)</i>	<i>SE block + BiLSTM + attention meningkatkan performa pada sinyal single-lead [24].</i>
4	2024	<i>CNN–BiLSTM–Attention untuk deteksi atrial fibrillation</i>	<i>Public ECG sets (chapman, CPSC, PTB-XL)</i>	<i>CNN–BiLSTM–Attention (CLA-AF)</i>	<i>Dilaporkan Acc / Sens / F1 (contoh: F1 tinggi untuk AF)</i>	<i>Pendekatan spatio-temporal + attention efektif untuk AF detection [25].</i>
5	2025	<i>TCN/ResNet + SE block untuk deteksi AF single-lead</i>	<i>Public single-lead ECG</i>	<i>TCN + SE (dengan ResNet elements)</i>	<i>Acc ≈ 98–99% pada pengujian terpilih</i>	<i>SE block konsisten memperbaiki akurasi model temporal [26].</i>
<i>RoBERTa(Laporan Medis Tekstual)</i>						
6	2022	<i>Clinical concept extraction (transfer/clinical RoBERTa)</i>	<i>i2b2, clinical corpora</i>	<i>RoBERTa–MIMIC / Clinical RoBERTa</i>	<i>F1 ≈ 0.80–0.90 (ber variasi per dataset)</i>	<i>RoBERTa yang dipretrain di data klinis unggul pada ekstraksi konsep klinis. [27]</i>

No	Tahun	Masalah	Dataset	Model	Result (Akurasi/F1)	Main Findings
7	2024	Ekstraksi temuan nodul paru dari laporan radiologi	<i>Radiology reports</i>	<i>RoBERTa-MIMIC fine-tuned</i>	$F1 \approx 0.9279$ (untuk nodule concepts)	RoBERTa-mimic mencapai F1 tinggi untuk konsep radiologi spesifik. [28]
8	2024	Deteksi temuan tak terduga di laporan radiologi	<i>Radiology reports</i>	<i>RoBERTafinetuned</i>	$Acc 86.15\%$ (lap. per penulis pada tugas test set)	Fine-tuning RoBERTa menunjukkan CNN/ChatGPT pada tugas ini [29]
9	2024	Adaptasi dan embedding LM untuk data kesehatan	EHR / berbagai corpora	RoBERTa/ Clinical fine-tuning approaches	(laporan: peningkatan micro-F1 $\approx 90\%$ pada tugas tertentu)	Penyesuaian pretraining pada data kesehatan meningkatkan kualitas embedding. [30]
10	2025	Fine-tuning LMs untuk menstrukturkan laporan radiologi	Large radiology corpora	RoBERTa variants / BERTrad	Token-level $F1 \approx 85.5\%$ (pada dataset radiology)	Model yang dipretrain pada laporan radiologi memberikan akurasi struktural yang lebih baik. [21]
<i>Hybrid (ResNet BiLSTM Attention & RoBERTa, Multimodal ECG & Teks)</i>						
11	2023	Multimodal fusion ECG + clinical notes	PTB-XL + clinical notes (research subset)	Multimodal fusion (ECG features + text embeddings)	Acc 92.8%, Macro-F1 84.6% (lap. penulis)	Menunjukkan teks klinis membantu prediksi kelas minor dan meningkatkan recall. [31]
12	2024	Multimodal representation learning (ECG + CXR + reports)	ECG, CXR, report subsets	Transformer-based multimodal (MoRE-like)	Peningkatan $F1 \approx 5\%-8\%$ vs unimodal (lap.)	Fusi multimodal meningkatkan performa pada tugas klinis kompleks. [16]

No	Tahun	Masalah	Dataset	Model	Result (Akurasi/F1)	Main Findings
13	2024	Fusi sinyal ECG dan metadata klinis untuk CVD prediction	ECG + EHR metadata	Hybrid ResNet feature + text LM fusion	Acc / AUC meningkat (lap. studi)	Metadata klinis + ECG memberi sinyal tambahan untuk deteksi risiko [32].
14	2025	Benchmark & prototipe multimodal ECG + teks (interpretable fusion)	Benchmarks baru (multimodal)	Lightweight attention fusion (ECG + text)	Hasil awal peningkatan robustnes; metrik bervariasi per benchmark	Munculnya benchmark multimodal awal menandai tren penelitian hybrid [33].
15	2025	Studi komparatif pipeline ECG (ResNet variants) dan fusi multimodal	PTB-XL, MIT-BIH, Chapman	ResNet variants + fusion strategies	Rekomendasi: ResNet-based pretrained models untuk stabilitas performa	Survey + empirical comparisons merekomendasikan ResNet pretrained untuk baseline [34]

Berdasarkan Tabel 2.1, berbagai penelitian terdahulu menunjukkan perkembangan signifikan dalam klasifikasi aritmia menggunakan pendekatan *deep learning* baik secara unimodal maupun multimodal. Pada kelompok studi dengan metode *ResNet–Bidirectional LSTM–Attention*, penelitian oleh [25], [26], [27], [28], dan [29] memperlihatkan kemajuan berarti dalam peningkatan akurasi klasifikasi sinyal *electrocardiogram (ECG)*. Model *Res-BiANet* yang diusulkan oleh [25] berhasil mencapai akurasi 92,22% dan F1-score 86,75%, menandakan peningkatan signifikan dibandingkan model *ablation ResNet* tunggal. Sementara itu, studi oleh [26] menggunakan pendekatan *DSE-ResNet* dengan transformasi sinyal menjadi citra 2D, yang terbukti meningkatkan kemampuan ekstraksi fitur spasial antar-lead. Penelitian lain oleh [27] menggabungkan *ResNet*, *Squeeze-and-Excitation (SE)*, *Bidirectional LSTM*, dan *Attention*, dan berhasil mencapai akurasi hingga 99% pada beberapa kelas dataset *MIT-BIH*, menunjukkan efektivitas blok SE dalam memperkuat fokus fitur penting pada sinyal ECG tunggal. Pendekatan *CNN–BiLSTM–Attention* pada studi [28] juga terbukti efektif untuk deteksi *Atrial*

Fibrillation (AF) pada berbagai dataset publik seperti *PTB-XL* dan *CPSC*, dengan nilai *F1* yang tinggi pada kelas *AF*. Di sisi lain, [29] menambahkan *Temporal Convolutional Network (TCN)* dengan elemen *ResNet* dan blok *SE* untuk memperbaiki akurasi deteksi hingga 98–99%, menegaskan bahwa kombinasi arsitektur spasial-temporal dan *attention* memberikan hasil optimal untuk analisis sinyal fisiologis.

Metode *RoBERTa* pada studi [30] [31] [32] [33] menyoroti keunggulan model *Transformer-based* tersebut dalam pemrosesan teks medis yang kompleks. Penelitian [30] menunjukkan bahwa *Clinical RoBERTa* yang dilatih pada korpus *i2b2* dan *MIMIC* mencapai *F1-score* hingga 0,90 pada tugas ekstraksi konsep klinis, mengungguli *BERT* dan model *NLP* konvensional. Studi [31] dan [32] memperluas penggunaan *RoBERTa* pada laporan radiologi, dengan hasil *F1* sebesar 0,9279 untuk ekstraksi nodul paru dan akurasi 86,15% untuk deteksi temuan tak terduga, membuktikan efektivitasnya dalam analisis teks berbasis domain medis. Penelitian oleh [33] menekankan pentingnya *fine-tuning* model pada korpus kesehatan untuk meningkatkan kualitas representasi semantik, dengan peningkatan *micro-F1* sekitar 90%. Sementara [24] memperkenalkan varian *BERT*, yaitu *RoBERTa* yang *pretrain* khusus untuk laporan radiologi, dengan hasil *token-level F1* 85,5%, menunjukkan peningkatan akurasi struktur kalimat dan konteks medis. Secara keseluruhan, *RoBERTa* terbukti lebih unggul dibandingkan model sejenis seperti *MediBERT* dalam tugas klasifikasi klinis, ekstraksi entitas medis, dan penyusunan ulang laporan medis. Pemilihan *RoBERTa* dibandingkan model lain didukung oleh bukti kuantitatif yang menunjukkan keunggulan akurasi dan generalisasi yang lebih tinggi dalam pemrosesan teks medis. Berdasarkan hasil perbandingan tahun 2024, *RoBERTa* menunjukkan peningkatan akurasi rata-rata 4,6–7,2% dan *F1-score* hingga 6,8% dibanding *MediBERT* pada berbagai tugas klinis. Selain itu, *HeartBERT-RoBERTa*, varian yang *di-fine-tune* khusus untuk data kesehatan, mencapai akurasi 88,48%, lebih tinggi dibanding *MediBERT* yang hanya 81–84% pada dataset yang sama. Dari segi efisiensi, *RoBERTa* memiliki mekanisme *dynamic masking* dan penghapusan *next sentence prediction* yang membuat

pretraining lebih stabil dan 35% lebih efisien. Dengan demikian, *RoBERTa* menjadi pilihan optimal untuk analisis teks medis yang kompleks dan mendukung integrasi multimodal dengan data fisiologis seperti ECG.

Pendekatan *hybrid* yang mengombinasikan *ResNet–BiLSTM–Attention* dan *RoBERTa* menjadi solusi integratif yang lebih komprehensif untuk deteksi aritmia. Berdasarkan studi [34], [19], [35], [36], dan [37], model hibrida ini mampu memanfaatkan kekuatan analisis sinyal dan teks secara simultan. Penelitian [34] menunjukkan bahwa fusi multimodal antara fitur ECG dan *clinical notes* menghasilkan akurasi 92,8% dan *macro-F1* 84,6%, terutama pada kelas minor dengan peningkatan *recall* yang signifikan. Sementara itu, [19] memperkenalkan arsitektur *Transformer-based multimodal (MoRE-like)* yang mampu meningkatkan F1-score sebesar 5–8% dibanding pendekatan unimodal. Studi [35] memperlihatkan bahwa integrasi metadata klinis dengan fitur ECG berbasis *ResNet* memberikan peningkatan AUC dan akurasi dalam prediksi penyakit kardiovaskular. Penelitian [36] memperluas pendekatan ini dengan memperkenalkan mekanisme *Lightweight Attention Fusion*, yang berfokus pada interpretabilitas model dan peningkatan *robustness* terhadap variasi dataset. Terakhir, [37] merekomendasikan penggunaan *ResNet pretrained models* untuk fusi multimodal karena kestabilannya dalam menangani data fisiologis dan tekstual secara bersamaan.

Penelitian yang menggunakan dataset PTB-XL untuk data fisiologis dan tekstual masih berfokus pada klasifikasi sinyal ECG menggunakan arsitektur konvensional seperti CNN, LSTM, atau kombinasi keduanya, dengan tujuan meningkatkan akurasi tanpa mempertimbangkan konteks klinis yang lebih luas. Pendekatan tersebut, meskipun menghasilkan akurasi tinggi, masih terbatas dalam aspek generalisasi dan interpretabilitas karena hanya memanfaatkan sinyal tunggal tanpa integrasi dengan informasi tekstual [26] [28]. Berbeda dari penelitian sebelumnya, studi ini berupaya menjembatani kesenjangan tersebut melalui pendekatan *multimodal hybrid* yang menggabungkan sinyal ECG dari dataset PTB-XL dengan laporan medis berbahasa Jerman berbasis *RoBERTa*. Dengan arsitektur

gabungan *ResNet BiLSTM Attention* dan *Transformer*, model yang diusulkan tidak hanya mampu mengenali pola spasial-temporal sinyal jantung, tetapi juga memahami konteks semantik dari deskripsi medis, sehingga memberikan interpretasi yang lebih bermakna dan meningkatkan akurasi dalam deteksi aritmia.

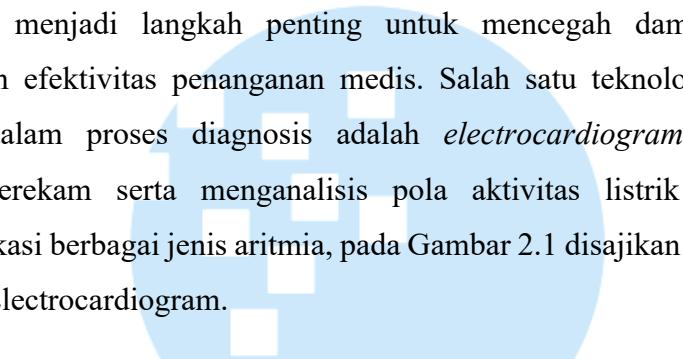
Berdasarkan Penelitian terdahulu berupa penggunaan metode untuk objek penelitian signal ECG dan *Medical Report* serta penggunaan dataset pada penelitian terdahulu, dapat disimpulkan bahwa integrasi antara analisis sinyal berbasis *ResNet BiLSTM Attention* dan pemahaman teks medis menggunakan *RoBERTa* didukung dengan data yang optimal berupa *PTB-XL* serta dibanding pendekatan *unimodal* tradisional mampu memberikan performa optimal dalam deteksi aritmia. Namun demikian, *research gap* masih ditemukan dalam hal implementasi real-time dan integrasi dengan laporan medis multibahasa, khususnya bahasa Jerman, yang menjadi tantangan penting untuk pengembangan sistem diagnosis otomatis. Oleh karena itu, penelitian ini mengusulkan model *Hybrid ResNet BiLSTM Attention* dan *RoBERTa* berbasis data multimodal ECG dan teks medis berbahasa Jerman. Pendekatan ini diharapkan mampu mendeteksi pola abnormalitas jantung secara kontekstual, cepat, dan efisien, serta mendukung implementasi klinis nyata, misalnya di lingkungan rumah sakit seperti *Dr. Sentosa Hospital*, guna meningkatkan akurasi dan kecepatan deteksi dini penyakit kardiovaskular.

2.2 Teori terkait Penelitian

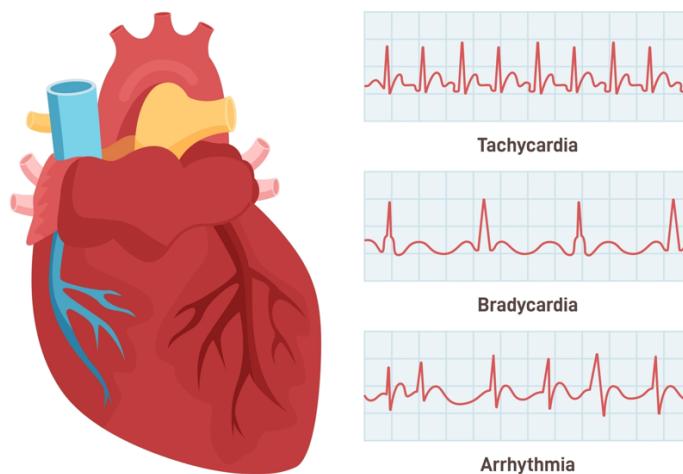
Penelitian ini didasarkan pada dua teori utama, yaitu teori medis dan teori sistem informasi. Dari sisi medis, sinyal ECG digunakan untuk menganalisis aktivitas listrik jantung guna mendeteksi *Arrhythmia*. Dari sisi sistem informasi, metode *ResNet BI LSTM Attention* memproses data ECG sebagai *time-series*, sedangkan *RoBERTa* memproses teks medis. Integrasi keduanya menghasilkan model *hybrid* yang mampu meningkatkan akurasi deteksi *Arrhythmia* secara komprehensif berikut penjelasan pada masing masing penggunaan teori:

2.2.1 *Arrhythmia*

Aritmia adalah kelainan irama jantung yang disebabkan oleh gangguan pada sistem impuls listrik yang mengatur detak jantung [35]. Kondisi ini dapat mengakibatkan jantung berdetak terlalu cepat yaitu *takikardia* [36], terlalu lambat yaitu *bradikardia*, atau secara tidak teratur [37]. Meskipun beberapa kasus aritmia bersifat ringan dan tidak menunjukkan gejala, gangguan ini juga dapat berujung pada komplikasi serius seperti stroke dan gagal jantung [38]. Oleh karena itu, deteksi dini menjadi langkah penting untuk mencegah dampak fatal dan meningkatkan efektivitas penanganan medis. Salah satu teknologi yang umum digunakan dalam proses diagnosis adalah *electrocardiogram (ECG)*, yang berfungsi merekam serta menganalisis pola aktivitas listrik jantung guna mengidentifikasi berbagai jenis aritmia, pada Gambar 2.1 disajikan Jenis Jenis Pola Pada signal Electrocardiogram.



Pathological Heart Rhythm



Gambar 2.1 Tipe Penyakit Aritmia

Sumber : [39]

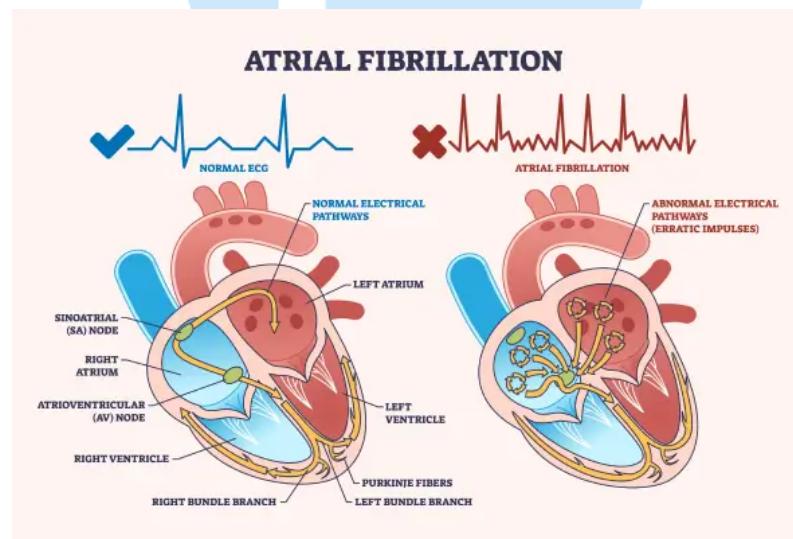
Berdasarkan Gambar 2.1, terdapat berbagai jenis *aritmia* yang dapat memengaruhi ritme jantung dengan mekanisme dan dampak yang berbeda-beda. *Normal rhythm* atau irama sinus merupakan kondisi di mana detak jantung berlangsung secara teratur dan stabil, dikendalikan oleh *Sinoatrial (SA) node* yang

berfungsi sebagai pemanas alami jantung dengan menghasilkan impuls listrik secara ritmis dan terkoordinasi [40] *Sinus arrhythmia* menggambarkan variasi kecil dalam irama jantung yang umumnya terjadi karena perubahan aktivitas sistem saraf otonom selama siklus pernapasan [41] kondisi ini biasanya bersifat fisiologis dan tidak menandakan adanya kelainan jantung yang serius. Sementara itu, *sinus tachycardia* terjadi ketika jantung berdetak lebih cepat dari normal, biasanya sebagai respons terhadap peningkatan kebutuhan oksigen atau stimulasi fisiologis seperti saat berolahraga, stres emosional, demam, atau konsumsi kafein [42]. Sebaliknya, *sinus bradycardia* merupakan kondisi di mana detak jantung lebih lambat dari normal, yang dapat terjadi secara normal pada individu dengan kebugaran tinggi, seperti atlet, namun juga dapat mengindikasikan adanya gangguan pada sistem konduksi listrik jantung apabila disertai gejala klinis tertentu [43].

Berbagai studi epidemiologis menunjukkan bahwa aritmia berkontribusi signifikan terhadap peningkatan morbiditas dan mortalitas kardiovaskular, terutama pada populasi lanjut usia dan pasien dengan komorbiditas penyakit jantung struktural [57], [58]. Oleh karena itu, deteksi dini menjadi langkah penting untuk mencegah dampak fatal dan meningkatkan efektivitas penanganan medis. Salah satu teknologi yang umum digunakan dalam proses diagnosis adalah electrocardiogram (ECG), yang berfungsi merekam serta menganalisis pola aktivitas listrik jantung guna mengidentifikasi berbagai jenis aritmia. ECG diakui secara luas sebagai modalitas diagnostik non-invasif yang esensial karena mampu menangkap perubahan temporal dan morfologis aktivitas listrik jantung secara real-time [59]. Stabilitas irama sinus mencerminkan integritas sistem konduksi jantung dan menjadi indikator utama fungsi elektrofisiologis jantung yang sehat [60]. *Sinus arrhythmia* menggambarkan variasi kecil dalam irama jantung yang umumnya terjadi karena perubahan aktivitas sistem saraf otonom selama siklus pernapasan [41]. Fenomena ini sering ditemukan pada individu sehat dan secara klinis dianggap sebagai respons fisiologis normal terhadap modulasi vagal [61].

2.2.2 Atrial Fibrillation

Atrial Fibrillation (AF) merupakan salah satu bentuk aritmia jantung yang paling umum dan kompleks, ditandai oleh aktivitas listrik atrium yang tidak teratur dan sering kali sangat cepat, sehingga menyebabkan kontraksi atrium yang tidak sinkron [44]. Kondisi ini berdampak pada hemodinamika jantung dan meningkatkan risiko komplikasi serius, termasuk stroke, gagal jantung, dan kematian akibat penyakit kardiovaskular [45]. AF umumnya bersifat paroksismal, seperti terlihat pada Gambar 2.2, namun jika tidak ditangani dengan tepat, dapat berkembang menjadi bentuk persisten atau permanen. Deteksi dini dan akurat terhadap *Atrial Fibrillation* (AF) menjadi sangat penting, karena intervensi medis yang tepat waktu dapat secara signifikan menurunkan risiko komplikasi jangka panjang serta meningkatkan prognosis pasien [46].



Gambar 2.2 Ritme Atrial Fibrillation

Sumber : [47]

Berdasarkan Gambar 2.2, ditunjukkan perbandingan antara ritme sinus normal dan *atrial fibrillation* dalam aktivitas listrik jantung. Pada ritme sinus normal, impuls listrik dimulai di nodus *sinoatrial* (SA) dan merambat secara teratur melalui atrium menuju nodus *atrioventricular* (AV) [48], sehingga menghasilkan kontraksi jantung yang terkoordinasi dan efisien, terlihat pada pola ECG dengan gelombang P-QRS-T

yang teratur [25]Sebaliknya, *atrial fibrillation* ditandai oleh impuls listrik yang kacau, menyebabkan atrium bergetar dan ventrikel menerima sinyal secara tidak teratur, menghasilkan detak jantung yang cepat dan tidak stabil Kondisi ini menghasilkan pola *electrocardiogram (ECG)* yang khas berupa hilangnya gelombang P yang jelas dan interval R–R yang tidak teratur, yang menjadi indikator utama dalam diagnosis klinis *Atrial Fibrillation (AF)* [49]. Penelitian longitudinal menunjukkan bahwa AF meningkatkan risiko stroke iskemik hingga lima kali lipat dibandingkan individu dengan irama sinus normal [62]. Deteksi dini dan akurat terhadap AF menjadi sangat penting karena intervensi medis yang tepat waktu dapat secara signifikan menurunkan risiko komplikasi jangka panjang [46]. Guideline kardiologi internasional menekankan pentingnya skrining AF berbasis ECG untuk pencegahan stroke primer dan sekunder [63].

2.2.3 *Ventricular Tachycardia*

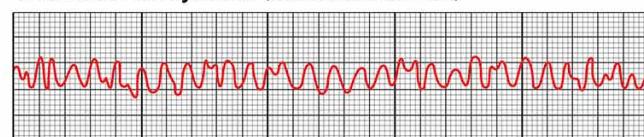
Ventricular Tachycardia (VT) adalah kondisi aritmia yang berasal dari ventrikel jantung, ditandai oleh irama cepat yang tidak normal dengan frekuensi lebih dari 100 denyut per menit [50] . VT merupakan masalah serius karena dapat menyebabkan penurunan output jantung yang signifikan, memicu gagal jantung akut, sinkop, atau bahkan kematian mendadak [51]. VT umumnya diklasifikasikan menjadi dua bentuk utama *sustained VT*, yang berlangsung lebih dari 30 detik seperti terlihat pada Gambar 2.3 dan memerlukan intervensi medis untuk penghentian, serta *non-sustained VT*, yang bersifat sementara dan berhenti sendiri. Secara klinis, VT dapat terjadi pada individu dengan atau tanpa penyakit jantung struktural [52]

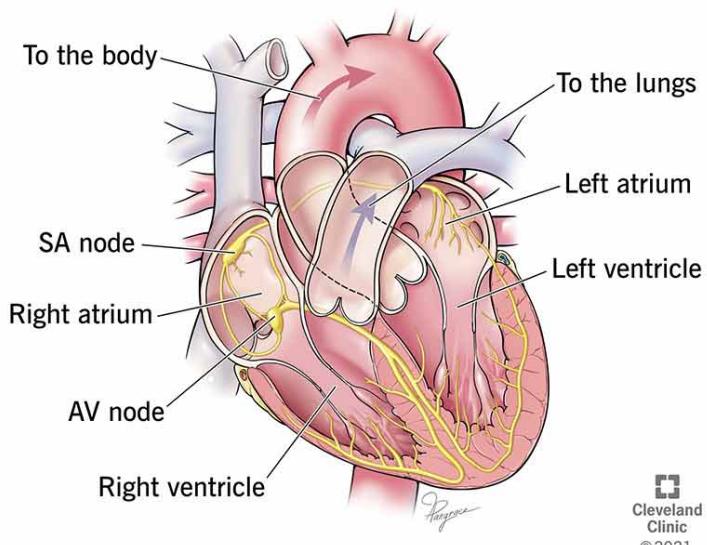
Ventricular arrhythmia

Normal sinus rhythm



Ventricular arrhythmia (fibrillation shown)





Cleveland
Clinic
©2021

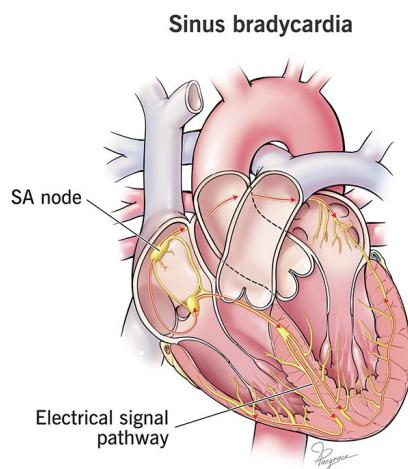
Gambar 2.3 Ritme *Ventricular Tachycardia*

Sumber :[53]

Berdasarkan Gambar 2.3, perbandingan antara ritme sinus normal dan *ventricular tachycardia* (VT) dalam aktivitas listrik jantung menunjukkan perbedaan yang cukup signifikan. Pada ritme sinus normal, impuls listrik dimulai di nodus *sinoatrial* (SA) dan merambat secara teratur melalui atrium menuju nodus *atrioventricular* (AV), kemudian diteruskan ke ventrikel, sehingga menghasilkan kontraksi jantung yang terkoordinasi dan efisien [54]. Proses ini memastikan pemompaan darah yang optimal ke seluruh tubuh dan tercermin pada pola ECG yang teratur dengan gelombang P-QRS-T yang konsisten dan dapat diprediksi. Sebaliknya, pada VT, impuls listrik abnormal berasal dari ventrikel itu sendiri, melewati jalur listrik normal atrium secara tidak sinkron [51]. Hal ini menyebabkan ventrikel berkontraksi secara cepat dan tidak terkoordinasi, sehingga kontraksi atrium dan ventrikel tidak selaras. Kondisi ini menghasilkan detak jantung yang sangat cepat, efisiensi pemompaan darah menurun, dan pola ECG menunjukkan gelombang QRS yang lebar dan abnormal [55]. VT secara klinis dikaitkan dengan risiko tinggi terjadinya sudden cardiac death, khususnya pada pasien dengan penyakit jantung iskemik dan kardiomiopati [64].

2.2.4 Normal Sinus Rhyhtm

Normal Sinus Rhythm (NSR) merupakan irama jantung normal yang terjadi ketika impuls listrik berasal dari nodus *sinoatrial* (SA node), yang berperan sebagai pacemaker alami jantung [39]. NSR ditandai dengan gelombang P yang teratur, diikuti oleh kompleks QRS dan gelombang T pada setiap siklus denyut. Interval antar gelombang R-R konsisten dan ritmis, menunjukkan bahwa aktivitas listrik jantung berjalan normal dari atrium menuju ventrikel. Pada orang dewasa, frekuensi denyut jantung pada NSR berkisar antara 60–100 kali per menit, seperti ditunjukkan pada Gambar 2.4, yang mencerminkan fungsi jantung yang efisien dan koordinasi kontraksi atrium serta ventrikel yang optimal [56]. Pola ini menunjukkan keseimbangan antara sistem saraf simpatis dan parasimpatis dalam mengatur aktivitas jantung secara fisiologis. NSR menjadi acuan penting dalam analisis sinyal *electrocardiogram* (*ECG*) karena setiap penyimpangan dari pola ini dapat mengindikasikan adanya gangguan irama atau kondisi patologis tertentu. Oleh karena itu, pemahaman karakteristik NSR sangat krusial dalam proses klasifikasi aritmia berbasis pembelajaran mendalam [44].



Gambar 2.4 *Normal Sinus Rhythm*

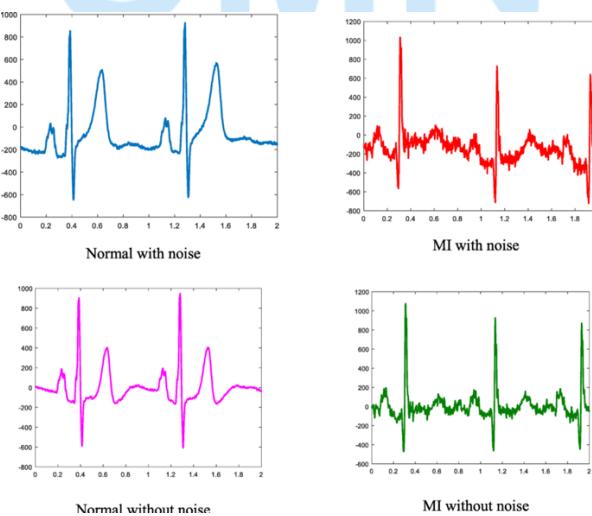
Sumber : [57]

Berdasarkan Gambar 2.4, secara literatur *Normal Sinus Rhythm* (NSR) didefinisikan sebagai irama jantung fisiologis yang mencerminkan fungsi normal sistem konduksi jantung, di mana nodus *sinoatrial* (SA) berperan sebagai pengatur

utama denyut [58]. Karakteristik utama NSR meliputi gelombang P positif pada sadapan standar (lead II), interval PR yang konsisten berkisar 0,12–0,20 detik setiap gelombang P diikuti oleh kompleks QRS sempit (<120 ms), serta interval R-R yang teratur [59]. Irama ini menunjukkan tidak adanya gangguan atrial maupun ventrikular, sehingga NSR menjadi acuan dasar dalam menilai kelainan jantung seperti aritmia, blok konduksi, atau *fibrillation* Myocardial dan Infarction Myocardial Infarction [45]. NSR digunakan sebagai baseline referensi klinis dalam evaluasi ECG karena mencerminkan kondisi elektrofisiologis jantung yang optimal dan stabil [65]. Pemahaman karakteristik NSR sangat krusial dalam proses klasifikasi aritmia berbasis pembelajaran mendalam [44].

2.2.5 Myocardial Infarction (MI)

Myocardial Infarction (MI) merupakan kondisi kritis yang terjadi akibat tersumbatnya aliran darah ke otot jantung, menyebabkan kerusakan jaringan jantung [60]. Pada sinyal *ECG*, MI biasanya ditandai dengan elevasi atau depresi segmen ST, gelombang T terbalik, dan munculnya gelombang Q patologis yang menandai area jantung yang mengalami infark. Gambar 2.5 memperlihatkan contoh sinyal *ECG* pasien dengan MI, di mana elevasi segmen ST dapat diamati pada beberapa lead sehingga menjadi indikator utama dalam deteksi diagnosis *Myocardial Infarction*.



Gambar 2.5 Signal Pada MI

Sumber : [45]

Berdasarkan Gambar 2.5, elevasi segmen ST menunjukkan adanya cedera akut pada otot ventrikel, sedangkan gelombang Q patologis menandakan jaringan jantung yang telah mati akibat infark [61]. Inversi gelombang T merefleksikan fase pemulihan ventrikel dan perubahan repolarisasi. Pola ini memungkinkan tenaga medis untuk menentukan lokasi dan luas infark, menilai keparahan kondisi, serta mengambil keputusan intervensi seperti terapi reperfusi atau pemantauan intensif. Analisis mendalam dari kombinasi perubahan ST, Q, dan T sangat penting untuk diagnosis dini dan pengelolaan klinis yang efektif [62].

2.2.6 STTC ST-T Change

ST-T Change (STTC) mencerminkan perubahan pada segmen ST dan gelombang T yang dapat menjadi indikator awal iskemia, stres jantung, atau gangguan elektrolit [63]. Pada sinyal *ECG*, STTC biasanya muncul sebagai depresi atau elevasi segmen ST, serta inversi atau perubahan bentuk gelombang T, tanpa munculnya gelombang Q patologis. Gambar 2.6 menampilkan pola STTC, di mana segmen ST terlihat menyimpang dari garis isoelektrik dan gelombang T tidak simetris atau terbalik, menandakan adanya stres pada otot jantung .



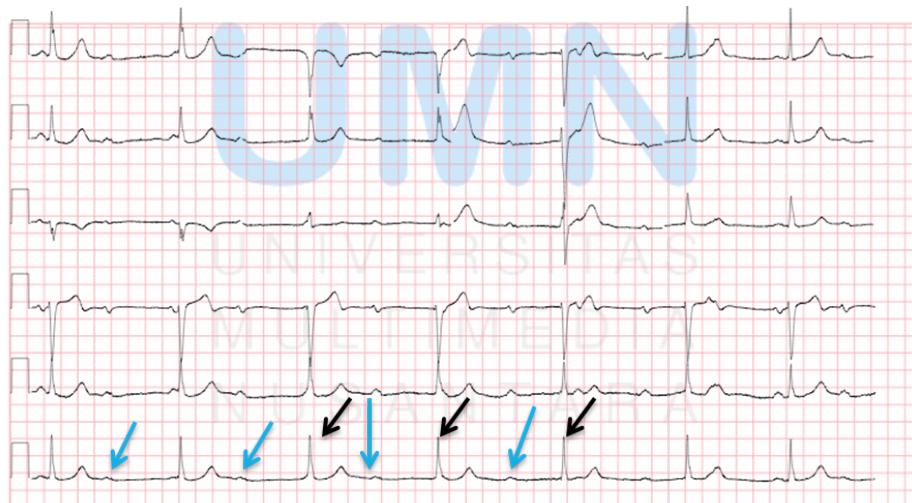
Gambar 2.6 STTC Signal

Sumber : [64]

Berdasarkan Gambar 2.6, perubahan segmen ST dan gelombang T memberikan informasi mengenai kemungkinan gangguan perfusi jantung atau kondisi sementara yang memengaruhi repolarisasi ventrikel [65]. Depresi segmen ST dapat menunjukkan iskemia subendokardial, sedangkan elevasi ST yang ringan bisa menandakan cedera jantung awal. Analisis STTC memungkinkan tenaga medis mendeteksi kondisi yang belum mengarah pada infark penuh, sehingga intervensi atau pemantauan dapat dilakukan lebih dini untuk mencegah komplikasi lebih lanjut [66].

2.2.7 Conduction Disorder

Conduction Disorder (CD) merupakan gangguan pada sistem konduksi listrik jantung, yang menyebabkan impuls listrik tidak menyebar secara normal dari atrium ke ventrikel [67]. Pada sinyal *ECG*, CD dapat muncul sebagai blok cabang kiri atau kanan, interval PR yang memanjang, atau kompleks QRS melebar [68]. Gambar 2.7 memperlihatkan contoh CD, di mana gelombang QRS melebar, interval R-R tidak konsisten, dan beberapa gelombang P tidak diikuti oleh kompleks QRS yang tepat.



Gambar 2.7 *Conduction Disorder Signal*

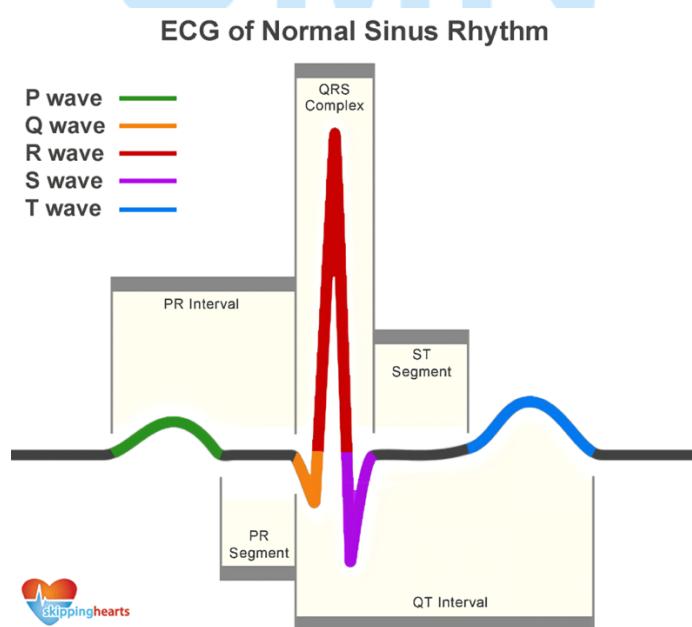
Sumber : [67]

Berdasarkan Gambar 2.7, pola *ECG* pada CD menunjukkan keterlambatan atau hambatan konduksi listrik dalam sistem jantung. Hal ini dapat menyebabkan

kontraksi atrium dan ventrikel tidak sinkron, menurunkan efisiensi pemompaan darah. Analisis gelombang QRS, interval PR, dan ritme R-R membantu menentukan tipe CD, seperti blok atrioventrikular atau blok cabang, serta menilai risiko komplikasi seperti gagal jantung atau aritmia serius [69]. Informasi ini sangat penting dalam menentukan strategi pengelolaan klinis, termasuk kebutuhan penggunaan *pacemaker* atau terapi obat spesifik. Selain itu, pemantauan jangka panjang menggunakan *Holter monitor* dapat membantu mendeteksi fluktuasi konduksi yang mungkin tidak muncul pada pemeriksaan *ECG* singkat di klinik.

2.2.8 Sinyal Elektrokardiogram

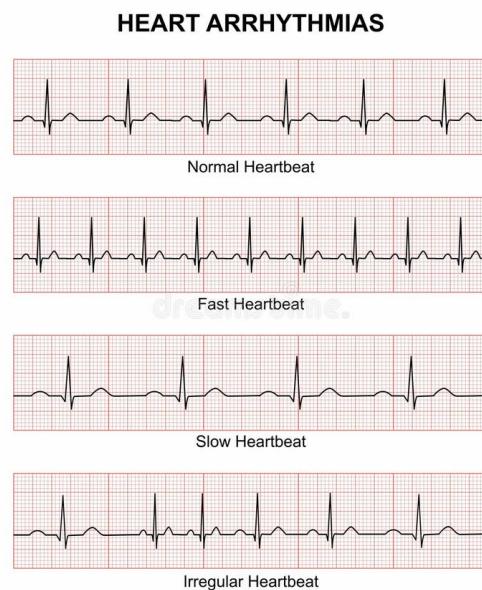
Sinyal elektrokardiogram (*EKG*) yang ditampilkan pada Gambar 2.8 merupakan representasi grafis dari aktivitas listrik jantung yang diukur melalui elektroda yang ditempatkan di permukaan tubuh. Sinyal ini menggambarkan impuls listrik yang dihasilkan oleh jantung selama proses kontraksi (sistol) dan relaksasi (diastol) [70], sehingga memberikan gambaran lengkap tentang fungsi listrik jantung. Selain menampilkan irama dan frekuensi denyut jantung, *EKG* juga menyediakan informasi tentang konduksi listrik melalui atrium dan ventrikel, pola gelombang P-QRS-T, serta potensi adanya kelainan elektrofisiologis seperti aritmia, blok konduksi, atau iskemia. Dengan demikian, *EKG* menjadi alat diagnostik penting untuk memantau kesehatan jantung secara non-invasif [71].



Gambar 2.8 Sinyal *Elektodiagram* Pada Deteksi Aritmia

Sumber : [72]

Perbedaan antara sinyal *ECG* (*Electrocardiogram*) normal dan abnormal ditunjukkan pada Gambar 2.8, yang menyajikan berbagai jenis detak jantung. Komponen utama sinyal *ECG* terdiri dari gelombang P, kompleks QRS, dan gelombang T [73]. Gelombang P mencerminkan aktivasi atrium, yaitu proses penyebaran impuls listrik melalui atrium (serambi jantung) yang memicu kontraksi untuk memompa darah ke ventrikel [51]. Kompleks QRS menggambarkan aktivasi ventrikel, di mana impuls listrik menyebabkan kontraksi ventrikel (bilik jantung) untuk memompa darah ke seluruh tubuh [74]. Gelombang T menunjukkan fase repolarisasi, yaitu saat ventrikel kembali ke kondisi istirahat setelah kontraksi [75]. Selanjutnya, Gambar 2.9 memperlihatkan pola-pola sinyal *electrocardiogram* yang dapat digunakan untuk deteksi aritmia, memberikan referensi visual penting dalam membedakan sinyal normal dari abnormal.



Gambar 2.9 Pola Elektodiagram pada deteksi Aritmia

Sumber : [4]

Detak jantung normal yang ditampilkan pada Gambar 2.9 menunjukkan pola teratur dengan interval seimbang antara gelombang P, QRS, dan T, mencerminkan aktivitas listrik jantung yang stabil dan sehat. Detak jantung cepat (*takikardia*) ditandai oleh frekuensi denyut yang tinggi dan interval gelombang yang sangat pendek, yang sering kali dipicu oleh stres, aktivitas fisik berat, atau kondisi medis tertentu [45]. Sebaliknya, detak jantung lambat (*bradikardia*) memperlihatkan interval gelombang yang lebih panjang, mencerminkan ritme yang lebih rendah, yang dapat disebabkan oleh gangguan konduksi listrik jantung atau efek samping obat tertentu [76]. Detak jantung tidak teratur (*aritmia*) memiliki pola gelombang yang tidak konsisten, dengan interval yang bervariasi, sering kali akibat gangguan pada sistem listrik jantung seperti *atrial* atau *ventricular fibrillation* [77]. Informasi ini menunjukkan bagaimana sinyal *ECG* dapat digunakan untuk mendiagnosis berbagai kondisi jantung berdasarkan variasi ritme dan frekuensi denyut. Analisis pola-pola tersebut sangat penting karena membantu tenaga medis dalam mendeteksi dini potensi gangguan kardiovaskular dan menentukan intervensi yang tepat untuk mencegah komplikasi serius.

2.2.9 *Medical report*

Laporan medis (*Medical report*) merupakan dokumen yang menyajikan informasi lengkap mengenai kondisi pasien, meliputi data identitas, riwayat penyakit, hasil pemeriksaan laboratorium, serta interpretasi dari tenaga medis. Dalam konteks penelitian kesehatan digital, laporan medis menjadi sumber data yang sangat bernilai karena menyediakan informasi yang dapat digunakan untuk analisis penyakit dan pengambilan keputusan klinis [78]. Saat ini, laporan medis dapat diproses secara otomatis untuk mengenali istilah-istilah medis, memahami konteks klinis, dan mengekstraksi informasi yang relevan.

Pemanfaatan laporan medis dalam penelitian kesehatan digital terus berkembang seiring kemajuan teknologi *Natural Language Processing* (NLP), khususnya model seperti *RoBERTa*[79]. Dengan teknik ini, data teks yang awalnya tidak terstruktur dapat diubah menjadi informasi yang tersusun dan siap untuk dianalisis secara otomatis. Pendekatan ini memungkinkan peneliti dan tenaga medis mengidentifikasi pola penting, misalnya hubungan antara gejala dan diagnosis, serta

mendukung pengembangan sistem *clinical decision support* berbasis kecerdasan buatan [80]. Gambar 2.10 menampilkan contoh laporan medis dalam bahasa Jerman.

```
■ Loading PTB-XL Database...
 
⌚ Report Rank 3783
Length: 397 characters
SCP Codes: {'ASMI': 50.0, 'IVCD': 100.0, 'IMI': 100.0}
⌚ Report:
sinusbradykardi extrem vänster el-axel ospecifict skänkelblock avvikande qrs(t) förlopp anteroseptal infarkt Alder ej bestÄ
mbar inferior infarkt Alder ej bestÄmbarconsensus label derived from: {'IMI': 100.0, 'ASMI': 100.0, 'IVCD': 100.0},{'IM
I': 100.0, 'ASMI': 50.0, 'LVH': 50.0, 'ISC_': 100.0, 'IVCD': 100.0},{'IMI': 100.0, 'ASMI': 50.0, 'LVH': 50.0, 'ISC_': 100.0, 'IVC
D': 100.0}
=====
⌚ Report Rank 1674
Length: 359 characters
SCP Codes: {'INJAS': 100.0, 'INJAL': 100.0, 'LVH': 100.0, 'VCLVH': 0.0, 'INVT': 0.0, 'STD_': 0.0, 'SR': 0.0}
⌚ Report:
sinus rhythm. r-s transition zone in v leads displaced to the left. voltages are high in limb leads suggesting lvh. st segments ar
e depressed in i, avl, v6. t waves are inverted in i, ii, avl, v2-6. this may be due to lv strain or ischaemia. prolonged qt, Edi
t: INJAS, INJAL, Sokolow 3,1 aber sehr hohe Amplituden in Standardableitungen(INJAS, INJAL, LVH 100)
=====
⌚ Report Rank 1358
Length: 340 characters
SCP Codes: {'ASMI': 100.0, 'ALMI': 100.0, 'IAVB': 100.0, 'SR': 0.0}
⌚ Report:
regelbunden rytm, inga p-vÄg funna lÄga qrs amplitudet i extremitetsavleddningarna avvikande qrs(t) förlopp anterior infarkt
mÖjigen fÄrsk hÖg lateral infarkt bÖr ÖvervÄgasconsensus label derived from: {'ASMI': 100.0, 'ALMI': 100.0, 'IAVB': 100.0, 'S
R': 0.0},{'ASMI': 100.0, 'ALMI': 100.0, 'IVCD': 100.0, 'IAVB': 100.0, 'SR': 0.0}
=====
⌚ Report Rank 4957
Length: 327 characters
SCP Codes: {'ASMI': 15.0, 'LVH': 100.0, 'ISC_': 100.0, 'LAFB': 100.0, 'VCLVH': 0.0, 'QWAVE': 0.0, 'STD_': 0.0, 'SR': 0.0}
⌚ Report:
sinus rhythm. left axis deviation. non-specific intraventricular delay. voltages are high in chest leads suggesting lvh. q waves,
small r waves and st segment elevation in v2,3 suggesting anteroseptal myocardial infarction of uncertain age, possibly recen
Edi
t: ASMI 50, LVH 100, ISC_, Sokolow 4,5(ASMI 15, LVH 100, ISC_, LAFB)
=====
```

Gambar 2.10 *Medical report* dalam Penyakit Aritmia

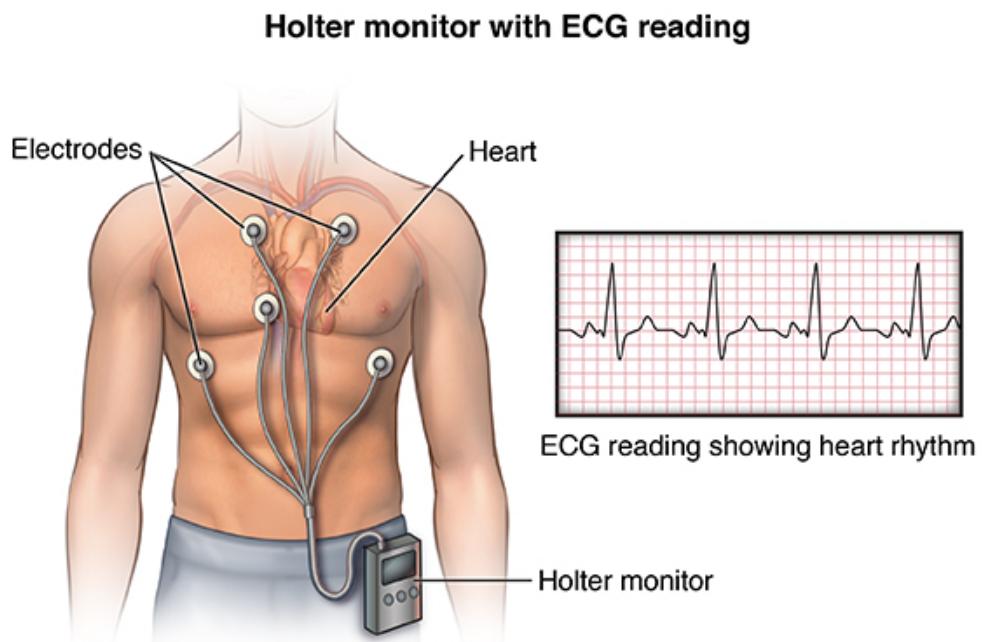
Sumber : [30]

Laporan medis yang ditampilkan pada Gambar 2.10 memuat informasi penting mengenai kondisi pasien, termasuk hasil pemeriksaan fisik, diagnosis awal, pengukuran klinis, serta interpretasi medis. Laporan ini ditulis dalam bahasa Jerman dan menggunakan terminologi medis khusus yang membutuhkan pemahaman mendalam agar dapat diolah secara otomatis [81]. Beberapa elemen yang tercakup dalam laporan tersebut antara lain catatan pemeriksaan jantung, indikasi gejala *arrhythmia*, serta parameter teknis seperti tekanan darah, denyut jantung, dan hasil tes *electrocardiogram (ECG)* [82]. Pemrosesan teks medis semacam ini dilakukan menggunakan model *RoBERTa*, yang terbukti unggul dalam memahami konteks bahasa medis domain-spesifik dengan akurasi lebih tinggi dibanding model bahasa umum. Dengan menggabungkan informasi dari laporan medis dan data sinyal *ECG*, sistem deteksi *arrhythmia* dapat menghasilkan diagnosis yang lebih akurat, cepat,

dan komprehensif, sehingga meningkatkan efisiensi dan kualitas praktik klinis modern.

2.2.10 *Holter Monitor*

Holter monitor adalah perangkat portabel yang digunakan untuk merekam aktivitas listrik jantung secara kontinu selama periode tertentu, biasanya antara 24 hingga 48 jam. Alat ini terdiri dari elektroda yang ditempelkan pada dada pasien dan terhubung ke unit perekam kecil yang dapat dibawa dalam kantong atau digantung pada sabuk [57]. Fungsi utama *holter monitor* adalah untuk mendeteksi kelainan ritme jantung (*arrhythmia*), iskemia, atau perubahan lain yang mungkin tidak terlihat pada rekaman *ECG* standar dengan durasi singkat. Gambar 2.11 memperlihatkan ilustrasi alat tersebut.



Gambar 2.11 *Holter Monitor* Pada Klasifikasi Aritmia

Sumber : [83]

Berdasarkan Gambar 2.11, *Holter monitor* mampu merekam sinyal *electrocardiogram* secara real-time selama aktivitas sehari-hari pasien, sehingga menghasilkan data yang lebih representatif dibandingkan pemeriksaan di klinik [84]. Data yang diperoleh umumnya disimpan dalam format .dat dan .hea, yang

berisi sinyal *ECG* mentah beserta metadata terkait pasien, durasi rekaman, dan parameter teknis lainnya. Dalam penelitian ini, *Holter monitor* berperan sebagai salah satu sumber data utama, karena format .dat dan .hea memungkinkan pemrosesan lebih lanjut menggunakan model *deep learning* untuk deteksi *arrhythmia* secara multimodal [85].

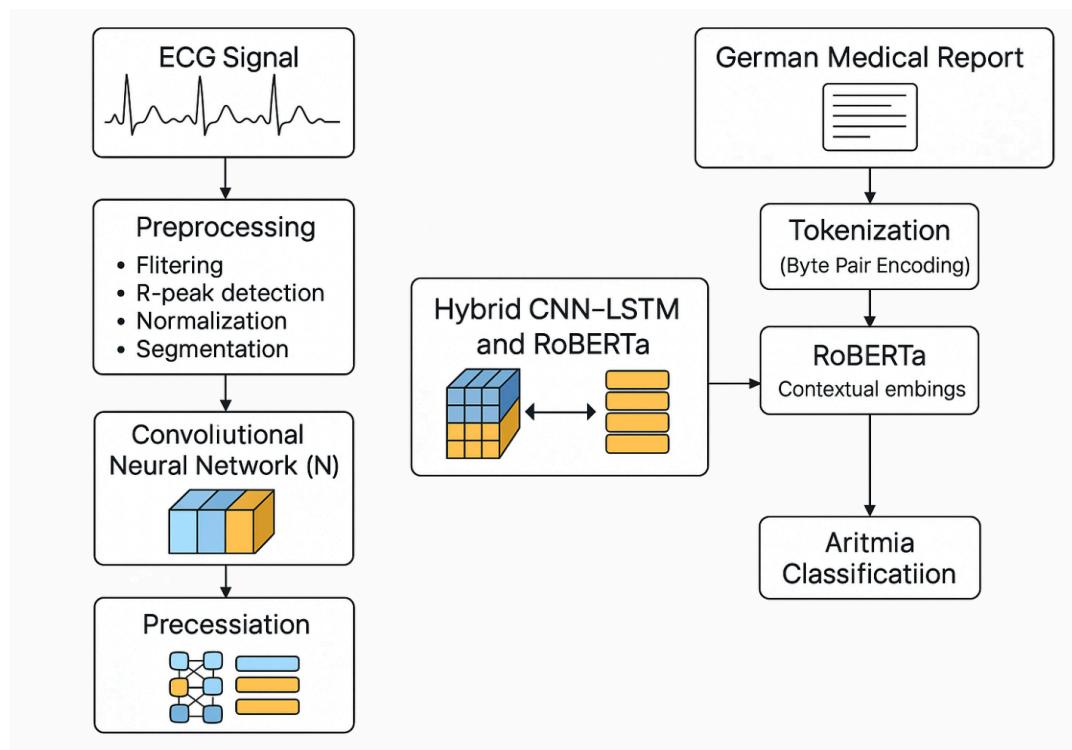
2.2.11 *ECG Signal* dan *Medical Report* menjadi Representasi Vektor

Sinyal ECG merupakan rangkaian nilai amplitudo terhadap waktu yang bersifat kontinu dan merepresentasikan aktivitas listrik jantung. Untuk dapat diolah oleh model *deep learning*, sinyal ini perlu dikonversi terlebih dahulu menjadi bentuk representasi numerik yang terstruktur atau vektor fitur [86]. Tahapan awal dilakukan melalui proses *preprocessing* guna mengekstraksi segmen sinyal dengan durasi tetap agar memiliki ukuran input yang konsisten. Selanjutnya, *ResNet* digunakan untuk mengekstraksi karakteristik spasial dari setiap segmen sinyal, mengubah pola gelombang menjadi *feature map* multidimensi yang merekam informasi morfologi sinyal [87]. Keluaran *ResNet* kemudian diratakan (*flattened*) menjadi vektor numerik yang menggambarkan bentuk dan karakteristik sinyal ECG. Vektor tersebut selanjutnya diproses untuk menangkap keterkaitan temporal antarsegmen secara berurutan, sehingga dihasilkan representasi akhir yang mampu mencerminkan dinamika serta keteraturan pola ritme jantung.

Medical Report berbahasa Jerman terlebih dahulu melalui proses *tokenization* menggunakan *Byte Pair Encoding (BPE)* [30]. Proses ini memecah teks menjadi unit sub-kata (*tokens*), yang kemudian diubah menjadi representasi numerik melalui *embedding layer* pada model *RoBERTa*. Setiap token diubah menjadi vektor berdimensi tinggi yang merepresentasikan makna semantik dan konteks linguistiknya. Melalui mekanisme *self-attention*, *RoBERTa* menghasilkan *contextual embeddings* yang mampu memahami istilah medis secara kontekstual [88].

Sinyal ECG berupa deretan amplitudo waktu, seperti $[0.12, 0.18, 0.15, 0.30]/[0.12, 0.18, 0.15, 0.30]$ [89], diproses oleh CNN untuk

mengekstraksi pola spasial dan menghasilkan vektor fitur, misalnya $[0.21, 0.34, 0.28, 0.15]$ $[0.21, 0.34, 0.28, 0.15]$, yang kemudian dianalisis oleh LSTM guna memahami pola ritme jantung. Sementara itu, *Medical report* berbahasa Jerman diolah melalui *tokenization* dan *embedding* RoBERTa, kata medis seperti “*Arrhythmie*” direpresentasikan dalam bentuk vektor $[0.21, -0.35, 0.48, 0.10]$ $[0.21, -0.35, 0.48, 0.10]$ [82]. Hasil dari kedua sumber data ini digabungkan menjadi representasi multimodal berikut Gambar 2.12 penjelasan pada alur Vektor. Gabungan representasi tersebut kemudian melewati lapisan *fusion* yang bertujuan untuk mengintegrasikan informasi fisiologis dari sinyal ECG dan konteks semantik dari laporan medis, sehingga model dapat melakukan klasifikasi aritmia dengan mempertimbangkan baik aspek biologis maupun klinis secara bersamaan.



Gambar 2.12 Proses Perubahan Kedua jenis Data Menjadi Vektor

Sumber : [83]

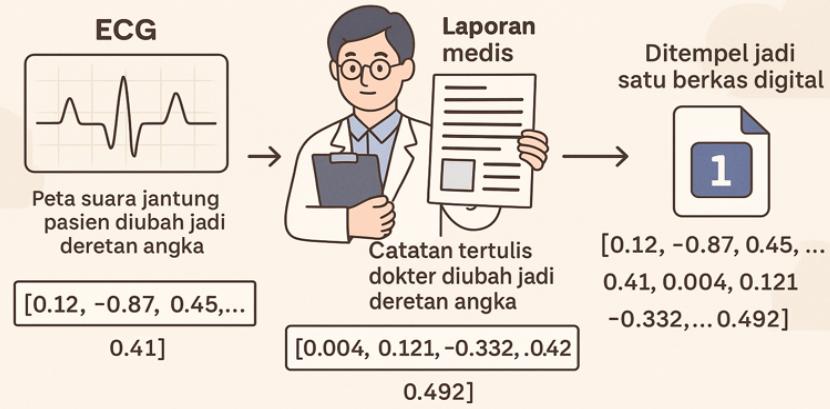
Setelah kedua modalitas, yaitu sinyal ECG dan teks medis, dikonversi menjadi bentuk vektor sebagaimana ditunjukkan pada Gambar 2.12, tahap berikutnya adalah melakukan proses penggabungan (*fusion*) terhadap kedua representasi tersebut [90].

Tujuan utama dari tahap ini adalah memungkinkan model untuk mempelajari hubungan antara kondisi fisiologis yang tercermin pada sinyal jantung dan informasi klinis yang dijelaskan dalam teks medis. Representasi vektor hasil yang menggambarkan pola temporal serta dinamika sinyal ECG, kemudian digabungkan dengan *contextual embedding* dari model RoBERTa menggunakan pendekatan *concatenation* atau *feature-level fusion* [91]. Hasil dari proses penggabungan ini berupa satu vektor multimodal yang memuat kombinasi informasi spasial-temporal dari sinyal ECG dan konteks semantik dari teks medis. Vektor multimodal tersebut menjadi representasi terpadu yang kaya akan informasi, sehingga model dapat memahami korelasi antara fitur fisiologis dan deskripsi klinis secara lebih mendalam. Setelah terbentuk, vektor gabungan ini diteruskan ke lapisan *fully connected* untuk proses klasifikasi jenis aritmia berdasarkan pola yang telah dipelajari dari kedua sumber data tersebut.

2.2.12 Integrasi ECG Signal dengan *Medical report*

Proses penggabungan fitur dalam sistem deteksi aritmia berbasis *multimodal deep learning* bertujuan untuk menyatukan dua sumber informasi yang berbeda namun saling melengkapi, yaitu sinyal fisiologis dan data textual. Sinyal elektrokardiogram (ECG) menggambarkan aktivitas kelistrikan jantung secara kontinu dan memiliki pola kompleks yang mencerminkan kondisi fisiologis pasien. Sementara itu, laporan medis berisi konteks klinis seperti deskripsi gejala, diagnosis, serta riwayat pengobatan yang memberikan pemahaman semantik terhadap kondisi pasien. Secara teoritis, kedua modalitas ini mewakili dua domain data yang berbeda numerik dan linguistik yang dapat dilihat pada Gambar 2.13 sehingga diperlukan pendekatan penggabungan (*feature fusion*) untuk menghasilkan representasi yang lebih informatif dan komprehensif bagi model klasifikasi [92]. Pendekatan multimodal terbukti mampu meningkatkan akurasi dan robustness sistem deteksi aritmia dibandingkan model unimodal, karena integrasi sinyal fisiologis dan informasi klinis memungkinkan model menangkap hubungan kompleks antara pola elektrofisiologis dan konteks medis pasien secara lebih holistik [93].

Dari Gelombang ke Kata, Mengapa Keduanya Bisa Menyatu?



Gambar 2.13 Integrasi Sinyal ECG dan *Medical report*

Sumber : [93]

Gambar 2.13 dalam proses penggabungan dilakukan, dengan masing masing modalitas diproses melalui model yang sesuai dengan karakteristik datanya. Sinyal ECG diolah menggunakan jaringan konvolusional untuk mengekstraksi fitur temporal dan morfologis yang merepresentasikan pola ritme jantung. Sementara itu, teks laporan medis dikonversi menjadi representasi vektor melalui model *language representation* seperti RoBERTa yang mampu memahami konteks dan makna antar-kata dalam kalimat. Hasil dari kedua proses ini adalah dua vektor fitur berdimensi tetap yang mencerminkan esensi informasi dari masing-masing modalitas [29]. Agar keduanya dapat dibandingkan atau digabungkan secara efektif, dilakukan proses normalisasi atau penyetaraan dimensi agar skala nilai fitur menjadi seimbang.

Penggabungan kedua vektor dalam Gambar 2.13 berupa fitur tersebut menggunakan teknik *feature-level fusion*. Metode yang paling umum adalah *concatenation*, yaitu penggabungan langsung antara vektor fitur ECG dan teks menjadi satu representasi terpadu yang kemudian diproses oleh lapisan dense untuk mempelajari hubungan non-linear antar-modalitas. Alternatif lainnya mencakup pendekatan yang lebih kompleks seperti *attention mechanism* atau *gated fusion*, yang memungkinkan model mempelajari bobot kontribusi masing-masing

modalitas secara adaptif [94]. Dengan mengintegrasikan informasi fisiologis dari sinyal ECG dan konteks klinis dari laporan medis, proses penggabungan ini menghasilkan representasi multimodal yang kaya, meningkatkan kemampuan model dalam mengenali pola aritmia secara lebih akurat dan kontekstual dibandingkan pendekatan berbasis satu modalitas saja.

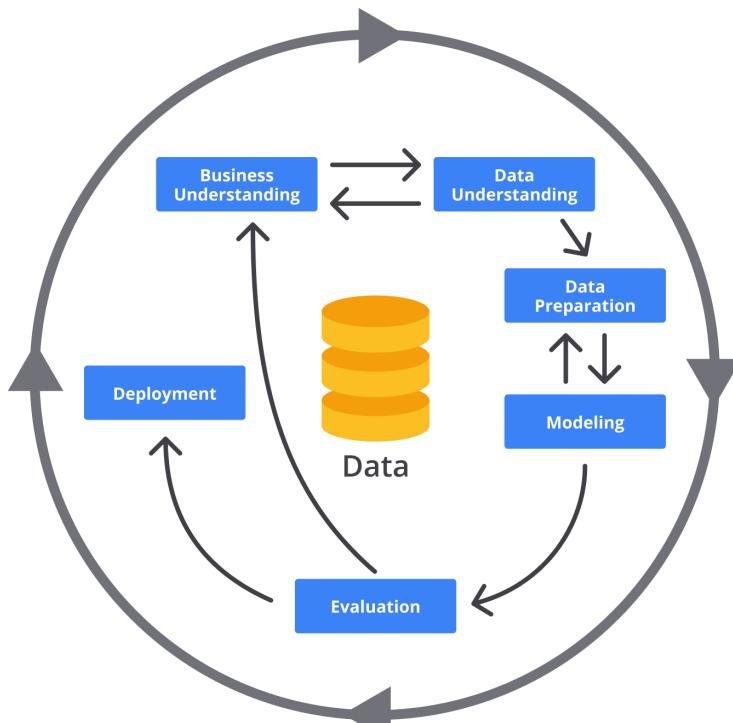
2.3 Framework dan Algoritma

Penelitian ini menerapkan kerangka kerja *Cross-Industry Standard Process for Data Mining (CRISP-DM)* sebagai tahapan utama dalam pengolahan data, dengan memanfaatkan arsitektur *ResNet–Bidirectional LSTM–Attention* untuk pemrosesan sinyal ECG dan model *RoBERTa* untuk analisis teks medis [81]. Kedua model ini kemudian digabungkan dalam suatu pendekatan *hybrid multimodal* yang dirancang untuk meningkatkan akurasi dan kemampuan generalisasi sistem dalam mendeteksi aritmia. Integrasi ini memungkinkan model untuk mempelajari hubungan antara pola fisiologis pada sinyal ECG dan konteks klinis dari laporan medis, sehingga menghasilkan sistem deteksi yang lebih adaptif.

2.3.1 Cross Industry Standard Process for Data Mining

Cross-Industry Standard Process for Data Mining (CRISP-DM) merupakan kerangka kerja yang banyak diadopsi dalam analisis data dan *data mining* untuk memastikan proses penggalian informasi dari data berlangsung secara sistematis, terstruktur, dan efisien [95]. Framework ini dirancang agar bersifat umum namun tetap fleksibel, sehingga dapat diadaptasikan di berbagai domain, termasuk dalam bidang medis. Pada penelitian deteksi aritmia, penerapan CRISP-DM berperan penting dalam menjaga konsistensi proses analisis mulai dari tahap pemahaman masalah, pengumpulan dan pemrosesan data, hingga evaluasi dan implementasi hasil [81]. Dengan demikian, setiap langkah yang dilakukan tidak hanya menghasilkan model dengan performa tinggi, tetapi juga keluaran yang memiliki relevansi dan interpretabilitas klinis. Kerangka CRISP-DM terdiri dari enam tahapan utama, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. sehingga hasil dari satu tahap dapat memengaruhi atau memperbaiki tahap lainnya [95]. Dalam konteks penelitian

ini, CRISP-DM menjadi landasan metodologis untuk mengintegrasikan data sinyal ECG dan teks medis ke dalam satu sistem analisis yang komprehensif dan berorientasi pada pemecahan masalah klinis. Keenam tahapan ini saling terhubung secara iteratif. Tahapan-tahapan utama dalam kerangka CRISP-DM digambarkan pada Gambar 2.14.



Gambar 2.14 Framework CRISP - DM

Sumber : [96]

Berdasarkan Gambar 2.14, framework CRISP-DM terdiri dari enam tahapan utama yang saling berkaitan dan bersifat iteratif, artinya setiap tahap dapat kembali ke tahap sebelumnya untuk melakukan penyesuaian atau perbaikan. Keenam tahapan tersebut membentuk siklus yang komprehensif untuk memastikan hasil analisis data selaras dengan kebutuhan dan tujuan penelitian. Tahapan-tahapan utama dalam framework CRISP-DM dijelaskan sebagai berikut [96]:

1. **Business Understanding**

Tahapan ini berfokus pada pemahaman mendalam terhadap tujuan utama penelitian atau kebutuhan klinis yang ingin dicapai. Dalam konteks deteksi aritmia, tahap ini mencakup identifikasi permasalahan seperti kebutuhan untuk mendeteksi kelainan irama jantung lebih dini atau meningkatkan

efisiensi analisis sinyal ECG oleh tenaga medis. Selain itu, peneliti juga menentukan tujuan analisis yang spesifik serta indikator keberhasilan yang akan digunakan untuk mengukur performa model, seperti tingkat akurasi dan interpretabilitas hasil deteksi.

2. ***Data Understanding***

Setelah tujuan penelitian dirumuskan, tahap berikutnya berfokus pada pemahaman dan eksplorasi awal terhadap data yang dikumpulkan. Data dapat berasal dari berbagai sumber, seperti rekaman sinyal ECG, laporan medis berbahasa Jerman, dan metadata pasien. Analisis deskriptif dilakukan untuk memahami karakteristik utama dataset, pola distribusi data, serta mengidentifikasi potensi kendala seperti *missing values*, *outliers*, atau ketidakkonsistenan antar sumber data. Langkah ini penting untuk memastikan bahwa data yang digunakan representatif dan siap diproses lebih lanjut.

3. ***Data Preparation***

Pada tahap ini, data dipersiapkan untuk digunakan dalam proses pemodelan. Aktivitas yang dilakukan meliputi *data cleaning* untuk menghapus duplikasi dan memperbaiki kesalahan, *data transformation* untuk menyamakan skala dan format data, serta *feature selection* untuk memilih atribut paling berpengaruh terhadap hasil analisis. Dalam penelitian berbasis sinyal ECG, proses ini mencakup segmentasi sinyal, normalisasi amplitudo, dan pembuatan vektor fitur.

4. ***Modeling***

Tahapan inti ini mencakup pembangunan dan pelatihan model menggunakan algoritma *deep learning*. Dalam penelitian ini, digunakan arsitektur *ResNet–Bidirectional LSTM–Attention* untuk pemrosesan sinyal ECG dan *RoBERTa* untuk pemrosesan teks medis. Proses ini juga mencakup *hyperparameter tuning* guna memperoleh kombinasi parameter yang memberikan hasil terbaik.

5. ***Evaluation***

Setelah model dibangun, tahap berikutnya adalah evaluasi untuk menilai performa dan validitasnya. Evaluasi dilakukan dengan menggunakan metrik

seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Dalam penelitian medis, tahap ini juga mencakup validasi klinis, yaitu memastikan bahwa hasil model dapat diinterpretasikan dan digunakan oleh tenaga kesehatan.

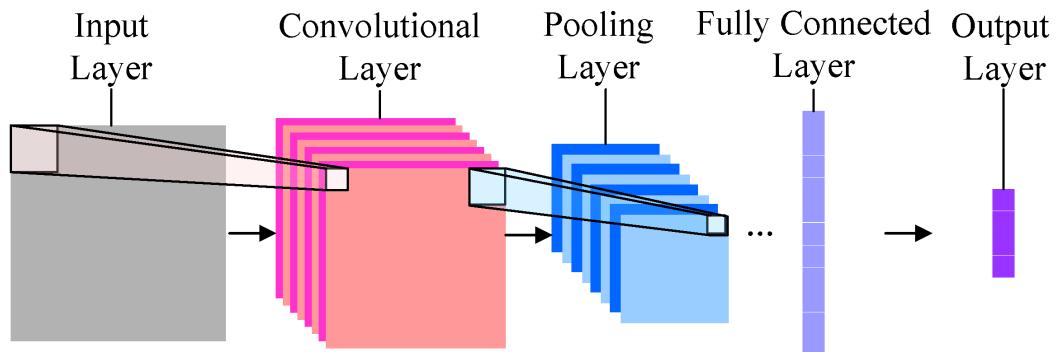
6. *Deployment*

Tahapan akhir adalah penerapan hasil analisis ke dalam sistem operasional atau aplikasi nyata. Dalam konteks *healthcare analytics*, hasil dari model dapat diintegrasikan ke dalam sistem monitoring pasien, untuk mendeteksi aritmia secara real-time.

Setiap tahapan dalam framework CRISP-DM memiliki keterkaitan erat satu sama lain dan berperan penting dalam menghasilkan pengetahuan yang valid, relevan, serta dapat diterapkan dalam konteks medis. Proses ini memastikan bahwa data mentah tidak hanya diolah menjadi model statistik, tetapi benar-benar diterjemahkan menjadi wawasan klinis yang bermanfaat [81]. Dalam penelitian klasifikasi penyakit aritmia, framework ini menjadi fondasi penting untuk mengintegrasikan analisis sinyal fisiologis dan teks medis secara sistematis,

2.3.2 *Convolutional Neural Network*

Convolutional Neural Network (CNN) merupakan arsitektur *deep learning* yang paling banyak digunakan dalam analisis data berbasis gambar dan sinyal, termasuk dalam bidang kesehatan seperti klasifikasi sinyal elektrokardiogram (ECG) untuk deteksi aritmia [97]. CNN mampu mengekstraksi fitur secara otomatis melalui lapisan konvolusi yang memproses data dalam bentuk spasial maupun temporal [98]. Hal ini menunjukkan potensi CNN dalam meningkatkan kecepatan dan ketepatan proses analisis sinyal medis. Arsitektur ini umumnya terdiri dari beberapa lapisan utama seperti *convolutional layer*, *pooling layer*, dan *fully connected layer* yang bekerja secara berurutan [99]. Setiap lapisan memiliki peran penting dalam mengekstraksi, mereduksi, dan mengklasifikasikan fitur dari data masukan. Dengan struktur berlapis ini, CNN mampu mempelajari representasi kompleks dari sinyal ECG secara efisien. Gambar 2.15 terkait arsitektur CNN.



Gambar 2.15 Arsitektur CNN

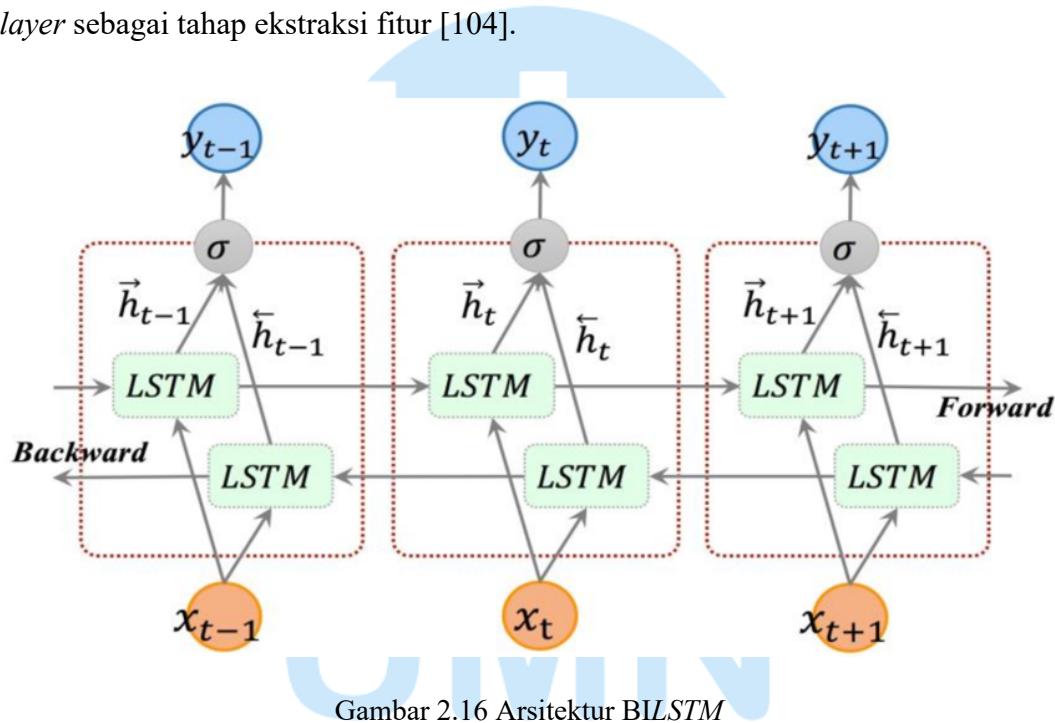
sumber : [100]

Berdasarkan Gambar 2.15 Arsitektur CNN terdiri dari beberapa lapisan utama, yaitu *input layer*, *convolutional layer*, *activation function (ReLU)*, *pooling layer*, *fully connected layer*, dan *output layer* [89]. Lapisan konvolusi berfungsi mengekstraksi fitur penting dari data menggunakan filter atau kernel. Selanjutnya, lapisan aktivasi ReLU mengintroduksi non-linearitas agar jaringan dapat mempelajari hubungan kompleks dalam data. *Pooling layer*, seperti *max pooling* [101], digunakan untuk mengurangi dimensi data sehingga mempercepat proses pelatihan dan mengurangi risiko overfitting. Lapisan *fully connected* menggabungkan semua fitur yang telah diekstraksi untuk menghasilkan prediksi akhir [99].

2.3.3 Bidirectional Long Short-Term Memory

Bidirectional Long Short-Term Memory (BiLSTM) merupakan pengembangan dari arsitektur *Long Short-Term Memory (LSTM)* yang dirancang untuk memproses data sekuensial dengan mempertimbangkan informasi dari dua arah, yaitu maju (*forward*) dan mundur (*backward*) [102]. Pendekatan dua arah ini memungkinkan model memahami konteks temporal secara lebih menyeluruh, karena setiap titik waktu dipengaruhi oleh urutan data sebelumnya dan sesudahnya. *BiLSTM* banyak digunakan dalam analisis sinyal *time-series*, termasuk sinyal elektrokardiogram (*ECG*), karena mampu mengenali pola ritme jantung yang kompleks dan bergantung pada konteks waktu. Gambar 2.16 menampilkan arsitektur umum dari *BiLSTM*. Dalam implementasinya pada sinyal *ECG*, *BiLSTM*

terbukti mampu meningkatkan performa klasifikasi dibandingkan model *LSTM* konvensional karena dapat menangkap hubungan temporal dari kedua arah sinyal. Hal ini sangat penting pada data medis yang bersifat dinamis, di mana setiap segmen sinyal memiliki keterkaitan dengan pola sebelumnya maupun sesudahnya [103]. Sejumlah penelitian juga menunjukkan bahwa penggunaan *BiLSTM* dapat meningkatkan akurasi deteksi aritmia hingga 3–7% dibandingkan model satu arah, terutama ketika dikombinasikan dengan mekanisme *attention* atau *convolutional layer* sebagai tahap ekstraksi fitur [104].



Gambar 2.16 Arsitektur BiLSTM

sumber : [105]

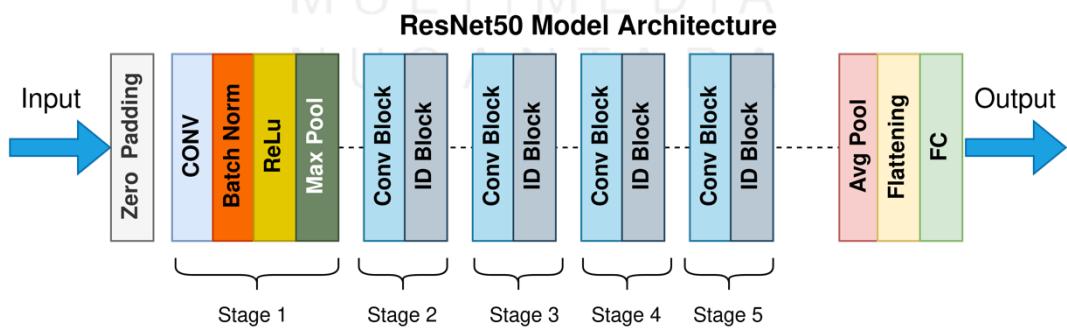
Arsitektur *BiLSTM*, seperti terlihat pada Gambar 2.16, terdiri dari dua lapisan *LSTM* yang berjalan secara paralel dalam arah berlawanan. Setiap unit *LSTM* di dalamnya memiliki tiga gerbang utama, yaitu *forget gate*, *input gate*, dan *output gate*, yang berfungsi untuk mengontrol aliran informasi di dalam sel memori [106]. *Forget gate* menentukan informasi mana yang perlu dihapus, *input gate* mengatur informasi baru yang disimpan, sedangkan *output gate* mengendalikan informasi yang diteruskan ke *hidden state*. Melalui kombinasi dua arah ini, *BiLSTM* mampu menangkap dependensi temporal jangka panjang maupun pendek secara lebih

akurat, menjadikannya sangat efektif untuk analisis sinyal *ECG* dalam mendeteksi pola aritmia yang halus dan dinamis [107].

2.3.4 Residual Network (RESNET)

Residual Network (ResNet) merupakan salah satu arsitektur *deep learning* yang dikembangkan untuk mengatasi masalah *vanishing gradient* pada jaringan saraf dalam (*deep neural networks*) [108]. Model ini dan menjadi tonggak penting dalam pengembangan *convolutional neural networks (CNN)* modern. ResNet menggunakan konsep *shortcut connection* atau *skip connection*, yaitu jalur langsung yang melewati satu atau beberapa lapisan, memungkinkan informasi dan gradien mengalir tanpa degradasi selama proses pelatihan. Pendekatan ini memungkinkan model untuk membangun jaringan yang sangat dalam tanpa mengalami penurunan performa akibat hilangnya informasi pada lapisan-lapisan terdalam [94].

Dalam konteks analisis sinyal elektrokardiogram (*ECG*), *ResNet* berperan penting dalam mengekstraksi fitur spasial yang kompleks dari bentuk gelombang sinyal. Setiap blok residual dalam *ResNet* bertugas mengenali pola morfologi jantung seperti gelombang P, kompleks QRS, dan gelombang T, yang menjadi indikator penting dalam diagnosis aritmia [109]. Dengan adanya koneksi residual, model dapat mempelajari representasi fitur yang lebih stabil dan efisien, terutama pada data dengan noise atau variasi antarindividu. Selain itu, arsitektur *ResNet* pada Gambar 2.17 terbukti meningkatkan konvergensi pelatihan dan ketahanan model terhadap overfitting.



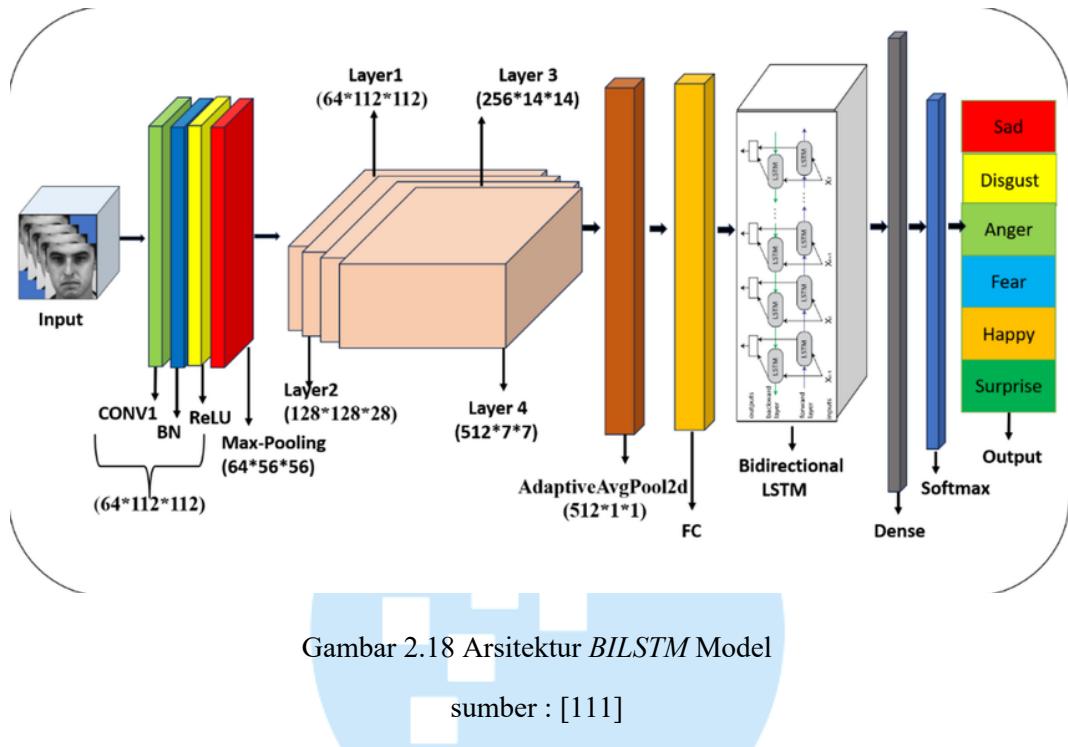
Gambar 2.17 ResNet Model Architecture

Sumber : [110]

ResNet terdiri beberapa *residual blocks* yang masing-masing memiliki dua hingga tiga lapisan konvolusi, diikuti oleh operasi normalisasi dan aktivasi (*batch normalization* dan *ReLU*) [101]. Hasil keluaran dari blok tersebut dijumlahkan dengan input awal melalui *shortcut connection*, seperti yang ditunjukkan pada Gambar 2.17. Mekanisme ini memungkinkan jaringan untuk belajar fungsi residu, bukan fungsi transformasi langsung, sehingga proses pembelajaran menjadi lebih efisien. Dalam penelitian deteksi aritmia, *ResNet* sering dikombinasikan dengan model sekuensial seperti *BiLSTM* dan mekanisme *attention* untuk menangkap baik fitur spasial maupun temporal dari sinyal ECG secara lebih menyeluruh [20].

2.3.5 *ResNet–Bidirectional LSTM–Attention*

ResNet–Bidirectional LSTM–Attention merupakan model hybrid yang menggabungkan kekuatan *Residual Network (ResNet)*, *Bidirectional Long Short-Term Memory (BiLSTM)*, dan mekanisme *Attention* dalam satu arsitektur yang terintegrasi dan berurutan [27]. Model ini dirancang untuk menangani data spasial dan temporal secara bersamaan, menjadikannya sangat efektif dalam tugas deteksi aritmia berbasis sinyal *Electrocardiogram (ECG)*. *ResNet* digunakan untuk ekstraksi fitur spasial yang kompleks dari sinyal, *BiLSTM* untuk menangkap dependensi temporal dua arah, sedangkan *Attention mechanism* berfungsi memberikan fokus pada bagian sinyal yang paling relevan terhadap proses klasifikasi. Kombinasi tiga komponen pada Gambar 2.18 memungkinkan model memahami struktur sinyal jantung secara menyeluruh, meningkatkan akurasi, dan memperkuat interpretabilitas hasil prediksi dalam konteks klinis. Secara konseptual, arsitektur hybrid *ResNet–BiLSTM–Attention* mengatasi keterbatasan model tunggal dalam menganalisis sinyal ECG yang bersifat nonlinier, nonstasioner, dan memiliki variasi temporal yang tinggi. *ResNet* dengan mekanisme skip connection mampu menjaga aliran gradien dan mengekstraksi fitur tingkat tinggi tanpa degradasi performa pada jaringan yang dalam, sementara *BiLSTM* memberikan kemampuan pemodelan dinamika waktu secara dua arah sehingga pola jangka pendek maupun jangka panjang pada sinyal jantung dapat ditangkap secara optimal.



Model *ResNet BiLSTM Attention* pada Gambar 2.18 bekerja melalui tiga tahap utama yang saling terhubung. Tahap pertama adalah *feature extraction* menggunakan *ResNet*, yang berperan mengekstraksi pola spasial dari sinyal ECG dengan memanfaatkan *residual connections* untuk mengatasi masalah *vanishing gradient* serta menjaga efisiensi pelatihan. Lapisan konvolusi pada *ResNet* mendeteksi pola penting seperti kompleks QRS, gelombang P, dan gelombang T, sementara blok residual memastikan bahwa informasi dari lapisan awal tetap terhubung dengan lapisan yang lebih dalam [94]. Hasil ekstraksi ini berupa *feature map* yang kemudian diratakan menjadi vektor fitur dan diteruskan ke tahap berikutnya, yaitu *temporal modeling* menggunakan *BiLSTM* [112]. Pada tahap ini, *BiLSTM* memproses sinyal dari dua arah (maju dan mundur), memungkinkan model memahami konteks temporal penuh dari urutan sinyal jantung, baik sebelum maupun sesudah setiap titik waktu. Mekanisme pada Rumus 2.1 penting untuk mendeteksi pola irama jantung yang kompleks, seperti fibrilasi atrium (AFib) [113]. Pendekatan pemodelan dua arah ini terbukti meningkatkan sensitivitas deteksi aritmia episodik, karena perubahan temporal yang bersifat halus dan tidak beraturan dapat dikenali secara lebih akurat [114].

Algorithm ResNet–Bidirectional LSTM–Attention

*Initialize model parameters θ (weights and biases)
Set learning_rate, epochs, batch_size*

For epoch in range(1, epochs + 1):

For each ECG segment d_seg in dataset D :

ResNet – based feature extraction
$F1 = Conv1D(d_seg, kernel_size = 3, stride = 1) + Bias1$
$F1 = BatchNormalization(F1)$
$F1 = ReLU(F1)$
$F1 = Conv1D(F1, kernel_size = 3, stride = 1) + Bias2$
$F1 = BatchNormalization(F1)$
$Skip1 = Add(d_seg, F1)$
$F_res1 = ReLU(Skip1)$

$F2 = Conv1D(F_res1, kernel_size = 3, stride = 1) + Bias3$
$F2 = BatchNormalization(F2)$
$F2 = ReLU(F2)$
$F2 = Conv1D(F2, kernel_size = 3, stride = 1) + Bias4$
$F2 = BatchNormalization(F2)$
$Skip2 = Add(F_res1, F2)$
$F_res2 = ReLU(Skip2)$

$F_pool = MaxPooling(F_res2)$

EndFor

$T_forward = LSTM(F_pool, direction = "forward")$
$T_backward = LSTM(F_pool, direction = "backward")$
$T_concat = Concatenate(T_forward, T_backward)$

*$Attention_weights = Softmax(W_a * T_concat)$*
*$Context_vector = Sum(Attention_weights * T_concat)$*

$FC_output = Dense(Context_vector)$
$Y_hat = Softmax(FC_output)$

$Loss = (Y_hat - Y_true)^2$
$Gradients = compute_gradient(Loss, \theta)$
*$\theta = \theta - learning_rate * Gradients$*

EndFor

$Y_pred = argmax(Y_hat)$

Return Y_pred

Rumus 2.1 Pseudocode dalam ResNet BI LSTM Attention

sumber : [114]

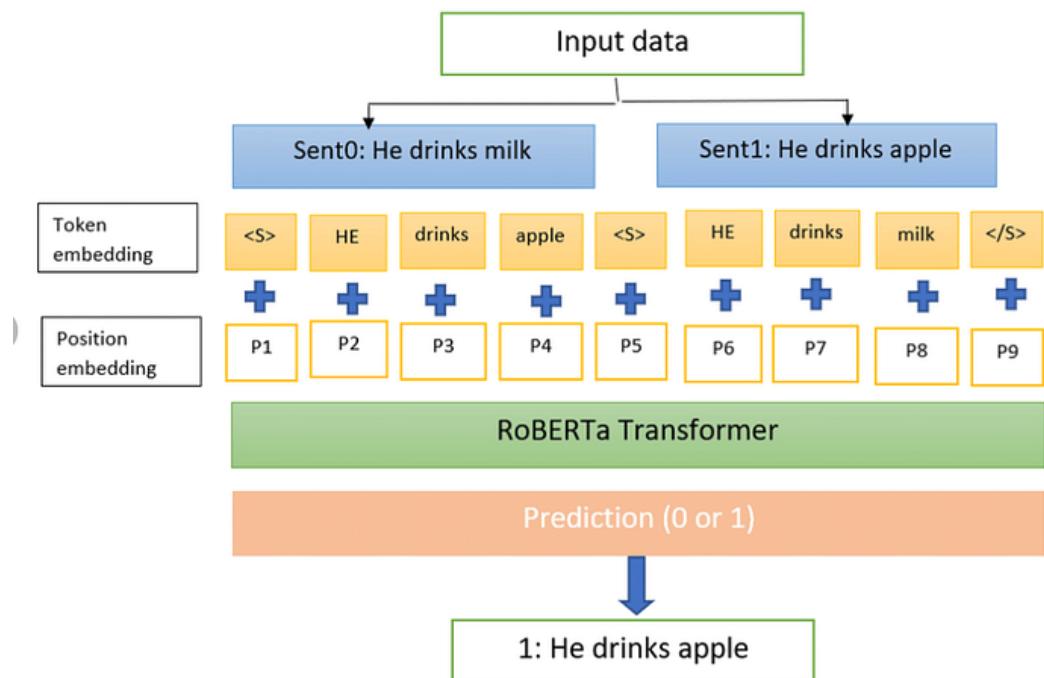
Berdasarkan Rumus 2.1 *Pseudocode* di atas menggambarkan alur pemrosesan sinyal elektrokardiogram (ECG) secara end-to-end menggunakan kombinasi tiga arsitektur deep learning. Tahap awal dilakukan oleh *ResNet*, yang mengekstraksi fitur spasial melalui operasi convolution, *batch normalization*, dan mekanisme skip connection untuk mempertahankan informasi penting dari lapisan sebelumnya serta mengatasi masalah vanishing gradient [115]. Hasil ekstraksi fitur kemudian diteruskan ke *Bidirectional LSTM (BiLSTM)*, yang mempelajari dependensi temporal sinyal dari dua arah maju dan mundur sehingga model memahami konteks waktu secara lebih menyeluruh [106]. Selanjutnya, *attention mechanism* digunakan untuk menimbang pentingnya setiap langkah waktu dalam urutan sinyal, dengan menghasilkan *context vector* yang menyoroti bagian paling relevan bagi klasifikasi. Output dari attention kemudian diteruskan ke lapisan *fully connected* dan diakhiri dengan fungsi aktivasi softmax untuk menghasilkan probabilitas kelas, seperti jenis aritmia tertentu [116]. Proses pelatihan dilakukan dengan menghitung fungsi loss dan memperbarui bobot menggunakan *backpropagation*, hingga model mencapai konvergensi dan mampu memberikan prediksi akhir dengan akurasi tinggi. Proses pembelajaran dilakukan melalui *backpropagation*, di mana bobot diperbarui berdasarkan *gradient loss* hingga model mencapai konvergensi dan menghasilkan prediksi akhir (Y_{pred}) dengan akurasi optimal khususnya dalam klasifikasi penyakit Aritmia [117].

2.3.6 *Robustly Optimized BERT Pretraining Approach*

RoBERTa(A Robustly Optimized BERT Pretraining Approach) merupakan model bahasa berbasis *Transformer* yang dikembangkan oleh Facebook AI sebagai pengembangan dari *BERT* dengan melakukan optimalisasi pada tahap *pretraining*. Studi literatur menunjukkan bahwa *RoBERTa* dilatih menggunakan jumlah data yang jauh lebih besar, dengan durasi pelatihan lebih lama, ukuran *batch* lebih besar, serta penghapusan *next sentence prediction (NSP)* yang ada pada *BERT* [29]. Pendekatan ini membuat *RoBERTa* mampu menghasilkan representasi teks yang lebih kaya, stabil, dan akurat pada berbagai tugas analisis bahasa alami. Dalam konteks penelitian ini, *RoBERTa* digunakan untuk menganalisis laporan medis berbasis teks dengan tujuan memahami pola semantik yang berhubungan dengan

kondisi aritmia, seperti deskripsi klinis, hasil diagnosis, atau catatan medis pasien. Model ini membantu mengidentifikasi kata atau frasa kunci yang relevan dengan kondisi jantung secara kontekstual, sehingga memperkuat interpretasi hasil klasifikasi berbasis sinyal.

Berbagai penelitian membuktikan bahwa *RoBERTa* mampu mengungguli *BERT* dalam hal akurasi karena model ini lebih sensitif terhadap konteks semantik dan sintaksis, sehingga sangat efektif digunakan dalam studi yang membutuhkan pemahaman bahasa alami secara mendalam, termasuk analisis teks medis, deteksi emosi, hingga klasifikasi dokumen ilmiah. Selain itu, fleksibilitas arsitektur *RoBERTa* memungkinkan integrasi dengan model lain dalam pendekatan multimodal, seperti penggabungan dengan *ResNet* atau *CNN* untuk data sekuensial non-teks, sehingga dapat meningkatkan performa dalam pemrosesan data kompleks [118].. Arsitektur *RoBERTa* pada gambar 2.19 berupa *encoder* bertingkat yang memproses input teks melalui mekanisme *self-attention* untuk menangkap hubungan antar kata dalam konteks yang luas [119].



Gambar 2.19 Arsitektur *RoBERTa*

Sumber : [29]

Gambar 2.19 menggambarkan arsitektur *RoBERTa* yang diperluas untuk klasifikasi dengan integrasi metadata. Proses dimulai dari input teks yang diproses oleh *RoBERTaFastTokenizer* lalu melewati beberapa transformer block guna menghasilkan representasi konteks. Lapisan *Citation Re-insertion* digunakan untuk memperkaya representasi dengan informasi tambahan sebelum menghasilkan vektor keluaran C [120]. Vektor ini kemudian digabungkan (*concatenation*) dengan fitur metadata eksternal untuk memperkuat konteks. Hasil gabungan diproses melalui beberapa lapisan tambahan, yaitu Dropout untuk regularisasi, Linear untuk transformasi dimensi, *Tanh* sebagai fungsi aktivasi non-linear, lalu kembali ke *Linear* dan *Softmax* untuk menghasilkan probabilitas prediksi kelas akhir P [121]. Arsitektur ini memungkinkan model tidak hanya memahami konteks linguistik dari teks, tetapi juga memanfaatkan informasi tambahan dari metadata sehingga meningkatkan performa klasifikasi. *Pseudocode* disajikan pada Rumus 2.2 dalam algoritma *RoBERTa*

Algorithm *RoBERTa* for classification

Input: Medical report (T), Metadata features (M)
Output: Predicted arrhythmia class (Y_pred)

Tahap 1: Preprocessing
tokens = RoBERTaFastTokenizer(T)
tokens = AddSpecialTokens(tokens) # [CLS], [SEP]

Tahap 2: Ekstraksi Representasi Kontekstual
H = RoBERTa(tokens) #Transformer Block
C = CitationReinsertion(H) #Domain knowledge

Tahap 3: Penggabungan dengan Metadata
ALL = Concatenate(C, M)

Tahap 4: Fully Connected Layers
Z1 = Dropout(ALL)
Z2 = Linear(Z1)
Z3 = Tanh(Z2)
Z4 = Dropout(Z3)
Z5 = Linear(Z4)

Tahap 5: Softmax untuk Klasifikasi
Ŷ = Softmax(Z5)

Tahap 6: Loss & Training

Algorithm RoBERTafor Classification

```
loss = CrossEntropy(Ŷ, Y_true)      # Y_true: label aritmia aktual
params_grad = Backprop(loss)
UpdateWeights(params_grad)

# Tahap 7: Prediksi Akhir
Y_pred = argmax(Ŷ)
Return Y_pred
```

Rumus 2.2 *Pseudocode* Terkait penggunaan metode RoBERTa

Sumber : [122]

Pseudocode *RoBERTa* pada Rumus 2.2 dalam klasifikasi *medical report* aritmia menjelaskan alur pemrosesan teks medis yang dimulai dari tahap *tokenisasi* menggunakan *RoBERTaTokenizer* dengan penambahan token khusus seperti *[CLS]* dan *[SEP]* untuk menandai awal dan akhir teks. Setiap token kemudian dikonversi menjadi *embedding* dan diproses melalui *transformer block* dari *RoBERTa*, yang berfungsi menangkap konteks semantik antar kata secara mendalam melalui mekanisme *self-attention* [123]. Dalam konteks penelitian ini, representasi kontekstual yang dihasilkan digunakan untuk memahami deskripsi klinis pasien, seperti diagnosis, hasil pemeriksaan, atau catatan dokter yang berkaitan dengan kondisi aritmia. Hasil representasi ini kemudian dapat dikombinasikan dengan informasi non-teks, seperti parameter sinyal *ECG* atau metadata pasien, guna memperkaya konteks interpretasi dan mendukung pendekatan multimodal. Selanjutnya, vektor hasil penggabungan diproses melalui beberapa lapisan *fully connected* dengan *dropout* sebagai teknik regularisasi untuk mencegah *overfitting*, serta menggunakan fungsi aktivasi *Tanh* untuk menjaga kestabilan representasi. Tahap akhir melibatkan lapisan *softmax* yang menghasilkan distribusi probabilitas untuk setiap kelas aritmia [124]. Selama proses pelatihan, model dioptimalkan menggunakan *cross-entropy loss* dan *backpropagation*, sedangkan pada tahap inferensi, prediksi akhir ditentukan berdasarkan nilai *argmax* dari distribusi probabilitas kelas [125]. Pendekatan ini memungkinkan *RoBERTa* tidak hanya memahami terminologi medis secara tekstual, tetapi juga menangkap hubungan kontekstual antarfrasa yang merepresentasikan kondisi klinis pasien secara lebih akurat dan bermakna.

2.3.7 *Embedding*

Embedding merupakan representasi vektor berdimensi tetap yang digunakan untuk memetakan data berdimensi tinggi, seperti teks dan sinyal biologis, ke dalam ruang laten yang lebih terstruktur sehingga lebih mudah diproses oleh model pembelajaran mendalam [116]. Pada data teks, *embedding* umumnya diperoleh melalui tokenisasi sub-kata dan transformasi oleh model *Transformer* seperti *RoBERTa* atau *XLM-RoBERTa*, yang menghasilkan representasi kontekstual seperti *[CLS] embedding* untuk keseluruhan kalimat [126]. Sementara itu, pada data sinyal seperti ECG, *embedding* dihasilkan melalui blok konvolusi atau LSTM yang merangkum karakteristik spasial dan temporal menjadi vektor fitur homogen. Fungsi utama *embedding* adalah memberikan representasi kompak yang mempertahankan informasi semantik atau morfologis penting, serta memungkinkan integrasi lintas-modal ketika digunakan dalam arsitektur multimodal [116]. Oleh karena itu, konsep *embedding* menjadi dasar dalam model hibrida yang menggabungkan sinyal fisiologis dan laporan klinis.

2.3.8 *MultiModal Fusion*

Multimodal fusion mengacu pada teknik integrasi dua atau lebih sumber data, seperti sinyal ECG dan teks laporan medis, untuk menghasilkan representasi komprehensif yang lebih informatif dibandingkan modalitas tunggal [127]. Terdapat tiga pendekatan utama fusion *early fusion*, yang menggabungkan fitur pada tahap awal sebelum pemodelan lanjutan *late fusion*, yang menggabungkan hasil prediksi masing-masing modalitas pada tahap akhir *hybrid fusion*, yang mengombinasikan kedua pendekatan tersebut dan banyak digunakan dalam penelitian modern karena mampu memanfaatkan kekuatan representasi lokal dan global secara bersamaan [128]. Selain itu, strategi *two-stage training* umum diterapkan dalam multimodal learning, yaitu tahap pertama dimana *backbone* model dibekukan (*freeze*) untuk melatih lapisan atas, dan tahap kedua di mana seluruh parameter model di-*unfreeze* untuk melakukan *fine-tuning* [15]. Pendekatan ini meningkatkan stabilitas pelatihan sambil menghindari *catastrophic forgetting*, terutama ketika modalitas memiliki kompleksitas yang berbeda.

2.3.9 *Fusion Head*

Fusion head adalah komponen akhir pada arsitektur multimodal yang dirancang untuk menggabungkan dua *embedding* dari modalitas berbeda menjadi representasi bersama sebelum proses klasifikasi [129]. Struktur *fusion head* umumnya terdiri atas beberapa lapisan *dense*, fungsi aktivasi non-linear, serta mekanisme *dropout* untuk mengurangi risiko *overfitting*. Tujuan utama *fusion head* adalah memastikan bahwa representasi teks dan sinyal dapat berinteraksi dalam ruang laten yang sama, sehingga model mampu mempelajari pola lintas-modal secara efektif [130]. Dalam konteks klasifikasi multi-label, *fusion head* juga bertanggung jawab membentuk representasi gabungan yang mampu menangkap hubungan antar-label pada kedua modalitas. Struktur dan kedalaman *fusion head* memiliki dampak signifikan terhadap kualitas integrasi fitur dan akurasi prediksi multimodal.

2.3.10 *Augmentation*

Augmentation merupakan teknik peningkatan variasi data melalui transformasi sintetis yang tetap mempertahankan karakteristik utama sinyal atau teks. Pada data ECG, teknik augmentasi dapat berupa *scaling*, *jittering*, *random cropping* [131], hingga penambahan noise dengan probabilitas tertentu untuk meningkatkan ketahanan model terhadap variasi sinyal nyata. Pada data teks, augmentasi umum dilakukan melalui *synonym replacement*, *random substitution*, atau *back-translation* dengan probabilitas tertentu untuk memperkaya kosakata dan mengurangi *overfitting* [132]. Implementasi augmentasi sangat penting dalam dataset medis yang cenderung terbatas dan tidak merata antar kelas, karena mampu meningkatkan generalisasi model sekaligus memperbaiki performa pada kelas minoritas.

2.3.11 *Weighted Focal Loss*

Weighted Focal Loss merupakan fungsi kerugian yang dirancang untuk menangani masalah ketidakseimbangan kelas dan memberikan fokus lebih pada sampel yang sulit diprediksi [133]. Fungsi ini merupakan perluasan dari *Focal Loss*

dengan menambahkan bobot kelas (*class weights*) yang dihitung berdasarkan distribusi label, misalnya menggunakan metode *Effective Number of Samples (ENS)* [134]. Parameter γ (*gamma*) berfungsi mengurangi kontribusi sampel yang mudah diprediksi, sedangkan α (*alpha*) menentukan tingkat perhatian pada kelas tertentu. Pendekatan ini sangat sesuai untuk klasifikasi aritmia karena jumlah kelas minoritas jauh lebih kecil dibandingkan kelas mayoritas [135]. Oleh karena itu, *Weighted Focal Loss* meningkatkan sensitivitas model terhadap sinyal dan laporan yang jarang muncul namun penting secara klinis.

2.3.12 Learning Rate Scheduler

Learning rate scheduler adalah mekanisme yang mengatur nilai *learning rate* selama proses pelatihan agar konvergensi model menjadi lebih stabil dan efisien. Dua pendekatan umum adalah *Polynomial Decay*, yang menurunkan *learning rate* secara non-linear berdasarkan progres epoch, dan *Cosine Decay*, yang mengikuti pola kurva kosinus untuk memberikan penurunan tajam pada awal pelatihan dan stabilitas pada tahap akhir [136]. Penggunaan scheduler diperlukan untuk menghindari kondisi pelatihan dengan *learning rate* terlalu besar yang menyebabkan divergensi, atau terlalu kecil yang menyebabkan pelatihan stagnan. Scheduler sangat relevan dalam model multimodal berdimensi besar seperti *RoBERTa* dan *ResNet*, karena membantu menjaga stabilitas *fine-tuning* [137].

2.3.13 Hyperparameter tuning

Hyperparameter tuning merupakan proses optimisasi parameter non-trainable, seperti *batch size*, jumlah *epoch*, *dropout rate*, ukuran embedding, jumlah blok konvolusi, unit LSTM, serta konfigurasi *learning rate*, untuk memperoleh performa terbaik pada model [138]. Proses tuning ini umumnya dilakukan melalui teknik seperti *grid search*, *random search*, atau *Bayesian optimization* [22]. Dalam konteks multimodal learning, tuning perlu dilakukan secara terpisah untuk cabang teks dan cabang sinyal karena masing-masing memiliki karakteristik berbeda. Pengaturan hiperparameter yang tepat sangat menentukan kemampuan model dalam menangkap pola relevan, menurunkan kesalahan pelatihan, serta meningkatkan generalisasi.

2.3.14 Lambda Layer

Lambda Layer merupakan komponen dalam *Keras* yang memungkinkan penerapan fungsi-fungsi custom atau transformasi tensor yang tidak tersedia dalam lapisan standar [19]. Lapisan ini sering digunakan untuk mengatasi perbedaan kompatibilitas, melakukan penyesuaian dimensi embedding, atau memodifikasi output model sebelum diberikan ke lapisan berikutnya. Dalam pemodelan multimodal, *Lambda Layer* digunakan untuk menyesuaikan output *RoBERTa* atau encoder lainnya agar selaras dengan struktur arsitektur keseluruhan, terutama ketika terjadi perubahan versi framework seperti *Keras 3* [139]. Penggunaan *Lambda Layer* memberikan fleksibilitas tinggi dalam pengembangan arsitektur kompleks tanpa memerlukan implementasi lapisan baru secara manual.

2.3.15 Evaluation Metric

Metrik evaluasi adalah komponen penting dalam menilai performa model pembelajaran mesin dengan pemilihan metrik yang tepat sangat bergantung pada jenis masalah yang dihadapi, seperti klasifikasi biner, multi-label, atau regresi [140]. Berikut metrik evaluasi utama yang digunakan dalam klasifikasi.

1. Accuracy

Accuracy adalah metrik evaluasi yang digunakan untuk mengukur seberapa baik model klasifikasi memprediksi dengan benar dibandingkan keseluruhan data dengan konteks klasifikasi biner maupun multi-label [75]. Rumus untuk menghitung *Accuracy* ditampilkan pada Rumus 2.3.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Rumus 2.3 *Accuracy* dalam Penelitian

Sumber : [111]

Rumus 2.3 menunjukkan bahwa *Accuracy* mengukur rasio total prediksi yang benar *True positives + True Negatives* terhadap keseluruhan jumlah data yang diuji. Metrik ini merepresentasikan seberapa sering model memberikan

prediksi yang benar tanpa memperhatikan distribusi kelas [76]. Dalam *dataset* seimbang, *Accuracy* memberikan evaluasi yang baik. Namun, untuk dataset dengan distribusi kelas yang tidak seimbang, model dengan performa rendah dapat memiliki *Accuracy* yang tinggi hanya karena bias pada kelas mayoritas.

2. *Precision*

Precision merupakan salah satu metrik evaluasi yang digunakan untuk mengukur kualitas prediksi dalam machine learning [78]. Metrik ini menggambarkan proporsi prediksi positif yang benar terhadap seluruh prediksi positif. Dengan *Precision*, kita dapat mengetahui seberapa akurat model dalam mengidentifikasi data positif [79]. *Precision* ditampilkan pada Rumus 2.4.

$$Precision = \frac{True\ Positive}{True\ positive + False\ Positive}$$

Rumus 2.4 *Precision* dalam Penelitian

Sumber : [97]

Rumus 2.4 menunjukkan bahwa *Precision* fokus pada pengurangan kesalahan prediksi *False Positives*. *Precision* tinggi berarti model memberikan prediksi positif yang menunjukkan bahwa model mampu membedakan dengan baik antara data positif dan negatif dalam klasifikasi [80]. Sebaliknya, *Precision* rendah mengindikasikan banyak prediksi positif yang salah, sehingga menurunkan kepercayaan terhadap hasil model tersebut [78]. s

3. *Recall (Sensitivity)*

Recall merupakan salah satu metrik evaluasi yang penting dalam machine learning, khususnya untuk menangani ketidakseimbangan data [81]. Metrik ini digunakan untuk mengukur seberapa baik model dalam mengenali data yang termasuk dalam kelas positif. *Recall* menunjukkan proporsi total kelas positif yang berhasil diidentifikasi dengan benar oleh model [82]. *Recall* ditampilkan pada Rumus 2.5. yang berfungsi untuk menangani ketidakseimbangan data.

$$Precision = \frac{True\ Positive}{True\ positive + False\ Negative}$$

Rumus 2.5 *Recall* dalam Penelitian

Sumber : [97]

Rumus 2.5 Menjelaskan *Recall* sebagai ukuran sensitivitas model terhadap Kelas Positif. *Recall* Tinggi menunjukkan Kemampuan model untuk menangkap hampir semua data kelas positif, meskipun terkadang disertai peningkatan *False Positives* [83]. Hal ini berarti model cenderung lebih fokus pada mendekripsi data positif, bahkan jika harus mengorbankan sebagian akurasi dengan menambahkan prediksi positif yang salah [84]. *Recall* sangat penting dalam situasi di mana kegagalan mendekripsi kelas positif dapat memiliki konsekuensi serius, seperti pada diagnosis penyakit atau deteksi penipuan [82].

4. *F1-Score*

F1-Score adalah metrik evaluasi yang digunakan untuk mengukur keseimbangan antara *Precision* dan *Recall* dalam sebuah model [85]. Metrik ini sangat berguna ketika terdapat trade-off antara keduanya, seperti saat model memiliki *Precision* tinggi tetapi *Recall* rendah, atau sebaliknya. Dengan menggunakan *F1-Score*, dapat mengevaluasi performa model secara lebih menyeluruh dan adil, terutama dalam kasus ketidakseimbangan data [86]. *F1-Score* ditampilkan pada Rumus 2.6.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Rumus 2.6 *F1 Score* dalam Penelitian

Sumber : [97]

Rumus 2.6, yaitu *F1-Score*, menggabungkan *precision* dan *recall* ke dalam satu metrik *F1-Score* relevan saat diperlukan keseimbangan antara

meminimalkan kesalahan positif palsu dan negatif palsu, terutama pada kasus dengan distribusi kelas tidak seimbang [87]. Dalam situasi diagnosis penyakit, keseimbangan antara meminimalkan kesalahan positif palsu dan negatif palsu sangat penting [88].

5. *Confusion Matrix*

Confusion matrix adalah metode evaluasi yang digunakan untuk menggambarkan kinerja model klasifikasi dengan cara membandingkan hasil prediksi model terhadap label sebenarnya. Matriks ini tidak hanya menunjukkan jumlah prediksi yang benar [107], tetapi juga memperlihatkan jenis kesalahan yang dilakukan model. Dengan demikian, confusion matrix menjadi alat penting dalam analisis performa, terutama pada kasus klasifikasi medis seperti deteksi aritmia, di mana kesalahan prediksi dapat berdampak serius pada diagnosis pasien. Bentuk umum *Confusion Matrix* ditampilkan pada Rumus 2.7.

$$\text{Confusion Matrix} = \begin{matrix} TP & FP \\ FN & TN \end{matrix}$$

Rumus 2.7 *Confusion Matrix* dalam Penelitian

Sumber : [97]

Berdasarkan Rumus 2.7 *True Positive* (TP) adalah jumlah kasus positif (aritmia) yang berhasil diprediksi dengan benar, sedangkan *True Negative* (TN) adalah jumlah kasus negatif (normal) yang diklasifikasikan dengan benar. *False Positive* (FP) menunjukkan jumlah kasus negatif yang salah diprediksi sebagai aritmia, sedangkan *False Negative* (FN) adalah jumlah kasus aritmia yang gagal terdeteksi karena salah diprediksi sebagai normal. Informasi dalam *confusion matrix* ini sangat penting karena menjadi dasar dalam menghitung berbagai metrik evaluasi, seperti akurasi, presisi, *Recall* (sensitivitas), *specificity*, dan *F1-score*.

6. ROC (*Receiver Operating Characteristic*)

ROC (*Receiver Operating Characteristic*) adalah metode evaluasi performa model klasifikasi biner yang menggambarkan hubungan antara *True Positive Rate (TPR / Sensitivitas / Recall)* dan *False Positive Rate (FPR)* pada berbagai nilai ambang (*threshold*) [141]. Kurva ROC dibentuk dengan memplot TPR pada sumbu Y dan FPR pada sumbu X, sehingga menunjukkan seberapa baik model dapat membedakan antara kelas positif (misalnya aritmia) dan kelas negatif (normal) [142]. Semakin mendekati sudut kiri atas grafik, semakin baik performa model karena TPR tinggi dan FPR rendah *True Positive Rate* dan *False Positive Rate* ditampilkan pada Rumus 2.8.

$$True\ positive\ Rate = \frac{TP}{TP + TN}$$

$$False\ positive\ Rate = \frac{FP}{FP + TN}$$

Rumus 2.8 ROC Curve dalam Penelitian

Sumber : [97]

Berdasarkan Rumus 2.8 *TPR* dihitung dengan membandingkan jumlah prediksi benar pada kelas positif (*TP*) terhadap seluruh kasus yang seharusnya positif, yaitu $(TP+FN)(TP + FN)(TP+FN)$, sehingga merepresentasikan kemampuan model dalam mendeteksi kasus positif secara benar. Sementara itu, *FPR* dihitung dari jumlah prediksi salah pada kelas negatif (*FP*) terhadap seluruh kasus yang sebenarnya negatif, yaitu $(FP+TN)(FP + TN)(FP+TN)$, sehingga menunjukkan seberapa besar kesalahan model dalam mengklasifikasikan data negatif sebagai positif [100].

7. AUC (Receiver Operating Characteristic)

Area Under Curve (AUC) adalah ukuran kuantitatif yang dihasilkan dari kurva *ROC* untuk menilai performa model klasifikasi secara keseluruhan [143]. *AUC* merepresentasikan luas area di bawah kurva *ROC*, dengan nilai berkisar antara 0 hingga 1. Semakin besar nilai *AUC* mendekati 1, semakin baik kemampuan model dalam membedakan kelas positif (misalnya aritmia) dan kelas negatif (normal), karena menunjukkan bahwa model mampu mencapai *True Positive Rate* (TPR) yang tinggi sekaligus menjaga *False Positive Rate* (FPR) tetap rendah pada berbagai nilai ambang (*threshold*). Rumus yang digunakan oleh *AUC* ditampilkan pada Rumus 2.9. Nilai *AUC* yang tinggi juga mencerminkan kestabilan model terhadap perubahan ambang keputusan, sehingga dapat digunakan sebagai indikator kinerja yang lebih andal dibandingkan hanya menggunakan akurasi.

$$AUC = \int_0^1 TPR(FPR) d(FPR)$$

Rumus 2.9 *AUC* dalam Penelitian

Sumber : [97]

Berdasarkan Rumus 2.9 *AUC* dihitung sebagai integral dari kurva *ROC*, yaitu dengan menjumlahkan nilai *True Positive Rate* (TPR) terhadap perubahan *False Positive Rate* (FPR) di sepanjang rentang 0 hingga 1. Secara matematis, hal ini berarti *AUC* mengukur luas area di bawah kurva *ROC* yang terbentuk dari plot TPR terhadap FPR pada berbagai nilai ambang (*threshold*) [144]. Semakin besar nilai area ini, semakin baik performa model karena menandakan model mampu mempertahankan TPR yang tinggi dengan FPR yang rendah. Dengan kata lain, *AUC* memberikan representasi kuantitatif atas probabilitas bahwa model akan memberikan skor prediksi lebih tinggi pada

sampel positif dibandingkan pada sampel negatif, sehingga menjadi indikator penting dalam mengevaluasi keandalan model klasifikasi.

2.4 Tools Penelitian

Penelitian ini digunakan perangkat bantu (*tools*) yang berperan penting dalam proses pengolahan data, analisis, serta pembangunan model. Tools yang dipakai adalah Jupyter Notebook dan Python. berikut penjelasan pada masing masing tools antara lain:

2.4.1 Jupyter Notebook

Jupyter Notebook adalah sebuah lingkungan pengembangan terintegrasi yang interaktif dan fleksibel untuk analisis data, pengembangan model, dan dokumentasi. *Jupyter Notebook* mendukung berbagai bahasa pemrograman melalui kernel, dengan Python sebagai salah satu bahasa yang paling umum digunakan [144]. Beberapa fitur utama *Jupyter Notebook* meliputi [78] :

1. *User interface* memungkinkan eksekusi kode secara langsung, disertai dengan output visual atau teks dalam satu sel.
2. *Multi language* mendukung lebih dari 40 kernel, termasuk Python, R, Julia, dan Scala, sehingga pengguna dapat bekerja dengan berbagai bahasa pemrograman.
3. Integrasi visualisasi Data Mendukung visualisasi data interaktif menggunakan pustaka seperti *Matplotlib*, *Seaborn*, dan *Plotly*.
4. Ekstensi dengan *Jupyter Notebook* dapat diperluas dengan ekstensi untuk menambahkan fungsionalitas baru, seperti penyorotan sintaks, diagram alur kerja, atau integrasi alat analitik lainnya.
5. Kolaborasi berupa *Notebooks* dapat dibagikan dalam format *file.ipynb*, memungkinkan kolaborasi yang mudah antarpeneliti atau pengembang.

Penerapan Jupyter Notebook dalam dunia kesehatan digunakan untuk analisis data medis seperti sinyal EKG atau data pasien, pengembangan model machine learning untuk prediksi penyakit (contohnya klasifikasi aritmia), visualisasi data klinis,

penelitian genomik untuk analisis DNA/RNA, dan pengembangan alat diagnostik berbasis kecerdasan buatan untuk mendukung pengambilan keputusan berbasis data di bidang kesehatan [145].

2.4.2 *Phyton*

Python adalah bahasa pemrograman tingkat tinggi yang bersifat general-purpose, mudah dipahami, dan mendukung paradigma pemrograman terstruktur, berorientasi objek, serta fungsional. Python terkenal karena sintaksnya yang sederhana dan pustaka yang kaya, sehingga menjadi pilihan utama dalam analisis data, pengembangan aplikasi, dan pembelajaran mesin *PHY*. Sebagai bahasa *open-source*, *Python* terus berkembang dengan sangat optimal khusunya untuk mendukung penelitian dan mendukung komunitas yang luas dan kompatibilitas dengan berbagai platform. Berikut Library populer dalam *Python* [145] :

1. *NumPy*: komputasi numerik dan operasi array
2. *Pandas*: manipulasi data dan analisis dataset.
3. *Matplotlib* dan *Seaborn*: visualisasi data dalam bentuk grafik dan diagram.
4. *Scikit-learn*: pembelajaran mesin dan analisis data prediktif.
5. *TensorFlow* dan *PyTorch*: pengembangan model deep learning.
6. *SciPy*: analisis ilmiah, termasuk pengolahan data statistik.

Penerapan Python dalam dunia kesehatan mencakup berbagai bidang, seperti pengolahan citra medis untuk mendeteksi penyakit menggunakan pustaka seperti OpenCV dan TensorFlow, analisis data pasien untuk menemukan pola atau anomali, pengembangan model prediktif untuk diagnosis penyakit seperti kanker atau aritmia jantung, hingga analisis genomik menggunakan alat seperti Biopython [146] Python juga digunakan untuk mengembangkan aplikasi kesehatan digital, seperti sistem manajemen rumah sakit, chatbot kesehatan berbasis *RoBERTa*, dan alat pemantauan kesehatan berbasis data real-time [147]. Fleksibilitas dan ekosistem pustaka yang luas menjadikan Python sebagai bahasa pemrograman yang sangat relevan untuk mendukung inovasi di bidang kesehatan modern.

Python juga berperan penting dalam pengembangan sistem decision support di bidang klinis, yang membantu tenaga medis dalam mengambil keputusan berbasis data [9]. Melalui integrasi dengan machine learning dan artificial intelligence, Python memungkinkan pembuatan sistem yang dapat memberikan rekomendasi pengobatan, prediksi risiko penyakit, serta identifikasi pasien dengan kondisi kritis secara lebih cepat dan akurat. Contohnya, algoritma klasifikasi berbasis *scikit-learn* atau *XGBoost* dan *RoBERT* dapat digunakan untuk memprediksi penyakit berdasarkan parameter klinis. Kemampuan Python dalam menangani data dalam jumlah besar, serta kemudahan integrasinya dengan database dan antarmuka pengguna, menjadikannya alat yang efisien dalam membangun solusi kesehatan berbasis teknologi. Selain itu, Python juga banyak dimanfaatkan dalam penelitian *multimodal learning* di bidang medis, di mana data teks, citra, dan sinyal fisiologis seperti ECG dapat digabungkan untuk menghasilkan sistem diagnosis yang lebih komprehensif dan berbasis konteks klinis.

